Title

Sensitivity of the response of Box-Girder Seat-type bridges to the duration of ground motions arising from crustal and subduction earthquakes
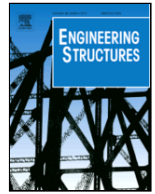
Authors

Fayaz, Jawad
Medalla, Miguel
Zareian, Farzin

Peer reviewed

Sensitivity of the response of Box-Girder Seat-type bridges to the duration of ground motions arising from crustal and subduction earthquakes

Jawad Fayaz [a], Miguel Medalla [a,b], Farzin Zareian [a,*]

[a] *Department of Civil and Environmental Engineering, University of California, Irvine, CA, USA*
[b] *Department of Structural and Geotechnical Engineering, Pontificia Universidad Catolica de Chile, Chile*

ARTICLE INFO

ABSTRACT

The design practice of Box-Girder Seat-Type (BGST) bridges in the Western U.S. is continuously evolving based on the results of advanced modeling and analysis techniques. This is mainly to help engineers and researchers to better understand the behavior of BGST bridges during seismic excitations. Within this backdrop, this study fills the gaps in the current knowledge of assessing the combined effect of strong motion duration and spectral shape on the response of bridges using a comprehensive set of numerical simulations and statistical analyses. Three-dimensional finite element models of two real BGST bridges are analyzed using a large set of ground motions obtained from crustal sources and subduction sources. By means of *Step-wise regression* – and other statistical procedures – the sensitivity of bridge response parameters to various ground motion parameters including Arias Intensity (I_a), *RotD50* spectral acceleration at the bridge's first natural period ($S_a\left(T_1\right)$), Significant Duration ($D_{5\text{-}95}$), mid-frequency (f), the derivative of the mid-frequency (f') and time at 30% of cumulated Arias Intensity (t_{mid}) are evaluated. Results indicate that in the case of ground motions arising from shallow crustal sources, I_a and $S_a\left(T_1\right)$ are the best predictors of the bridge response, and strong motion duration ($D_{5\text{-}95}$) has no statistically meaningful impact on the response of bridges. However, it is observed that the $D_{5\text{-}95}$ of the ground motions ascending from the subduction sources highly affects the bridge response; utilizing $D_{5\text{-}95}$ alongside $S_a\left(T_1\right)$, or I_a, can significantly increase the accuracy of bridge response estimates. Hence, it is concluded that $D_{5\text{-}95}$ is not an important ground motion intensity measure for ground motion selection for bridges located in areas with crustal earthquakes. In contrast, $D_{5\text{-}95}$ is important in subduction zone ground motions and must be given proper consideration in the design and analysis of BGST bridges.

1. Introduction

Highway bridges are among the vulnerable components of any urban infrastructure. They are expected to sustain minor damage and maintain their functionality after major man-made as well as natural disasters such as Earthquakes [11]. During the last three decades, bridges designed according to seismic design codes were observed to occasionally show poor performance during major earthquakes [7]. With the onset of sophisticated methods of analysis, the design and analysis codes for bridge construction (e.g., [11]) have evolved to look beyond single bridge behavior and aim at developing sustainable transportation network systems [1]. Caltrans SDC [11] recommends bridge design with a global target column ductility demand (μ_D) between 3 and 5; ductility is used as a surrogate for more direct performance measures. However, there are certain gaps in the current bridge analysis procedures. One of the major gaps present in the current methodology is the obliviousness of Intensity Measures (IMs) that carry more infor-

mation about the nature of impeding ground motions compared with the conventional pseudo-spectral acceleration towards quantifying the response of bridges. By this assertion, this study investigates the merits of adding strong motion duration, or Arias intensity, as a measure input energy, to the pseudo-spectral acceleration for estimating the response of bridge structures.

The current practices in bridge design are widely based on the recommendations of AASHTO Specifications (2011), which mainly incorporate four methods of analysis. These include: i) Equivalent Static Analysis (ESA), ii) Elastic Dynamic Analysis (EDA), iii) Inelastic Static Analysis (ISA), and iv) Nonlinear Time History Analysis (NLTHA). While the first three methods are appealing due to low computational overhead ranging from multimodal analysis to pushover, NLTHA requires that the bridge model must be analyzed under at least seven independent time histories applied in two orthogonal directions [12]. The peak response of bridge members for each set of time histories shall be recorded by applying the ground excitations in four orienta-

* Corresponding author.

tions (0, 30, 6030, 60, 90 degrees). For each orientation (and for each of the seven sets of time histories), the peak response at each pertinent Degree of Freedom (DOF) is recorded. This will result in (4 orientations) × (7 sets of times histories) = 28 peak responses at each pertinent DOF. The bridge is designed for the average of the recorded peak responses at each degree of freedom of interest. If fewer than seven sets of time histories are used, then the maximum rather than the average response shall be used for design. The selection of these ground motions is primarily based on a matching of ground motion spectra with the probabilistic Uniform Hazard Spectrum (UHS) of the location. Hence, it is deduced that these methods consider only the amplitude and frequency content of the ground motions, and the duration of ground motions is neglected.

Duration characteristic of a ground motion is generally identified in terms of its strong motion duration, also known as *Significant Duration* (D_s). D_s was initially measured using visual inspection (e.g., [40]). Current common ways to quantify D_s include (i) *Bracketed Duration* which is the time elapsed between the first and last excursions of the accelerogram above a specified acceleration threshold, e.g., 0.05 g [32,8]; (ii) the time interval over which a specific percentage range, e.g., 5–75%, 5–95% (denoted as D_{5-75}, and D_{5-95}, respectively), of the integral $\int_0^{t_{max}} a^2(t)\,dt$ is accumulated, where $a(t)$ represents the ground acceleration at time t, and t_{max} represents the length of the accelerogram [39]. While these metrics correspond to the canonical definition of ground motion duration, other duration metrics that are only implicitly correlated to the duration of strong ground motion, and may not even possess units of time, have also been proposed. $ID = \int_0^{t_{max}} a^2(t)\,dt/(PGA \times PGV)$, for example, is a dimensionless duration metric proposed by Cosenza and Manfredi [16]. Although the form and nature of a duration metric do influence its usability in structural design and assessment, a more important criterion is its correlation to key structural demands [13].

One of the earliest studies to employ nonlinear dynamic analyses to examine the effect of duration on structural response was done by Clough et al. [15]. This study analyzed a nonlinear model of a 20-story building and assessed the effect of duration on the peak lateral displacements, member ductility demands, and column axial forces by using sections of an accelerogram recorded from the 1940 El Centro earthquake, of different lengths, and scaled to different intensities. Several other studies, including Mahin [22], Bommer et al. [9], Ruiz-García and Miranda [30], Oyarzo-Vera and Chouw [27], and Mantawy and Anderson [23], have since employed nonlinear dynamic analyses to assess the influence of duration on structural demands. The conclusions reached by each of these studies were found to depend on (i) the nature of ground motions used; (ii) the metrics used to quantify duration; (iii) the characteristics of the structural models employed; and (iv) the structural demand parameters considered. The broad consensus that one may draw from this wide body of research is that ground motion duration generally does not affect peak demands, like peak story drifts and peak member forces, but significantly influences cumulative damage metrics (e.g., total dissipated hysteretic energy and accumulated plastic strain). Studies that simulated structural collapse also found that duration of ground motions affect collapse capacity or the capacity of a structure to resist structural collapse, as typically quantified by a collapse fragility curve [28]. Recently, Chandramohan [13] compared the structural response effect of duration and spectral shape of ground motions originating from crustal and subduction sources. This study analyzed building structures from three different locations in North America and concluded that the long duration and spectral shape of ground motions originating from subduction sources play a significant role in changing the response of the structures. However, all these studies were primarily based on analyzing the impact of strong mo-

tion duration on buildings or single bridge columns, and there is no exclusive literature available which has precisely investigated the effect of strong motion duration on the bridge response considering the complete dynamic response of the bridge model.

The study summarized herein, primarily investigates the impact of utilizing various IMs (features of the statistical relations), or their combination, on the response of two reinforced concrete (RC) Box-Girder Seat-Type (BGST) Bridges. These IMs include *RotD50* spectral acceleration at the bridge's natural period ($S_a(T_1)$), Arias Intensity (I_a) [2], Significant Duration (D_{5-95}), mid-frequency (f_{mid}) (Dabaghi and Der Kiureghian (2017) [17]), time at 30% of cumulated Arias Intensity (t_{mid}) [17] and the derivative of the mid-frequency (f') at t_{mid}. The bridges are analyzed under two groups of 500 ground motions, the first group arising from crustal sources and the second group from subduction sources. The analysis is conducted by rotating the two orthogonal components of the ground motions through 180 degrees in an interval of 9 degrees. The median of maximum column drift ratio (*Rot50CDR*) and dissipated energy (*Rot50DE*) in each direction is obtained and compared with the significant duration D_{5-95} of the ground motions. It should be noted that in this study the EDP of Column Drift Ratio (CDR) and Dissipated Energy (DE) is termed in the form of *RotppEDP*, where *Rot* indicates the rotation of ground motion components, *pp* indicates the percentile value used for the measure (e.g. "00", "50" and "100" correspond to minimum, median and maximum values, respectively; the median value is used in this study), and *EDP* indicates that the measure is an Engineering Demand Parameter (i.e., Column Drift Ratio *CDR* or Dissipated Energy *DE*). The motivation behind using such an EDP is to be consistent with the current state of the art Intensity Measure (IM), i.e., *RotD50* spectral acceleration; *RotDpp*, where *Rot* indicates the rotation of the two orthogonal components of the ground motions, *D* indicates the period dependency and *pp* corresponds to the percentile value (mainly limited to the 50 percentiles, i.e., the median value). *RotD50* is the median value of the spectral acceleration obtained when the two horizontal components are applied orthogonally and rotated throughout 180 degrees on a Single Degree of Freedom (SDOF) system with a varying spectral period [10]. Conclusively, in this research *RotD50* should not be confused with *Rot50EDP*, while the former is a measure of IM obtained after rotating the 2 ground motions components on an SDOF with varying period, latter is a measure of the EDP (Column Drift Ratio *CDR* and Dissipated Energy *DE*) obtained after rotating the two components of ground motions through 180 degrees on the MDOF bridge models. To avoid any confusion and for the sake of brevity, in this study, the *RotD50* spectral acceleration at bridge's first mode period is termed as $S_a(T_1)$ or S_a, S_a and $S_a(T_1)$ are used interchangeably in this paper.

Using regression tools and machine learning algorithms, the obtained *Rot50CDR*, and *Rot50DE* are used as the target variables for various predictors of ground motion and response parameters for the two groups of 500 ground motions. This is followed by a statistical sensitivity analysis wherein the predictive power of the different predictors (features, also dubbed as Intensity Measures) is estimated through *Step-wise regression*, *Random Forest*, and *Neighborhood Component Analysis* (*NCA*). Through the series of statistical analysis, the sensitivity of each predictor is observed in terms of statistical values (which include: R^2_{adj}, AIC, BIC, Relative Feature Importance *RI*, NCA Feature Weight *FW*; described later). Finally, conclusions are provided of the above-mentioned features that are statistically significant in predicting the response of the bridges (hence are related to the EDP) and whether the duration of the strong motions substantially affects the response of the bridges. It should be noted that based on the earlier studies [21,31] the vertical component of ground motions is considered to have a negligible effect on the displacement demands of bridge columns and hence, the vertical component of ground shaking is not considered in this study.

2. Bridge models

Two representative RC highway bridge structures are selected for the statistical analysis. Table 1 includes the details of the two ordinary bridges with seat-type abutments which reflect the common bridge engineering practice in the western United States. Finite Element models of the bridges were developed in *OpenSees* [24]. The finite element models are comprised of: seat-type abutments, shear keys, column bents, elastomeric bearing pads, backfill soil, and superstructure. An illustration of the model is provided in Fig. 1. The concrete and steel used in modeling possess a compressive strength $f_c' = 35$ MPa with

Table 1
Bridge Characteristics.

Bridge	A	B
Number of Spans	2	2
Span Lengths	33.6 m, 33.6 m	45.7 m, 45.7 m
Column Bent	Single-column	Two-column
Deck Width	8.4 m	23.0 m
Column Radius	0.84 m	0.85 m
Column Height	6.7 m	6.7 m
Reinforcement of Column Section	Long: 44 #11 (bundles of 2)	Long: 44 #11 (bundles of 2)
	$\rho_l = 2.00\%$	$\rho_l = 1.95\%$
	Trans: Spiral, #6 @ 8.5 cm	Trans: Spiral, #4 @ 15.2 cm
Site V_{s30}	222.95 m/sec	187.60 m/sec
Fundamental Period	0.61 sec	0.83 sec

a modulus of elasticity $E_s = 27789$ MPa and yield strength = 470 MPa with a modulus of elasticity $E_s = 200000$ MPa, respectively.

The models are based on the bridge models presented in Omrani et al. [26]; however, their structural component models are upgraded, and associated modeling parameters are updated. Caltrans SDC [11] recommends the superstructure be designed to remain elastic during earthquake events; therefore, the superstructure is modeled with elasticBeamColumn using uncracked section properties. To capture the dynamic response accurately, the mass of the superstructure is distributed throughout the length of the deck with each span's mass being distributed in ten intervals. The bridge columns are modeled using beamWithHinges element (two Gauss integration points) with fiber-discretized cross sections to model confined concrete for the core, unconfined concrete for the cover and steel rebars. The plasticity of columns is concentrated at two plastic hinges at the opposite ends connected by a linear elastic element. The plastic hinge length is determined based on Caltrans SDC [11]. Assuming a monolithic construction of cap beam and columns, the cap beam is modeled as a rigid bent using elasticBeamColumn element with high torsional, in-plane, and out-of-plane stiffnesses. The concrete and steel are modeled using Concrete01 and ReinforcingSteel materials, respectively, which are available in *OpenSees*.

The base of Bridge A and Bridge B are simulated as fixed and pinned connections, respectively, with the stiffness of connections arising from piles beneath. However, the flexibility of base connections arising from piles beneath is specifically modeled. To describe the translational behavior of pile foundations, a linear elastic spring model is used, and the vertical movement is restricted. In most practical situations, the response of laterally loaded piles is only controlled by the uppermost part of the pile length, l_c [14]. The length, l_c, is typically on the order of 5 to 10 pile diameters and is a function of the material
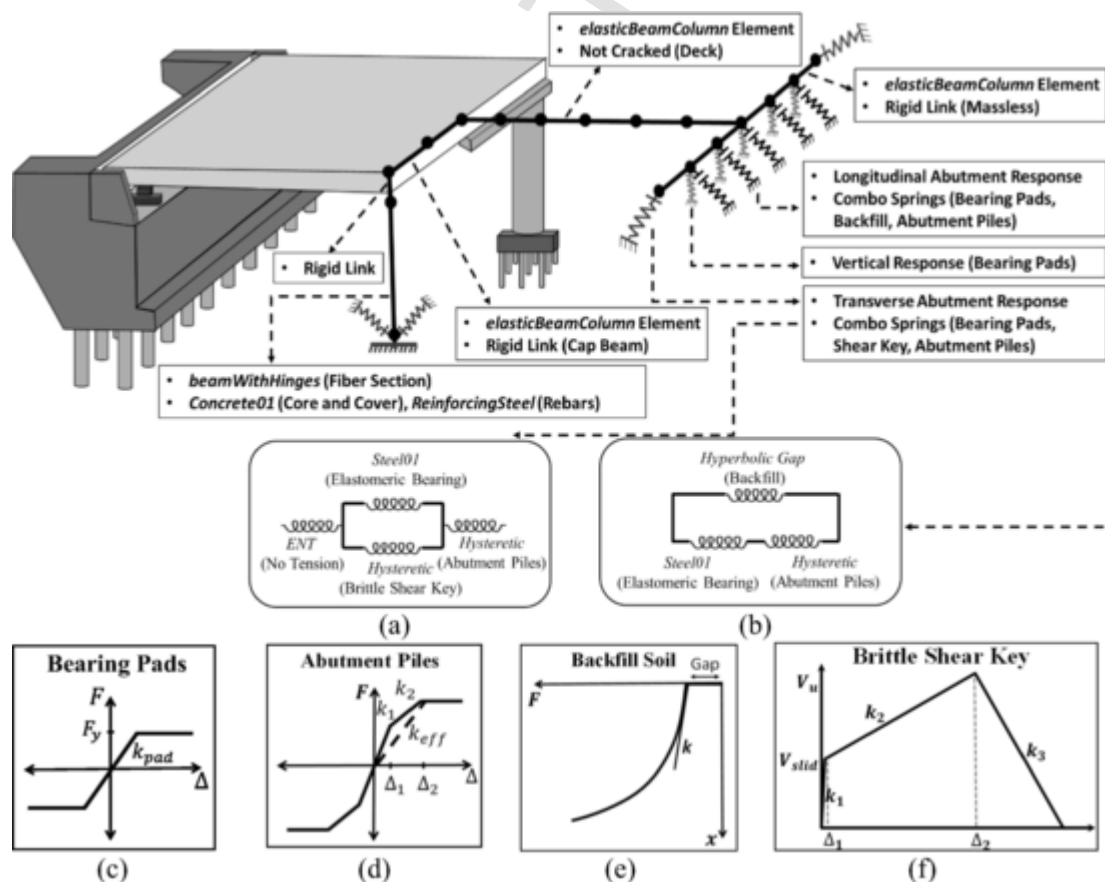

Fig. 1. Details of the finite-element model of bridge structures: a) Transverse abutment combo spring, b) Longitudinal abutment combo spring, c) Bearing pads response, d) Abutment pile response, e) Backfill soil response, and f) Brittle shear key response.

of the pile with respect to the soil. Eq. (1) presents simple algebraic expressions given by Fang [18], for estimating l_c of a circular solid pile with diameter, d, Young's modulus, E_p, for parabolic increases of soil modulus, $\hat{E}_s$, is the reference Young's modulus of the soil at depth, $z = d$, and d is the diameter of a pile. Eq. (2) is then used to calculate the total horizontal stiffness of pile footing ($K_{h,total}$). In Eq. (2), N represents the number of piles in the footing and the term inside brackets represent the stiffness of single pile and are valid for piles with length; $L>l_c$. In this study, the reference modulus of sand, 35 MPa is used for the conservatism and piles are assumed to be made of concrete ($E_p = 21994$ MPa).

$$l_c = 2d\left(E_p/\hat{E}_s\right)^{0.22} \tag{1}$$

$$K_{h,total} = N\left[0.8d\hat{E}_s\left(E_p/\hat{E}_s\right)^{0.28}\right] \tag{2}$$

Shear keys are designed and modeled in a brittle/isolated manner using the hysteretic spring model available in *OpenSees*. The model is defined with a trilinear backbone curve as given in Fig. 1f. The abutments are assumed as high abutments and the isolated shear keys are designed as per Caltrans SDC (2013) with the area of vertical reinforcement (A_{vsk}) calculated as per Eq. (3), with P_{dl}^{sup} is the superstructure dead load reaction at the abutment and f_{ye} is the yield strength of steel rebars. To be conservative, $\alpha = 1$ is used in Eq. (3).

$$A_{vsk} = \frac{\alpha \times P_{dl}^{sup}}{1.8 \times f_{ye}} \quad 0.5<\alpha<1 \tag{3}$$

Based on past experimental observations detailed in Kottari [20], the sliding shear resistance of an isolated shear key is associated with two states: i) shear resistance at first sliding (V_{slid}) and ii) ultimate sliding shear resistance (V_u) right before the rupture of the dowel bars. Assuming a smooth construction joint, the shear resistance due to the dowel action (F_d) of the vertical dowel bars is calculated using Eq. (4) which leads to the calculation of V_{slid} in Eq. (8) through Eqs. (5), 6, and 7. Based on the equilibrium of the horizontal and vertical forces [5]. V_u is calculated as per Eq. (9).

$$F_d = \sum_{i=1}^{\#of\ Vertical\ bars} \sqrt{2.M_{pl,i}.f_{cb,i}.d_{b,i}} \tag{4}$$

$$M_{pl,i} = \frac{f_y.d_{b,i}^3}{6} \tag{5}$$

$$f_{cb,i} = a_i.f_c'^{1.2} \tag{6}$$

$$a_i = 2.0 + \frac{0.5}{d_{b,i}} \tag{7}$$

$$V_{slid} = \frac{T + F_d}{\left(1 - \mu_f.\tan\beta\right)} \tag{8}$$

$$V_u = \frac{\mu_f.\cos\gamma + \sin\gamma}{1 - \mu_f.\tan\beta}.A_{vsk}f_{su} \tag{9}$$

In these equations, $M_{pl,i}$ is the plastic moment capacity of bar i, and the compressive strength of confined concrete, $f_{cb,i}$, f_c' is the uniaxial concrete compressive strength (in *ksi*), $d_{b,i}$ is the diameter of bar i (in *inches*), β is the angle of the inclined face of the shear key with respect to a vertical plane, T is the cohesive force, and μ_f is the coefficient of friction of the smooth construction joint ($\mu_f = 0.36$)[20]. γ is the angle of inclination of the vertical dowel bar with respect to the vertical (Angle of Kink). It is assumed that a bond breaker is applied to the construction joint, hence $T = 0$ in Eq. (8). The value of γ is obtained

from Kottari [20] through interpolation for the provided diameter of dowel bars. The initial stiffness (k_1) of the backbone curve is computed through the summation of shear and flexural responses of the concrete cantilever action of the shear key[26]. The stiffness of hardening (k_2) and softening (k_3) branches are expressed as a percentage of k_1 (ranging from 0.5% to 2.5% for various rebar diameters). In this study, k_2 and k_3 are assumed to be expressed with the same percentage of k_1; the ratio of k_2/k_1 is interpolated to reach dowel fracture as per Kottari [20].

The model of abutment comprises of i) abutment piles, ii) backfill soil, and iii) elastomeric bearing pads. Piles of the abutments are modeled through a trilinear hysteretic spring model in *OpenSees* with the backbone curve defined as per Choi [14]. The backbone is presented in Fig. 1d. Caltrans recommendation for pile stiffness [45], 119 kN/pile, is considered in this study. The model assumes that piles become plastic at a deformation of 25.4 mm and first yielding occurs at a displacement equal to 30% of the ultimate deformation and the yielding force is 70% of the ultimate force. The initial stiffness is assumed to degrade with soil surface yielding. The parameters of the model are given in Eq. (10) to (14).

$$K_{eff} = 7.0\ \text{kN/mm/pile} \times \text{Numberof Piles} \tag{10}$$

$$K_1 = 2.333 \times K_{eff} \tag{11}$$

$$\Delta_1 = 7.62\ mm \tag{12}$$

$$K_2 = 0.428 \times K_{eff} \tag{13}$$

$$\Delta_2 = 25.4\ mm \tag{14}$$

The backfill soil is modeled using the HyperbolicGapMaterial material with a Generalized Hyperbolic Force-Deformation (GHFD) backbone [35]. The backbone is presented in Fig. 1e. The details of the model can be found in Omrani (2015). Hence, the active resistance of the abutment is provided by the piles, while the passive action includes resistance due to the piles and backfill soil. The elastomeric bearing pads are modeled using the Steel01 material as shown in Fig. 1c [29]. The initial stiffness, k_{pad}, of the bearing pad is calculated using Eq. (15), where G is the shear modulus, A is the cross-sectional area, and h is the thickness of the bearing pad. The yield force, F_y, is calculated by multiplying the normal force, N, acting on the bearing with the coefficient of friction, μ, of the pad. Scharge [42] presented an expression for the coefficient of friction, specific to elastomer on concrete, based on experimental tests and is a function of the normal stress, σ_n (in MPa), as presented in Eq. (16). Five parallel springs, representing abutment behavior, at each short side of the deck is modeled to capture resistance to the longitudinal and rotational movements. These springs are connected by a rigid link representing the short edge of the deck. In the transverse direction, one spring on each corner of the deck is modeled to represent the behavior of the abutment. The arrangement of the longitudinal and transverse springs are shown in Fig. 1a and 1b, respectively.

$$k_{pad} = \frac{GA}{h} \tag{15}$$

$$\mu = 0.05 + \frac{0.4}{\sigma_n} \tag{16}$$

3. Ground motions

Two groups of ground motions are used in this research. Group 1 includes 500 ground motions generated from subduction sources from all over the world, while group 2 contains 500 ground motions originating from crustal sources. For group 1, ground motions arising from subduction sources with $6.0 < M_w < 9.5$ and $10\ kms < R_{Rup} < 350\ kms$. For group 2,500 ground motions are selected from a total set

of 7000 main-shock ground motions available in NGAWest2 database[38]. Crustal ground motions are selected based on the spectral match with the subduction ground motions, which means among the 7000 crustal ground motions, the selected 500 ground motions have the closest spectra (between 0 and 5.0 secs) to the spectra of 500 subduction ground motions. The response spectrum of the subduction ground motion is discretized at period ranges of 0.05 to 1.0 sec, 1.0 to 3.0 sec, and 3.0 to 5.0 sec at intervals of 0.05 sec, 0.10 sec, and 0.20 sec, respectively. This results in a target set of 500 vectors of pseudo-spectral accelerations ($\mathbf{S}_i$). For the same periods, vectors of spectral accelerations ($\mathbf{C}_j$) of all 7000 crustal ground motions are calculated. Then the spectrum of each subduction ground motion ($\mathbf{S}_i$) is compared against the spectra of all 7000 crustal ground motions individually using the sum of squared error, $SSE = \sum \left(S_i - C_j\right)^2$. The crustal ground motions with the lowest SSE is considered as the best match for that subduction ground motion.

In contrast with prior research (e.g., [13]), the crustal ground motions are not scaled to match the subduction target spectrum; hence, the ground motions used in this study are as recorded in both of the databases. Once the best match is identified, the target subduction ground motion and the spectrally equivalent crustal ground motion are selected and removed from their corresponding databases. This process is repeated for all the 500 subduction ground motions, and their spectrally matching 500 crustal ground motions are selected. The resulting ground motions in group 2 arise from crustal sources with $4.5 < M_w < 8.0$ and $1 \text{ km} < R_{Rup} < 150 \text{ km}$. The unscaled *RotD50* spectra of the as-recorded ground motions from group 1 and group 2 are shown Fig. 2a and 2b, respectively, along with the Design Uniform Hazard Spectrum (UHS) of bridges A and B. Also, the distributions of the strong motion durations (D_{5-95}) for the two groups of ground motions are shown in Fig. 3. It is noticed from Fig. 3, ground motions arising from crustal sources tend to have a shorter significant duration (D_{5-95}) than subduction sourced ground motions.

4. Statistical analysis

The methodology for investigating the significance of features (also known as Intensity Measures, IMs, and is used interchangeably throughout the text) in the prediction of *EDP* is based on statistical tools and machine learning algorithms to develop regression equations through which sensitivity of each predictor on the goodness of fit is compared and conclusions are deduced. The two groups of ground motions are used to conduct Nonlinear Time History Analysis of the two BGST bridges. Since the response of a bridge is different in the two orthogonal directions and depends heavily on the intercept angle of ground motion, for each ground motion, the two orthogonal components of each record are rotated through 180 degrees at 9 degrees in-

crements. Hence a total of 42,000 (2 bridges × 500 ground motions × 21 angles × 2 groups) nonlinear simulations are conducted throughout this study. As mentioned in the Introduction section, two types of EDPs are considered in this study: *Rot50CDR* and *Rot50DE*. While the former is associated with the peak response of column drift ratio, the latter corresponds to the energy dissipated throughout the time history of ground motion.

While peak Column Drift Ratio (*CDR*) is obtained from the history of column drift ratio, Dissipated Energy (*DE*) is obtained by monitoring the accelerations, velocities, and displacements of each node and then solving the product of equation of motion (Eq. (17), [6]) by subtracting inertial energy ($\int M\ddot{u}du$) from the input energy, as shown in Eq. (18). In these equations, M and C denote the mass and damping matrices of the MDOF system, and f_s is the restoring force. $\ddot{u}_g$ is the ground acceleration and $\ddot{u}$ and $\dot{u}$ are the relative acceleration and velocity of each mass with respect to the ground. The energy balance of the Eq. (17) is valid throughout the duration of the ground motion. Eq. (17) consists of three terms on the left-hand side which denote the relative kinetic energy, damping energy, and absorbed energy. The first term denotes the relative kinetic energy of the system with respect to the ground, which characterizes the energy temporarily stored in the kinematics of the system. The second term represents the damping energy dissipated by viscous damping, while the third denotes the absorbed energy that consists of the irrecoverable hysteretic energy and the recoverable strain energy. Despite the presence of the recoverable part, when the integration is carried out until the system comes to rest, the strain energy essentially vanishes. The final term on the right-hand side of the equation is the relative input energy induced by the ground motion to the system, as measured relative to the ground, excluding any rigid body translation. Still, if integration is carried out to the time when the system comes to rest, this is essentially equivalent to the absolute input energy [43].

Thus, the actual input energy induced to a system during an earthquake event is dissipated entirely by means of viscous damping and hysteretically absorbed energies (represented by the second and third term), which leads Eq. (18) where *DE* denotes the Dissipated Energy. Eq. (18) is applied to each node for each degree-of-freedom during each time step of ground motion to calculate *DE* for each node. Due to the additive nature of energy, *DE* of each node for each degree of freedom is added scalarly throughout the response history to obtain the overall dissipated energy.

$$\int M\ddot{u}\dot{u}dt + \int C\dot{u}\dot{u}dt + \int f_s\dot{u}dt = -\int M\ddot{u}_g\dot{u}dt \tag{17}$$

$$DE = \int M\ddot{u}_g\dot{u}dt - \int M\ddot{u}\dot{u}dt \tag{18}$$
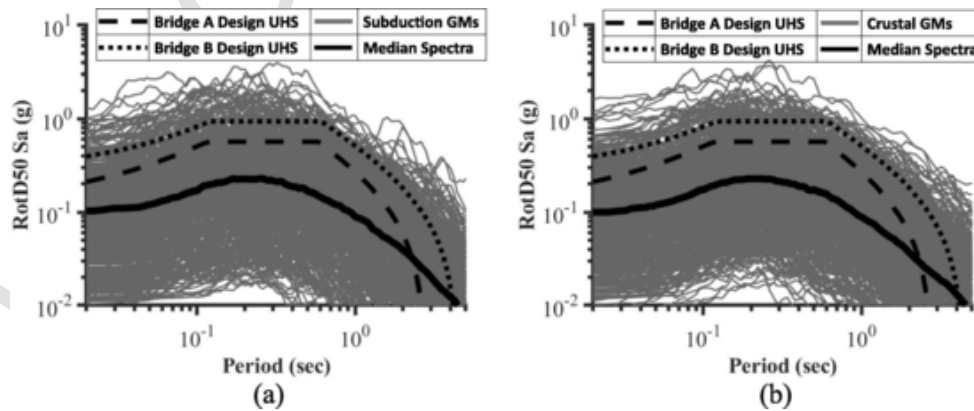

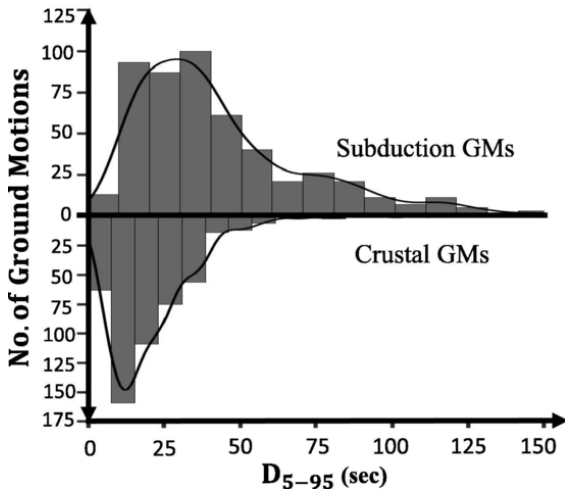**Fig. 2.** RotD50 Spectra of: (a) Subduction GMs, (b) Crustal GMs.

Fig. 3. Significant Durations (D_{5-95}) of Crustal GMs and Subduction GMs.

For conducting the sensitivity analysis, the ground motion time histories are characterized by several IMs which include: Arias Intensity (I_a), Significant Duration (D_{5-95}), mid-frequency (f_{mid}), time at 30% of cumulated Arias Intensity (t_{mid}), and the derivative of the mid-frequency (f') at t_{mid}. Significant Duration (D_{5-95}) is the time interval over which 5–95% of the integral $\int_{0}^{t_{max}} a^2(t)\,dt$, is accumulated, where $a(t)$ represents the ground acceleration at time t, and t_{max} represents the length of the accelerogram [39]. I_a used in this study is *SRSS* of the Arias Intensities of the two orthogonal components. An illustration to calculate the f_{mid} and f' for a record is provided in Fig. 4. The upper left of the figure represents the time history of the ground motion. At every instance when the ground motion history crosses up the zero-level, the cumulative number of zero-crossing is noted at the time of crossing as shown in the lower left of Fig. 4. Then, a second-order polynomial ($f(t)$) is fitted to the cumulative count of zero-level up-crossings. The fitted polynomial is differentiated to obtain a linear estimate of the filter frequency as a function of time. The value of this function at t_{mid} represents the estimate of f_{mid}, and its slope represents the estimate of f'. IMs, including D_{5-95}, f_{mid}, and f' are obtained from the component of the ground motion that possesses a higher Arias In-

tensity. As stated earlier, another parameter that is used as a feature (i.e., IM) in the statistical regression analysis is the *RotD50* spectral acceleration at the bridge natural period (termed as $S_a(T_1)$).

Apart from these general parameters, the crustal ground motions are further classified as pulse-like or non-pulse-like using Shahi and Baker [34] algorithm. Examples of the acceleration and velocity time-histories for non-pulse-like and pulse-like ground motions are provided in Fig. 5a and 5b, respectively. As can be seen in the figures, non-pulse-like ground motions tend to have similar amplitudes throughout the acceleration time history, while for pulse-like ground motions, high amplitudes are noticed only at a few peaks. The pulse characteristic is highly noticeable in the velocity time history of the pulse-like ground motion. Based on Shahi and Baker [34], among the 500 crustal ground motions, 62 ground motions are classified as pulse-like. For these 62 ground motions, pulse period (T_p) and pulse velocity (V_p), obtained from the Shahi and Baker [34] algorithm, are added as additional predictive features in the statistical analysis. Some of the limitations that this study possesses include that the bridge models used are assumed to represent the general bridge configurations and do not explicitly consider the effects of Soil-Structure Interaction (SSI). Also, there are no explicit considerations of near-fault condition and basin effects in the models.

Before performing the statistical analysis to correlate EDPs to IMs, general outliers are removed from the dataset of both crustal and subduction ground motions. This is done by first scaling the entire data set using the *Feature Scaling Standardization* algorithm [4], which involves normalization by nullifying the bias. The scaled data points are then used to fit Eq. (19) (without the pulse parameters V_p and T_p, i.e., indicator function $I = 0$) and the corresponding residuals are calculated. In Eq. (19), the variable I represents an indicator function which takes a value of 1 if the ground motion is pulse-like else is allotted a value of 0. The outliers in the resulting residuals (and hence outliers in the dataset) are then removed using the anomaly detection algorithm of *K-Means* clustering. *K-Means* clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the mean, which has the smallest Euclidean distance to the observation, serving as a prototype of the cluster. Outliers in the fitted data points are detected by clustering the residuals vs. M_w and residuals vs. R_{rup} of the fit into K clusters. The cluster of residuals farthest from the rest is classified as the outliers and the corresponding data is removed from the data set. The value of K is decided to be the *elbow point* of the plot
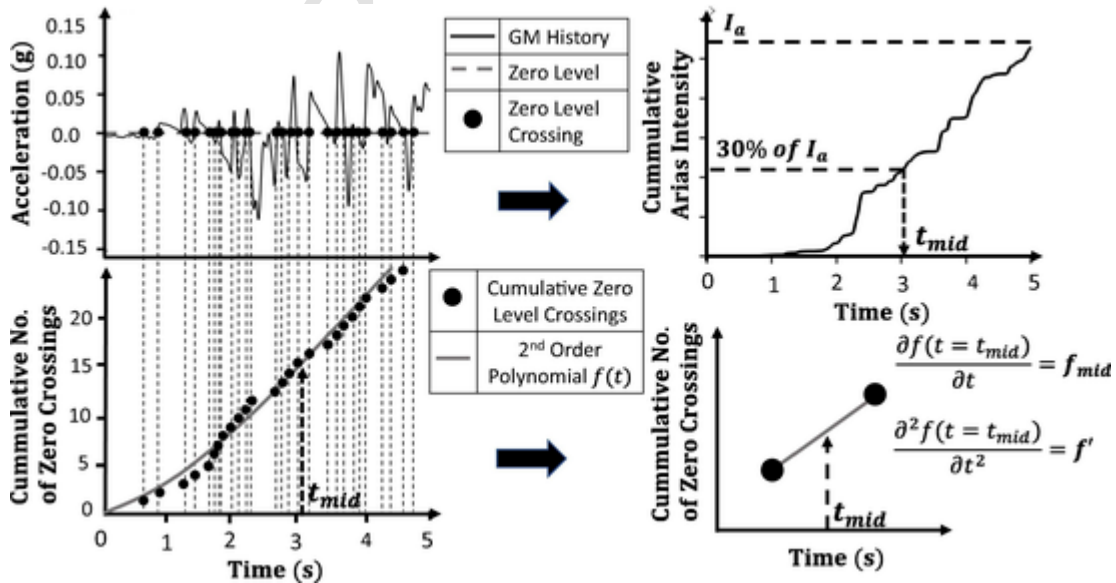

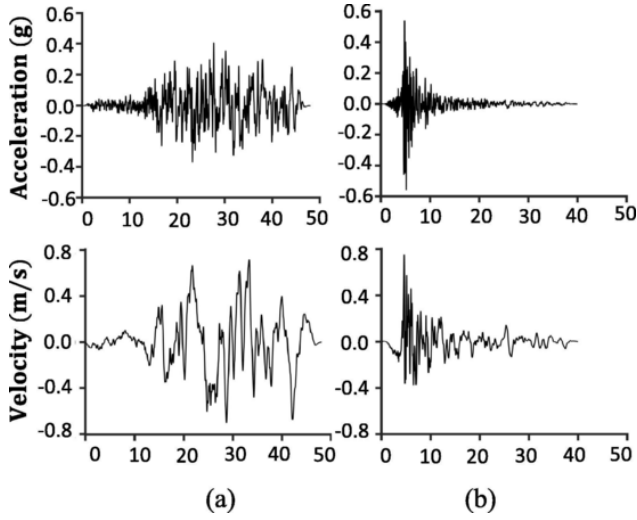Fig. 4. Illustration for f_{mid} and f'.

Fig. 5. Example of GMs classified as (a) Non-Pulse-Like, (b) Pulse-Like.

K vs. *Average within-cluster distance to the centroid*. This minimizes the average distance between the data points and the centroid within a cluster while limiting the value of *K*. To prevent removal of outliers from only certain magnitudes and epicentral distances, the outlier detection is conducted in batches. The residuals for the fitting of each formulation are clubbed into batches based on the magnitudes of their respective earthquake event, such as residuals of each formulation for $6.5 < M_w < 7.0$ form one batch, $7.0 < M_w < 7.5$ form another batch and so on. Then, the above-mentioned anomaly detection process is performed on each batch and a cluster of the outliers is removed. The process is conducted for both crustal and subduction ground motions for both types of *EDPs*.

This method of outlier removal, in case of crustal ground motions, reduces the data set from 500 to 469 and 475 for Bridge A and Bridge B, respectively; while for subduction group, the data set of 500 is reduced to 471 and 476 for Bridge A and Bridge B, respectively. The removed (outlier) data is not used in any further data processing. The sensitivity analysis of bridge response to all the features (including $I_a, D_{5-95}, f_{mid}, t_{mid}, f\hat{\ }, T_p, V_p$) is then conducted using the three algorithms including *Step-wise regression*, *Random Forests*, and *Neighbourhood Component Analysis* (explained in the following sections). The main benefit of using these algorithms as compared to traditional methods of gradient-based sensitivity analysis is that they do not require any closed functional forms to compute the partial derivatives, and the conclusions are derived from the trends observed in the data. The particular advantages and drawbacks of each algorithm are mentioned under their respective sections.

$$\ln(\boldsymbol{EDP}) = a_0 + a_1 \ln(\boldsymbol{I_a}) + a_2 \ln(\boldsymbol{D_{5-95}}) + a_3 \ln(\boldsymbol{f_{mid}}) + a_4 \ln(\boldsymbol{t_{mid}}) + a_5\boldsymbol{f'}$$
$$a_6\ln(\boldsymbol{T_p})(I) + a_7\ln(\boldsymbol{V_p})(I) + a_8\ln(\boldsymbol{S_a})$$

4.1. Step-wise regression

The general equation form used in the regression analysis is expressed in Eq. (19). Different combinations of Eq. (19) are formulated to predict the natural log of *EDP* (i.e*., Rot50CDR* and *Rot50DE*) using *Step-wise regression*. Particularly *backward selection* algorithm is used to conduct the *Step-wise regression*. In the *backward selection* algorithm, the regression is conducted with a null model containing all the available features as the predicting variables and then at each step, the least predictive variable is dropped from the regression until only one predictor is left. In other words, the coefficients of different features are set to zero in a stepwise manner to create a flow of alternative and sim-

pler formulations. The different formulations of Eq. (19) used in this study, are mentioned in Table 2, named as formulation *I – IX*. These formulations are then compared with each other. The features ($I_a, D_{5-95}, f_{mid}, f', t_{mid}, V_p, T_p, S_a$) used in prediction equations in the statistical analysis are widely classified as ground motion features (which include: $I_a, D_{5-95}, f_{mid}, f', t_{mid}, V_p, T_p$) and response feature ($S_a$). The formulations of the Eq. (19) are carefully made to minimize the use of ground motion features and response features in the same functional form. Hence, except for the inclusion of Significant Duration (D_{5-95}) in the same equation with S_a (i.e. formulation *VIII*) no other formulation includes both ground motion features and response features. For pulse-like cases, the *backward selection* algorithm does not specifically lead to these formulations in the same order. While for non-pulse-like cases, formulations *III* to *VII* (including ground motion features) and formulations *VIII* to *IX* (including response feature) are the average result of *backward selection* algorithm where f', t_{mid}, f_{mid}, and D_{5-95} are dropped in the same order as per their predictive power to estimate *Rot50CDR* and *Rot50DE*. Nonetheless, the selected formulations are mainly presented to properly emphasize the effect of ground motion duration and other ground motion and spectral parameters on the response of bridge structures.

To conduct the *Step-wise regression*, the reduced dataset containing EDPs (*Rot50CDR* and *Rot50DE*) and their corresponding ground motion and spectral parameters ($I_a, D_{5-95}, f_{mid}, f', t_{mid}, V_p, T_p, S_a$), are randomly arranged and split into Training, Validation and Testing sets, with 70% of data in the Training set, 15% in the Validation set, 15% in the Testing set. The Training set in used in fitting the data points using nonlinear regression. To prevent overfitting, regressions are conducted via *k-Fold* cross-validation with *k* set to 10. The obtained coefficients are then tested for goodness of fit on the combined dataset. The statistical measures that are used to measure the goodness of fit include Adjusted R-Squared (R_{adj}^2), Akaike Information Criterion (*AIC*), and Bayesian Information Criterion (*BIC*). These measures are carefully chosen to test the goodness-of-fit as they penalize the addition of each predictor (complexity) that is included in the fitted equation. Hence, unless the addition of a predictor/feature produces a better predictive equation, these measures tend to support the equations of a lower order. R_{adj}^2 is a variation of Coefficient of Determination (R^2) which is used in model selection. Unlike R^2, the R_{adj}^2 increases only when the increase in R^2 (due to the inclusion of a new explanatory variable) is more than one would expect to obtain randomly. R_{adj}^2 is calculated using Eq. (20). *AIC* [3] and *BIC* [33] are estimators of the relative quality of statistical models for a given set of data and are founded in information theory. Hence the models that minimize the information loss are considered to be of higher quality. *AIC* and *BIC* values of a model are calculated using Eq. (21) and Eq. (22) [36], respectively. For Eqs. (20) to (22), *n* represents the number of data points, *p* represents

Table 2
Different formulations of Eq. (19).

Formulation	Fitting Equation: $\ln(EDP) =$
I	$a_0 + a_1 \ln(I_a) + a_2 \ln(D_{5-95}) + a_3 \ln(f_{mid})$ $+ a_4 \ln(t_{mid}) + a_5 f\hat{\ } + a_6\ln(T_p)(I) + a_7\ln(V_p)(I)$
II	$a_0 + a_1 \ln(I_a) + a_2 \ln(D_{5-95}) + a_3 \ln(f_{mid})$ $+ a_4 \ln(t_{mid}) + a_5 f\hat{\ } + a_6\ln(T_p)(I)$
III	$a_0 + a_1 \ln(I_a) + a_2 \ln(D_{5-95})$ $+ a_3 \ln(f_{mid}) + a_4 \ln(t_{mid}) + a_5 f\hat{\ }$
IV	$a_0 + a_1 \ln(I_a) + a_2 \ln(D_{5-95})$ $+ a_3 \ln(f_{mid}) + a_4 \ln(t_{mid})$
V	$a_0 + a_1 \ln(I_a) + a_2 \ln(D_{5-95}) + a_3 \ln(f_{mid})$
VI	$a_0 + a_1 \ln(I_a) + a_2 \ln(D_{5-95})$
VII	$a_0 + a_1 \ln(I_a)$
VIII	$a_0 + a_2 \ln(D_{5-95}) + a_8\ln(S_a)$
IX	$a_0 + a_8\ln(S_a)$

the number of estimated parameters (features) and $\hat{\sigma}$ is the maximum value of the likelihood function for the model.

$$R_{adj}^2 = 1 - \left(1 - R^2\right)\frac{n-1}{n-p-1} \tag{20}$$

$$AIC = \log\hat{\sigma}^2 + \frac{n+2p}{n} \tag{21}$$

$$BIC = \log\hat{\sigma}^2 + \frac{plogn}{n} \tag{22}$$

Tables 3, 4, and 5 present results of *Step-wise regression* for Bridge A for pulse-like crustal, non-pulse-like crustal, and subduction ground motions, respectively. The tables indicate values of R_{adj}^2, *AIC*, and *BIC* for the different formulations of Eq. (19) (formulations *I* to *IX*, as indicated in Table 2) for both target EDPs (*Rot50CDR* and *Rot50DE*). Results for *Rot50CDR* in Table 3 show that for pulse-like ground motions, formulation *I* (consisting of I_a, D_{5-95}, f_{mid}, f', t_{mid}, V_p, T_p) lends the highest R_{adj}^2 = 0.592, and lowest *AIC* and *BIC* values of 39.89 and 44.1, respectively. While for *Rot50DE*, formulation *II* (consisting of I_a, D_{5-95},

f_{mid}, f', t_{mid}, T_p) is the desired with R_{adj}^2 = 0.530, and lowest *AIC* and *BIC* values of 63.19 and 73.00, respectively. It is also observed that formulations that do not contain pulse parameters (V_p, T_p), which include formulations *III* to *VII*, there is poor fitting for both *Rot50CDR* and *Rot50DE*. This is justifiable since the characteristics of pulse-like ground motions are dominated by their pulse parameters; hence, the results are statistically poor if these are not included in the formulation for estimating the EDPs. In terms of response parameter S_a, formulations *VIII* and *IX* seem to show a better fit compared with other formulations. However, it is essential to notice here that the value of R_{adj}^2 increases from 0.719 to 0.735 and 0.628 to 0.639 when D_{5-95} is dropped from the regression to estimate *Rot50CDR* and *Rot50DE*, respectively. Also, it is observed that the values of *AIC* and *BIC* decrease when D_{5-95} is dropped as a feature in the drop from formulation *VIII* to *IX*. A similar trend is observed between formulations *VI* and *VII* involving features I_a and D_{5-95} for both types of EDPs. This shows that the effect of duration in crustal pulse-like ground motions is highly insignificant.

Table 4 presents the results of non-pulse-like crustal ground motions. In general, it is observed that the goodness of fit in predict-

Table 3
Results of *Step-wise regression* for Bridge A under Pulse-Like Crustal GMs.

Formulation	R_{adj}^2		AIC		BIC	
	Rot50CDR	Rot50DE	Rot50CDR	Rot50DE	Rot50CDR	Rot50DE
I	0.592	0.520	39.89	63.84	44.10	73.29
II	0.531	0.530	42.63	63.19	45.44	73.00
III	0.295	0.221	54.34	85.42	60.74	89.95
IV	0.301	0.273	55.88	82.07	59.88	87.08
V	0.306	0.288	55.89	81.14	58.50	86.75
VI	0.280	0.319	55.19	78.42	57.40	81.62
VII	0.314	0.334	51.83	77.42	54.64	80.22
VIII	0.719	0.628	29.57	62.90	28.01	71.10
IX	0.735	0.639	25.21	60.01	26.77	69.81

Table 4
Results of *Step-wise regression* for Bridge A under Non-Pulse-Like Crustal GMs.

Formulation	R_{adj}^2		AIC		BIC	
	Rot50CDR	Rot50DE	Rot50CDR	Rot50DE	Rot50CDR	Rot50DE
III	0.845	0.762	586.27	1207.66	601.44	1234.14
IV	0.849	0.787	584.87	1204.66	599.84	1230.70
V	0.867	0.795	581.73	1203.47	598.51	1227.01
VI	0.851	0.810	587.82	1201.23	604.40	1221.83
VII	0.877	0.817	574.87	1198.66	595.84	1219.63
VIII	0.929	0.850	238.05	1133.12	250.02	1142.39
IX	0.939	0.857	236.53	1132.81	244.92	1141.51

Table 5
Results of *Step-wise regression* for Bridge A under Subductions GMs.

Formulation	R_{adj}^2		AIC		BIC	
	Rot50CDR	Rot50DE	Rot50CDR	Rot50DE	Rot50CDR	Rot50DE
III	0.756	0.556	526.99	1190.41	546.88	1193.29
IV	0.786	0.561	518.89	1182.32	538.96	1191.39
V	0.820	0.562	515.90	1178.52	531.16	1183.77
VI	0.841	0.574	513.73	1174.56	525.17	1181.00
VII	0.753	0.509	526.52	1179.99	540.15	1185.62
VIII	0.937	0.594	241.32	1171.32	272.76	1177.76
IX	0.857	0.537	259.35	1179.45	286.98	1187.08

ing EDPs (*Rot50CDR* and *Rot50DE*) is better in non-pulse-like ground motions than the pulse-like ground motions. This table shows that for both target variables (*Rot50CDR* and *Rot50DE*), among the formulations based on ground motion features (i.e., formulation *III, IV, V, VI,* and *VII*), formulation *VII* (consisting only of I_a) results in the highest R_{adj}^2 = 0.877, and R_{adj}^2 = 0.817 in estimating *Rot50CDR* and *Rot50DE*, respectively. Formulation *VII* also possesses lowest *AIC* and *BIC* values of 574.87 and 595.84 in estimating *Rot50CDR*, and values of 1198.66 and 1219.63 in estimating *Rot50DE*. This means that among the ground motion features, I_a possesses the highest predictive power in estimating the EDPs. Similar to pulse-like cases, it is observed from Table 4 that for non-pulse-like crustal ground motions for formulation *VIII* to *IX*, R_{adj}^2 increases from 0.929 to 0.939 and 0.850 to 0.857 when D_{5-95} is dropped from the regression to estimate *Rot50CDR* and *Rot50DE*, respectively. Also, it is observed that values of *AIC* and *BIC* decrease when D_{5-95} is dropped as a feature in the drop from formulation *VIII* to *IX*. Similar trend is also observed between formulations *VI* and *VII* involving features I_a and D_{5-95} for both types of EDPs. This shows that the effect of duration in crustal non-pulse-like ground motions is highly insignificant, and I_a and S_a alone can be used in the ground motion selection process.

Table 5 shows the results of the goodness of fit measures for Bridge A for subduction ground motions. It is observed from the table, that for both target variables (*Rot50CDR* and *Rot50DE*), among the formulations containing ground motion features, i.e., formulations *III* to *VII*, formulation *VI* (consisting of I_a and D_{5-95}) gives highest R_{adj}^2 = 0.841, and R_{adj}^2 = 0.574 in estimating *Rot50CDR* and *Rot50DE*, respectively. Formulation *VI* also possesses lowest *AIC* and *BIC* values of 513.73 and 525.17 in estimating *Rot50CDR*, and values of 1174.56 and 1181.00 in estimating *Rot50DE*. Also, it is noticed that dropping D_{5-95} in formulation *VII* results in 0.09 (=0.841−0.753) decrease in R_{adj}^2 and increase of 12.80 (=526.52−513.72) and 14.98 (=540.15−525.17) in *AIC* and *BIC* in estimating *Rot50CDR*. Similarly, it is noticed that dropping D_{5-95} in formulation *VII* results in 0.065 (0.574−0.509) decrease in R_{adj}^2 and increase of 5.43 (=1179.99−1174.56) and 4.62 (=1185.62−1181) in *AIC* and *BIC* in estimating *Rot50DE*. This means that among the ground motion features, I_a along with D_{5-95} possesses the highest predictive power in estimating the EDPs; hence duration cannot be ignored in subduction ground motions. It is further observed from Table 5 that for subduction ground motions for formulation *VIII* to *IX*, R_{adj}^2 decreases from 0.937 to 0.857 and 0.594 to 0.537 when D_{5-95} is dropped from the regression to estimate *Rot50CDR* and *Rot50DE*, respectively. Also, it is observed that the values of *AIC* and *BIC* in-

crease significantly when D_{5-95} is dropped as a feature in the drop from formulation *VIII* to *IX*. These results indicate that in subduction ground motions, S_a alone is not a good estimator of the EDPs of bridge structures.

4.2. Random forest

Random Forests algorithm is based on two methods: 1) Bagging, and 2) Random subspace method [44]. In this study, the method of Bagging is primarily used to create randomized decision trees. Given the number of trees (*S*) used in developing the forest, *S* datasets are created from random sampling (with replacement) of the data. This is called the bootstrap dataset. For each bootstrap dataset, a *Classification and Regression Tree CART* (K_i) is created. Due to the 'with-replacement' nature of the process, the bootstrap datasets can have duplicate data records and can have missing data records from the original datasets. Bootstrap datasets that do not contain a particular record from the original dataset are called Out-Of-Bag (OOB) examples [44]. Out-Of-Bag (OOB) estimate for the generalization error is the aggregation of errors of the OOB examples. The first step in measuring the variable importance in a data set is to fit a random forest to the data. During the fitting process, the OOB error for each data point is recorded and averaged over the forest. To measure the importance of the jth feature after training, the values of the jth feature are permuted among the data, and the OOB error is again computed on this perturbed data set. The importance score for the jth feature is computed by averaging the difference in OOB error before and after the permutation over all trees. The score is normalized by the standard deviation of these differences. Features that produce large values for this score are ranked as more important than features that produce small values [44]. Due to their non-parametric nature, *Random Forests* are fairly robust and have a high power of handling large data sets with higher dimensionality. Random forests also work very well with smaller datasets, and contrary to previously used *Step-wise regression, Random Forests* handle the outliers better and can deduce the highly nonlinear relationships among the features and target variable. Using the *Random Forests* algorithm, the relative importance (*RI*) of each feature in predicting the target variable (EDP) is obtained and compared.

Fig. 6 and Fig. 7 present relative importance (*RI*) of the different features in predicting EDPs for crustal (non-pulse-like and pulse-like) ground motions and subduction ground motions, respectively. Since S_a is a response parameter and *Random Forests* algorithm computes the relative importance of the features, to understand the feature importance properly, the random forest is applied twice for each case, first to all the features except S_a and then all features including S_a. Hence

Fig. 6. Relative Importance (*RI*) of Crustal GMs for Bridge A with target variable: (a) *Rot50CDR*, (b) *Rot50CDR* (including S_a as predictor), (c) *Rot50DE*, and (d) *Rot50DE* (including S_a as predictor), for Non-Pulse-Like GMs; and, (e) *Rot50CDR*, (f) *Rot50CDR* (including S_a as predictor), (g) *Rot50DE*, and (h) *Rot50DE* (including S_a as predictor), for Pulse-Like GMs.
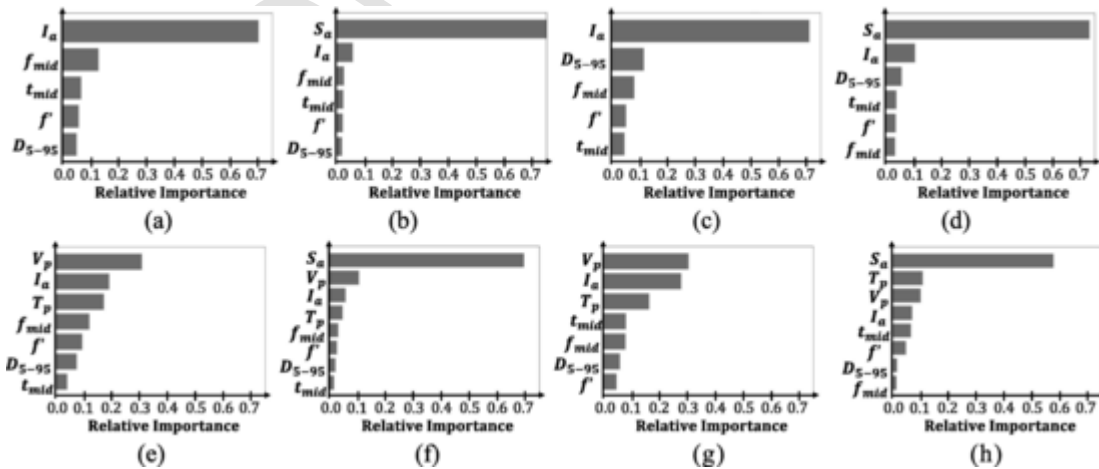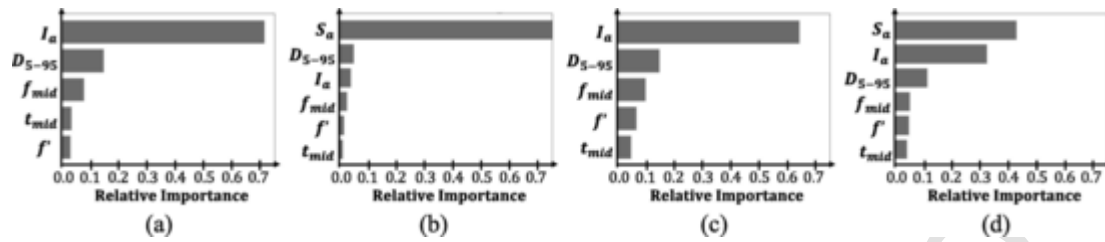
Fig. 7. Relative Importance (*RI*) of Subduction GMs for Bridge A with target variable: (a) *Rot50CDR*, (b) *Rot50CDR* (including S_a as predictor), (c) *Rot50DE*, and (d) *Rot50DE* (including S_a as predictor).

for crustal ground motions, there are results for total 8 cases (2 EDPs × two types of motions × 2 cases) of random forests, which include 2 EDPs (*Rot50CDR* and *Rot50DE*), two types of motions (non-pulse-like and pulse-like), and 2 cases (excluding S_a and including S_a). While for subduction ground motions, results are presented for 4 cases (2 EDPs × 2 cases).

Results in Fig. 6a and 6c illustrate that for non-pulse-like crustal ground motions, while predicting the *Rot50CDR* and *Rot50DE*, I_a is heavily rated as the best predictor of response EDPs among the ground motion features. Other features such as f_{mid} and t_{mid}, do show some importance; however, they are deemed as not significant as compared to I_a. It is observed that when S_a is included as a predictor (Fig. 6b and c), S_a tends to overpower all the other predictors and shows the highest *RI* in estimating both types of EDPs. For pulse-like crustal ground motions, it is observed from Fig. 6e and g, that pulse parameters V_p and T_p are as significant as I_a in predicting the *Rot50CDR* and *Rot50DE*, respectively. Among these three, V_p is classified as the most important feature with the highest *RI* followed by I_a and T_p. Similar to non-pulse-like case, when S_a is included as a predictor (Fig. 6f and h), S_a shows the highest *RI* in estimating both types of EDPs. This is again expected since S_a is clearly a response parameter itself and does not solely depend on the ground motion time-history. However, it is further observed that the *RI* for S_a in pulse-like ground motions is lower than the *RI* for S_a in non-pulse-like ground motions, which shows the importance of pulse parameters and the stochasticity involved in the EDPs arising from pulse-like ground motions. For both types of ground motion, D_{5-95} is heavily deemed insignificant in estimating the EDPs. D_{5-95} still tends to show greater importance for non-pulse-like ground motions when predicting the energy-based EDP (*Rot50DE*); for all other cases, it is clear that D_{5-95} is not a good estimator of the EDPs.

Results in Fig. 7a and c show that even for subduction ground motions, while predicting the *Rot50CDR* and *Rot50DE*, I_a is heavily rated as the best predictor of response EDPs among the ground motion features. However, it is clear from the figures that the *RI* of the duration parameter D_{5-95} is significantly higher for subduction ground motions than the crustal ground motions. D_{5-95} consistently rises to be the second most important feature among the ground motion features. This proves that the consideration of duration in subduction ground motions is highly important as compared to crustal ground motions. Furthermore, even when S_a is included as a predictor (Fig. 7b and d), though S_a shows the highest *RI* in estimating both types of EDPs, D_{5-95} continues to show significant *RI* for both types of EDPs. This further bolsters the conclusion that duration must be given proper attention while selecting subduction ground motions.

4.3. Neighborhood component analysis (NCA)

Previously used Random Forests algorithm has some drawbacks for determining feature importance. For example, if the data contains groups of correlated features of similar relevance for the target variable, then smaller groups are favored over larger groups. *Random Forests* algorithm sometimes leads to overfitting. Neighborhood compo-

nent analysis (NCA) is a non-parametric and embedded method for selecting features to maximize prediction accuracy of regression and classification algorithms [19]. Functionally, it serves the same purposes as the *K-nearest neighbors* (KNN) algorithm but makes direct use of stochastic nearest neighbors. NCA aims at learning a distance metric by finding a linear transformation of input data such that the average leave-one-out (LOO) classification performance is maximized in the transformed space. The critical insight to the algorithm is that matrix $\boldsymbol{A}$ corresponding to the transformation can be found by defining a differentiable objective function for $\boldsymbol{A}$, followed by the use of an iterative solver such as *conjugate gradient descent*. In this case, the aim is to predict the response given the dataset. NCA randomly picks a point ($Ref(x)$) from the dataset as the 'reference point' for x and then sets the response value at x equal to the response value of the reference point ($Ref(x)$). One of the benefits of this algorithm is that the number of classes k is determined as a function of $\boldsymbol{A}$, up to a scalar constant. This makes the algorithm simplistic for model selection and assigning feature weights (*FW*) to the different predictors. Values of *FW* close to zero represents that the particular features can be ignored in estimating the target variable. NCA learns the feature weighting vector by maximizing the expected leave-one-out classification accuracy. The algorithm makes no parametric assumptions about the distribution of the data and scales naturally to multiclass problems [41]. Unlike many other objective functions (where good optima are not necessarily deep but rather broad) for this algorithm, the larger drive of objective function $f(\boldsymbol{A})$ results in better performance. In other words, overfitting is highly reduced. Let $T = \left\{ \left(\boldsymbol{x}_1, y_1 \right) \cdots \left(\boldsymbol{x}_N, y_N \right) \right\}$ be a set of training samples, where $\boldsymbol{x}_i$ a d-dimensional feature vector, and N is the number of samples. The goal of the algorithm is to find a weighting vector $\boldsymbol{FW}$ that lends itself to select the feature subset when the distance between two samples x_i and x_j is given by Eq. (23).

$$D_{FW}\left(x_i, x_j\right) = \sum_{l=1}^{d} FW_l^2 \left|x_{il} - x_{jl}\right| \tag{23}$$

Figs. 8 and 9 present the results for feature weights (*FW*) for crustal and subduction ground motions, respectively. Since the *FW* are absolute values, all the features (I_a, D_{5-95}, f_{mid}, f', t_{mid}, V_p, T_p, and S_a) are used to conduct NCA on the EDPs (*Rot50CDR* and *Rot50DE*). Fig. 8a and b, show the *FW* for the non-pulse-like crustal ground motions in estimating *Rot50CDR* and *Rot50DE*, respectively. It is observed that for *Rot50CDR*, S_a and I_a show the highest *FW* while all other features have *FW* close to 0, which signifies that the other features are not important. While for *Rot50DE*, apart from S_a and I_a, an increased *FW* is noticed for D_{5-95}, showing that duration is an important feature in estimating the dissipated energy. From Fig. 8c and 8d, it is observed that for pulse-like crustal ground motions, the most important features consist of pulse parameters V_p, and T_p, and response parameter S_a for both EDPs. Similar to the non-pulse-like case, an increased *FW* is noticed for D_{5-95} and I_a for estimating *Rot50DE*. NCA for Bridge A under subduction ground motions is shown in Fig. 9a and b for *Rot50CDR* and *Rot50DE*, respectively. From the figures, it is noticed that for subduction ground motions S_a, I_a, and D_{5-95} possess the highest *FW* and hence are
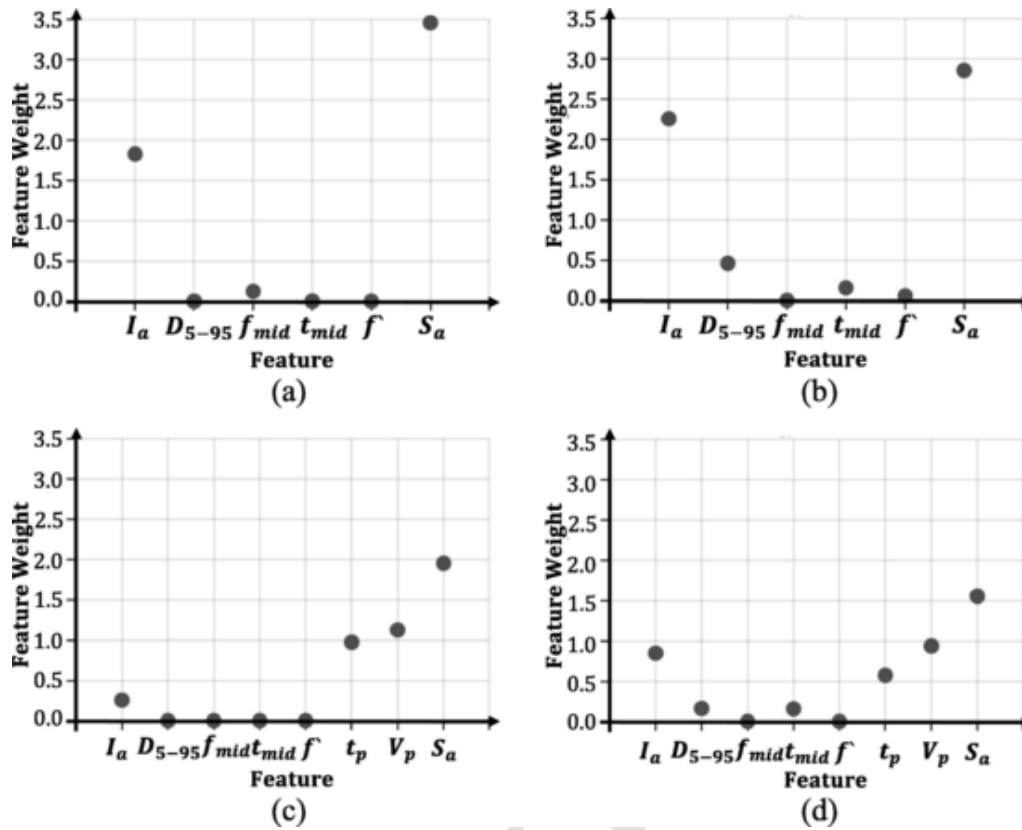
Fig. 8. Feature Weights (*FW*) of Crustal GMs for Bridge A with target variable: (a) *Rot50CDR*, and (b) *Rot50DE*, for Non-Pulse-Like GMs; and, (c) *Rot50CDR*, and (d) *Rot50DE*, for Pulse-Like GMs.
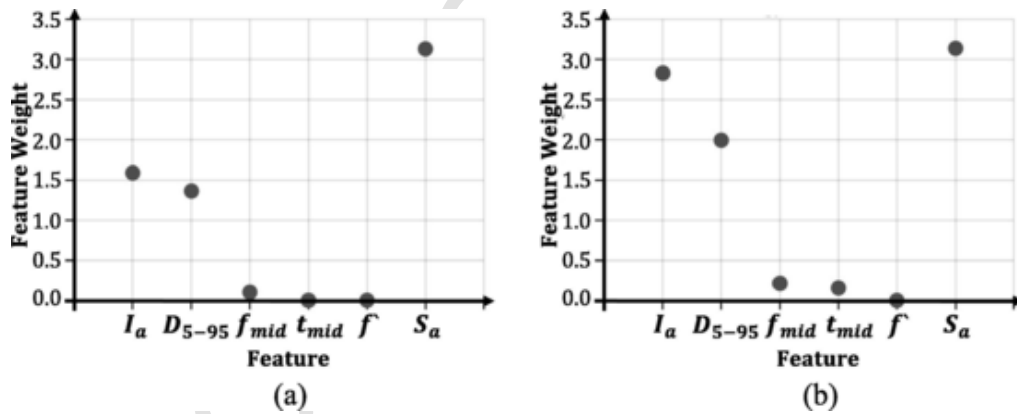
Fig. 9. Feature Weights (*FW*) of Subduction GMs for Bridge A with target variable: (a) *Rot50CDR*, and (b) *Rot50DE*.

highly important features to estimate both types of EDPs (*Rot50CDR* and *Rot50DE*).

5. Results and discussion

While the results have been mainly presented for Bridge A, identical results were also observed for Bridge B; hence, general conclusions are made for the BGST bridges. Based on the results of the three statistical procedures, in general, it is observed that for non-pulse-like ground motions when estimating the peak response (*Rot50CDR*), duration of the ground motion has minimal impact and the response is mainly controlled by the Arias Intensity (I_a) and the Spectral Acceleration at first mode period (S_a). However, when estimating the energy-based EDP, *Rot50DE*, it is observed that the impact of duration rises and is usually the third most important feature following Arias Intensity (I_a)[2]and the Spectral Acceleration at first mode period (S_a). This is intuitive

as longer duration leads to more deterioration cycles causing higher energy dissipation.

Nevertheless, it is observed that the absolute importance of the duration is still much lesser than the I_a and S_a for estimating *Rot50DE* using non-pulse-like crustal ground motions. Hence it can be said that I_a and S_a are sufficient features for estimating the target EDP. For pulse-like crustal ground motions, the main features that are deemed significant in estimating the EDPs include pulse parameters T_p and V_p, and spectral parameter S_a. All other features are seen to show an insignificant impact on the prediction of both types of EDPs. In pulse-like crustal ground motions, strong motion duration is deemed insignificant even for the energy-based EDP (*Rot50DE*). This is because pulse-like ground motions create the EDPs due to one strong pulse and the residual ground motion by itself is quite benign and hence does not create damage. On the contrary, for subduction ground motions, it is ob-

served that on average S_a, I_a, and D_{5-95} are highly correlated with the EDPs. Though the importance of D_{5-95} in estimating peak response, *Rot50CDR* is observed to be quite high; the importance increases to a much higher level while predicting the dissipated energy *Rot50DE*. This is as per the intuition since the subduction ground motions are usually long with high amplitude and lead to a higher number of deterioration cycles causing both greater damage and higher energy dissipation.

The results obtained using formulations *VI* (including I_0 and D_{5-95}) and *VIII* (including S_a and D_{5-95}) are presented in Figs. 10, 11, 13, and 14. They show the effect of D_{5-95} on *Rot50CDR* (using crustal GMs), *Rot50DE* (using crustal GMs), *Rot50CDR* (using subduction GMs), and *Rot50DE* (using subduction GMs), respectively. The ground motions and corresponding EDPs are discretized by clubbing I_a and S_a into five sets. The sets are roughly decided based on 15, 32, 49, 66, and 83 percentiles of data of I_a and S_a. The data for discretization is mentioned in the tables included in each sub-figure of Figs. 10, 11, 13, and 14. Figs. 10 and 11 present the *Rot50CDR* vs D_{5-95} and *Rot50DE* vs D_{5-95}, respectively, for Bridge A and Bridge B for crustal ground motions. Apart from the discretization of data based on S_a and I_0 values, the data obtained from crustal ground motions are further classified as pulse-like or non-pulse-like based on Shahi and Baker [34]. For each discretized set, the average values of S_a and I_a are obtained and termed as μ_{S_a} and μ_{I_a}, respectively, and then the regression line is fitted to present the trend in that discretized set. This means that for each discretized set, regression lines $a_0 + a_2\ln\left(\boldsymbol{D}_{5-95}\right) + a_1\ln\left(\mu_{I_a}\right)$ and $a_0 + a_2\ln\left(\boldsymbol{D}_{5-95}\right) + a_8\ln\left(\mu_{S_a}\right)$ for formulations *VI* and *VIII*, respectively, are used to display the trend, where the vector $\boldsymbol{D}_{5-95}$ is the vector containing the duration of ground motions of that particular discretized set and μ_{I_a} and μ_{S_a} are the mean values of vectors $\boldsymbol{I_a}$ and $\boldsymbol{S_a}$ for the discretized set, respectively. In Figs. 10 and 11, sub-figures (*a*) and (*c*) show EDP vs D_{5-95} along with the regression trends for formulation *VI*

(i.e., $a_0 + a_2\ln\left(\boldsymbol{D}_{5-95}\right) + a_1\ln\left(\mu_{I_a}\right)$) while sub-figures *b* and *d* show EDP vs D_{5-95} along with the regression trends for formulation *VIII* (i.e., $a_0 + a_2\ln\left(\boldsymbol{D}_{5-95}\right) + a_8\ln\left(\mu_{S_a}\right)$). Figs. 13 and 14 show similar figures for Bridge A and Bridge B under subduction ground motions for *Rot50CDR* and *Rot50DE*, respectively. All these plots are developed to single out the impact of D_{5-95} on both the bridge EDPs *Rot50CDR* and *Rot50DE*.

It is noticed in Fig. 10 that for both Bridge A and Bridge B undergoing ground motions from crustal sources, the trend of the regression lines is almost horizontal and parallel to the *x*-axis for both formulation *VI* and *VIII*, which signify that an increase in duration of crustal ground motions does not contribute to the peak response *Rot50CDR* of the bridges. Moreover, for ground motions with lower I_a levels for both bridges A and B, a decreasing trend is noticed between *Rot50CDR* and D_{5-95} (Fig. 10a and 11a). It is observed that this decrease in trend is mainly due to the short duration ground motions possessing strong velocity pulses; hence an increase in arias intensity levels do not significantly impact the response. In general, it is noticed that crustal ground motions with a short duration that cause high *Rot50CDR* are classified as pulse-like motions by Shahi and Baker [34]. Similar trends are also noticed for dissipated energy, *Rot50DE*, in Fig. 11. However, a slight increase in trends is noticed for the bridges with *Rot50DE* as the EDP (especially for higher I_a and S_a levels), this goes with the intuition that the dissipated energy is affected by an increase in the strong motion duration of the ground motions. However, the slope of the trend lines is not high, and hence the increase is still not significant enough to conclude that the duration of crustal ground motions causes a huge change in the *Rot50DE* EDP of the bridge structures. Additionally, Fig. 12 shows the trends in the slope of D_{5-95} in formulations *VI* and *VIII* for the different discretized seismic intensity levels (normalized by their corresponding S_a and I_a of design hazard level termed as $S_{a,haz}$ and $I_{a,haz}$).

Non Pulse	Pulse	**Fitting Formulation *VI***			Non Pulse	Pulse	**Fitting Formulation *VI***			Non Pulse	Pulse	**Fitting Formulation *VI***		
		I_a Discretization	μ_{I_a} for Bridge				I_a Discretization	μ_{I_a} for Bridge				I_a Discretization	μ_{I_a} for Bridge	
			A	B				A	B				A	B
◦	◇	$I_a < 5$	1.87	1.87	●	◆	$15 < I_a < 40$	26.47	26.47	●	◆	$I_a > 100$	307.7	307.7
◦	◇	$5 < I_a < 15$	9.31	9.31	●	◆	$40 < I_a < 100$	66.47	66.47					

(a) BRIDGE A — ln(Rot50CDR) vs ln(D₅₋₉₅)

(b) BRIDGE B — ln(Rot50CDR) vs ln(D₅₋₉₅)

Non Pulse	Pulse	**Fitting Formulation *VIII***			Non Pulse	Pulse	**Fitting Formulation *VIII***			Non Pulse	Pulse	**Fitting Formulation *VIII***		
		S_a Discretization	μ_{S_a} for Bridge				S_a Discretization	μ_{S_a} for Bridge				S_a Discretization	μ_{S_a} for Bridge	
			A	B				A	B				A	B
◦	◇	$S_a < 0.05$	0.02	0.02	●	◆	$0.10 < S_a < 0.25$	0.14	0.13	●	◆	$S_a > 0.50$	0.65	0.60
◦	◇	$0.05 < S_a < 0.10$	0.06	0.07	●	◆	$0.25 < S_a < 0.50$	0.27	0.28					

(c) BRIDGE A — ln(Rot50CDR) vs ln(D₅₋₉₅)

(d) BRIDGE B — ln(Rot50CDR) vs ln(D₅₋₉₅)

Fig. 10. Effect of D_{5-95} on *Rot50CDR* using crustal GMs on: a) Bridge A (Formulation *VI*), b) Bridge B (Formulation *VI*), c) Bridge A (Formulation *VIII*), and d) Bridge B (Formulation *VIII*).

Non Pulse	Pulse	Fitting Formulation *VI*			Non Pulse	Pulse	Fitting Formulation *VI*			Non Pulse	Pulse	Fitting Formulation *VI*		
		I_a Discretization	μ_{I_a} for Bridge A	B			I_a Discretization	μ_{I_a} for Bridge A	B			I_a Discretization	μ_{I_a} for Bridge A	B
○	◇	$I_a < 5$	1.87	1.87	●	◆	$15 < I_a < 40$	26.47	26.47	●	◆	$I_a > 100$	307.7	307.7
●	◆	$5 < I_a < 15$	9.31	9.31	●	◆	$40 < I_a < 100$	66.47	66.47					

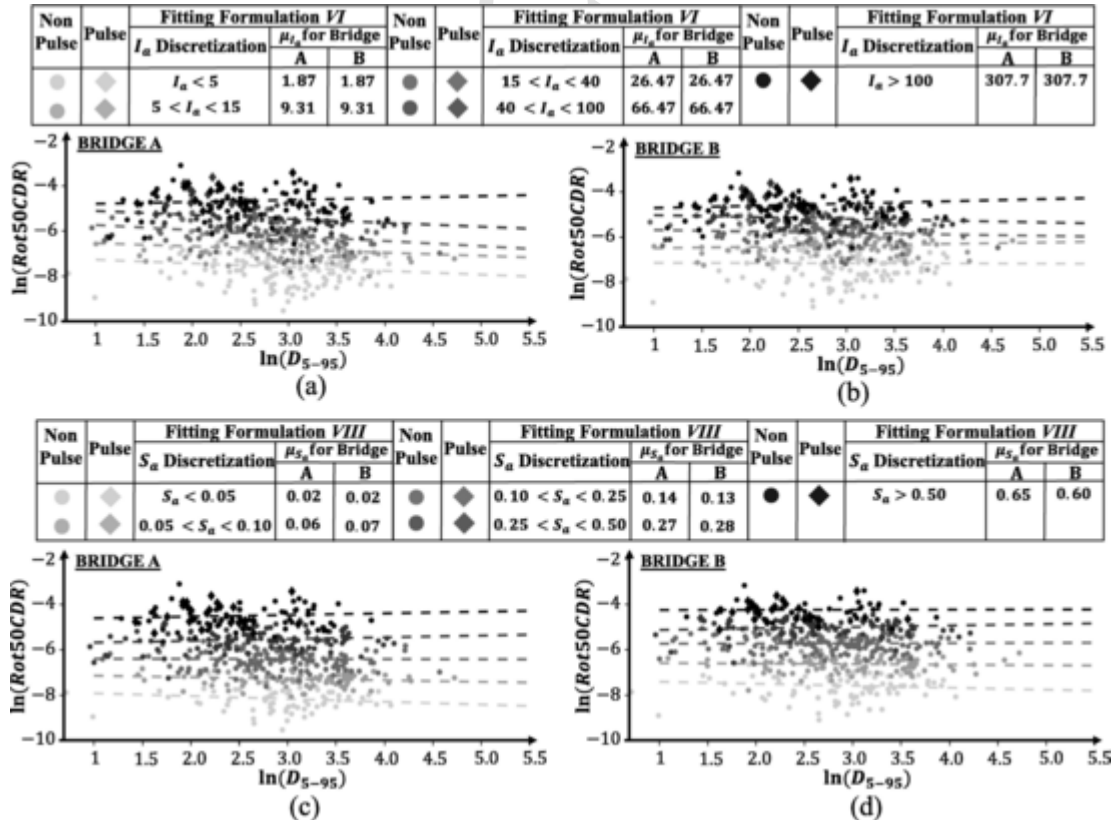| Non Pulse	Pulse	Fitting Formulation *VIII*			Non Pulse	Pulse	Fitting Formulation *VIII*			Non Pulse	Pulse	Fitting Formulation *VIII*		
		S_a Discretization	μ_{S_a} for Bridge A	B			S_a Discretization	μ_{S_a} for Bridge A	B			S_a Discretization	μ_{S_a} for Bridge A	B
○	◇	$S_a < 0.05$	0.02	0.02	●	◆	$0.10 < S_a < 0.25$	0.14	0.13	●	◆	$S_a > 0.50$	0.65	0.60
●	◆	$0.05 < S_a < 0.10$	0.06	0.07	●	◆	$0.25 < S_a < 0.50$	0.27	0.28					

Fig. 11. Effect of D_{5-95} on *Rot50DE* using crustal GMs on: a) Bridge A (Formulation *VI*), b) Bridge B (Formulation *VI*), c) Bridge A (Formulation *VIII*), and d) Bridge B (Formulation *VIII*).
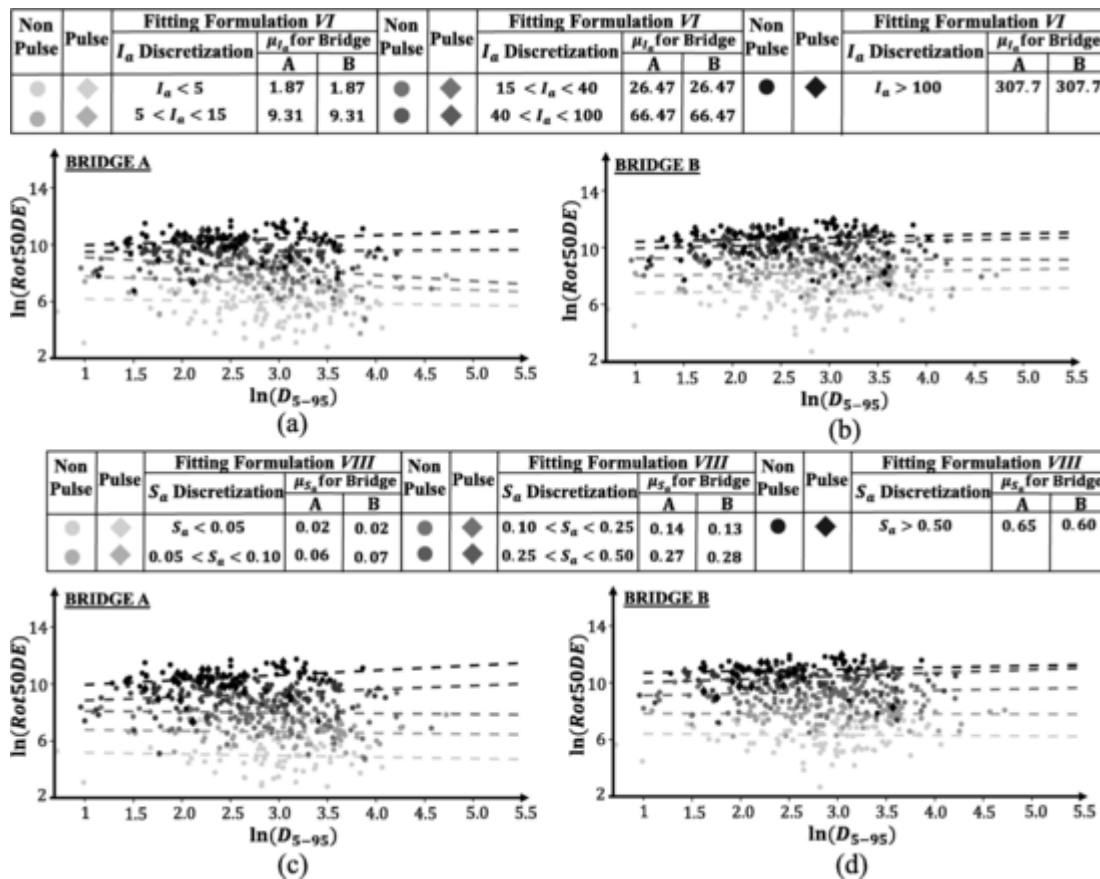
It is observed that as there is an increase in the ratio of the mean S_a (μ_{S_a}) and design hazard S_a ($S_{a,haz} = 0.53\ g$ for Bridge A and $= 0.65\ g$ for Bridge B) and in the ratio of the mean I_a (μ_{I_a}) and design hazard I_a ($I_{a,haz} = 225.67$ cm/sec for Bridge A and $= 155.36$ cm/sec for Bridge B), there is no significant increase in regression coefficients (a_2) associated with D_{5-95} in formulations *VI* and *VIII*. This minimal impact of including D_{5-95} in the different formulations for various intensity levels indicates the insignificance of duration as a parameter for characterizing the response bridges in crustal earthquakes.

On the contrary, Figs. 13 and 14 portray the effect of duration of ground motions arising from subduction sources on EDPs *Rot50CDR* and *Rot50DE,* respectively; it is noticed that an increase in the duration of ground motions is associated with an increase in the EDPs. A linear increasing trend between *Rot50CDR* and D_{5-95} (Fig. 13) and *Rot50DE* and D_{5-95} (Fig. 14) portray that the duration of subduction ground motions cannot be neglected in estimating seismic demands of bridges. Fig. 13 shows that the regression lines have a very high slope in relating D_{5-95} with *Rot50CDR*, especially for formulation *VIII,* which associates D_{5-95} with S_a. This is due to the fact that formulation *VI* contains I_a which is highly correlated with D_{5-95}; hence, D_{5-95} does not bring a high amount of information into the regression model. A higher slope, hence a higher increasing trend is observed between D_{5-95} and *Rot50DE* (Fig. 14) for both bridges, which further solidifies the conclusion that the duration of subduction ground motions is highly associated with the EDPs of bridges especially the dissipated energy (*Rot50DE*). It is observed from the figures that for both EDPs for both bridges, the slopes of the regression lines are higher for higher levels of intensity measures (I_a and S_a). This is since the higher intensity measure causes higher seismic demands that increase the level of structural nonlinear behavior, where both strength and stiffness degradation are imposed by

the large number of nonlinear cycles in ground motions with higher D_{5-95}. This conclusion is further supported in Fig. 15, where the slopes of D_{5-95} regression coefficients (a_2) for formulation *VI* and *VIII* are observed to increase with an increase in normalized average seismic intensity measures ($S_{a,haz}$ and $I_{a,haz}$) of discretized sets. It is observed that an increase in the ratio of the mean S_a (μ_{S_a}) and design hazard S_a ($S_{a,haz} = 0.53\ g$ for Bridge A and $= 0.65\ g$ for Bridge B) and in the ratio of the mean I_a (μ_{I_a}) and design hazard I_a ($I_{a,haz} = 902.60$ cm/sec for Bridge A and $= 841.25$ cm/sec for Bridge B), leads to an increase in regression coeffecients (a_2). The value of a_2 gets as high as 1.5 for *Rot50DE* (Fig. 15c and d) while for *Rot50CDR*, a_2 gets near to 0.6 (Fig. 15a and b).

6. Conclusion

Previous studies have investigated the effects of ground motion duration on engineered building structures; however, there is a scientific gap in quantifying these effects derived from nonlinear analysis of sophisticated 3-dimensional bridge models and explicit statistical measures. Similar studies on quantifying the impact of ground motion duration on building structures have led to a general conclusion that the duration of strong motion arising from crustal ground motions is not impactful, while the significant duration of large subduction ground motions is significant in seismic response assessment of buildings. This research further strengthens the hypothesis through exhaustive statistical reasoning and for bridge structures.

This study is based on a statistical approach of quantifying the impact of the duration of ground motions on two types of three-dimensional Box-Girder Seat-Type (BGST) bridge models. Two conforming and spectrally equivalent groups of 500 bi-directional ground motions originating from crustal sources and subduction sources are used to an-
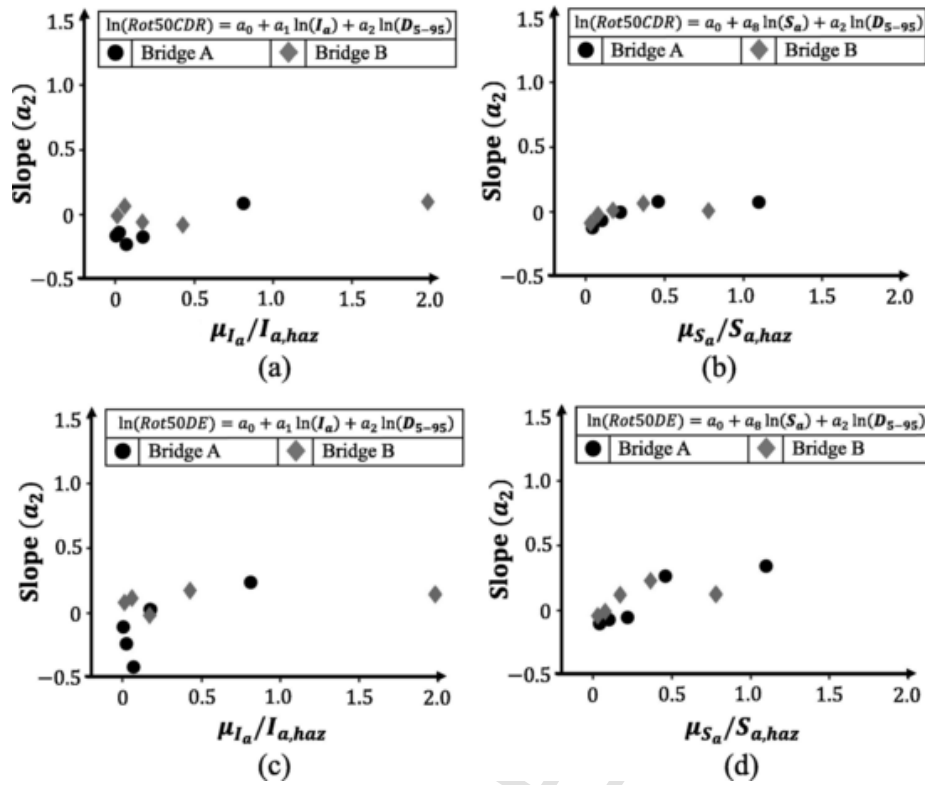
Fig. 12. Regression coefficient (a_2) of D_{5-95} representing the slope for crustal GMs for EDP: a) *Rot50CDR* (using Formulation *VI*), b) *Rot50CDR* (using Formulation *VIII*), c) *Rot50DE* (using Formulation *VI*), and d) *Rot50DE* (using Formulation *VIII*).

alyze the bridge models. The sensitivity of the bridge response is analyzed against 6 ground motion features which include *RotD50* spectral acceleration at the bridge's natural period ($S_a(T_1)$ interchangeably used as S_a), Arias Intensity (I_a), Significant Duration (D_{5-95}), mid-frequency (f_{mid}), time at 30% of cumulated Arias Intensity (t_{mid}) and the derivative of the mid-frequency (f) at t_{mid}. Since the response of the bridges is different in the longitudinal and transverse directions, the two orthogonal components of the ground motions are rotated through 180 degrees with an intercept increment of 9 degrees (21 angles). For each bridge, the median of the 21 obtained column drift ratios for each ground motion is termed as *Rot50CDR* and used as an EDP for feature importance for the ground motions parameters. Since the duration of ground motions is highly correlated with the number of deterioration cycles of the bridge components, another EDP termed as *Rot50DE,* which represents the median dissipated energy, is used in the study. The sensitivity of the bridge response with respect to the six ground motion parameters is obtained through three statistical procedures that include *Step-wise regression*, *Random Forest*, and *Neighborhood Component Analysis* (*NCA*). Nine different formulations of Eq. (19) (given in Table 2) are used to perform the *Step-wise regression* analysis for both groups of ground motions and both types of bridges.

Results indicate that for both types of bridges, in case of ground motions descending from crustal sources, the duration of strong motions can be neglected from seismic response assessment independently of different seismic intensity levels. It is also concluded that among the ground motion features, a single feature I_a is a good estimator of the *Rot50CDR* and *Rot50DE* in the case of crustal ground motions. Furthermore, it is concluded that the current procedure of only using spectral response hazard is a satisfactory methodology since *Rot50CDR* and *Rot50DE* of the bridges are highly associated with the $S_a(T_1)$. Nevertheless, results in the case of ground motions ascending from subduction sources show that the D_{5-95} of the ground motions is an important esti-

mator of the *Rot50CDR* of the bridges. The importance of D_{5-95} of the subduction ground motions is higher in estimated the dissipated energy *Rot50DE* and hence, D_{5-95} cannot be neglected. It is also noticed, unlike the case of crustal ground motions, no single feature can satisfactorily estimate the response of the bridges (i.e., *Rot50CDR* and *Rot50DE*). Among the tested parameters, the combination of I_a and D_{5-95}, and the combination of $S_a(T_1)$ and D_{5-95} can be used in estimating the bridge response. It is concluded that in the case of subduction ground motions, the significance of $S_a(T_1)$ as a predictor of EDPs can be highly improved by adding D_{5-95}; this is particularly important for in collapse assessment where strong ground motions are the main culprits.

It is concluded that duration of subduction strong motions must be given proper attention for ground motion selection, while the duration of crustal ground motions can be neglected in the process. Since the current seismic design philosophies for structures are progressing towards maintaining a uniform risk of collapse, the findings of this study look particularly interesting for bridge engineers and bridge design code developers in places prone to subductions ground motions such as the northwest of the U.S. and Canada. In these places, code calibration and risk assessment is mainly conducted using the NGAWest2 database that is currently dominated with the crustal ground motions. Based on the findings of this study, crustal ground motions clearly lack the impact caused by the significant characteristics of subduction ground motions and the two cannot be used interchangeably. The conclusion of this paper is based on large-scale nonlinear simulations wherein ground motions are studied by virtue of different intensity measures, however, only the two most common bridge structural configuration are analyzed in this study. Further studies, that consider more bridge configurations and randomness in the capacity side are required to widely characterize the results and generalize the conclusions on the differences in the impact of crustal and subduction ground motions.
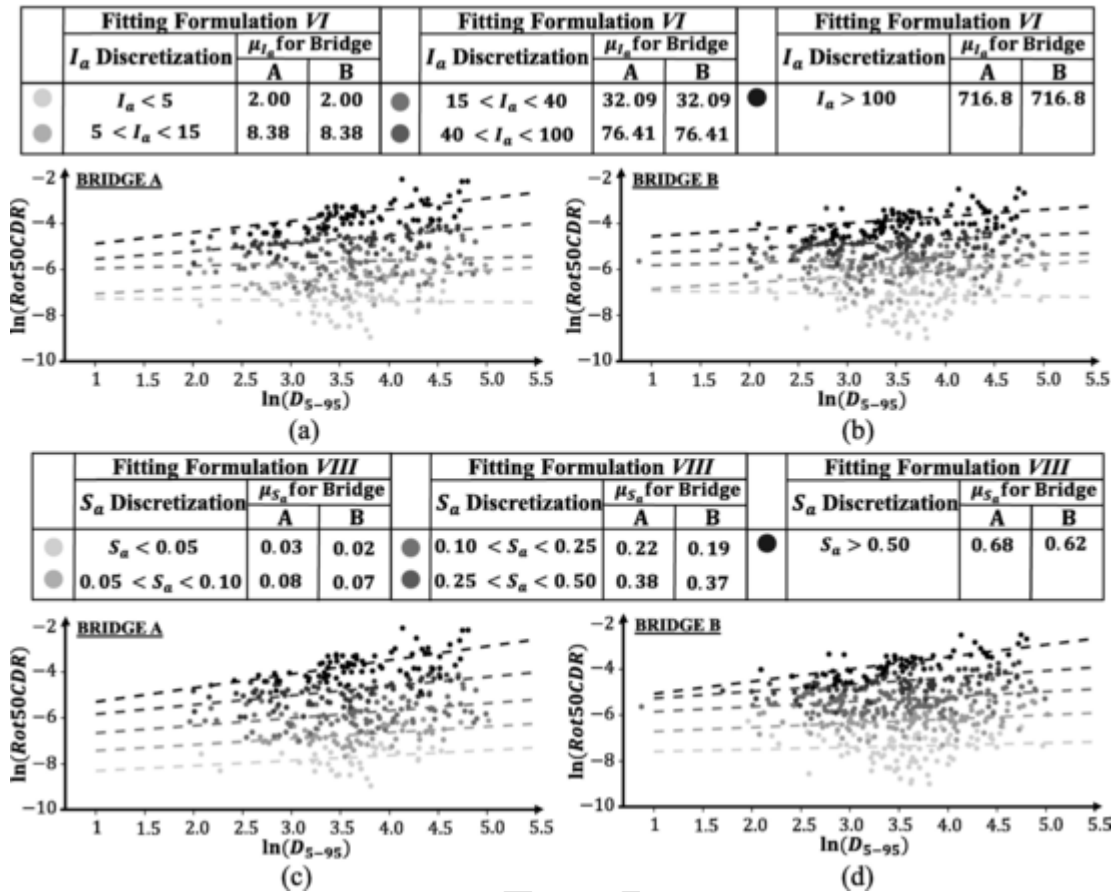
Fitting Formulation *VI*				Fitting Formulation *VI*				Fitting Formulation *VI*		
	I_a Discretization	μ_{I_a} for Bridge			I_a Discretization	μ_{I_a} for Bridge			I_a Discretization	μ_{I_a} for Bridge
		A	B			A	B			A	B
○	$I_a < 5$	2.00	2.00	◐	$15 < I_a < 40$	32.09	32.09	●	$I_a > 100$	716.8	716.8
◔	$5 < I_a < 15$	8.38	8.38	◕	$40 < I_a < 100$	76.41	76.41				

Fitting Formulation *VIII*				Fitting Formulation *VIII*				Fitting Formulation *VIII*		
	S_a Discretization	μ_{S_a} for Bridge			S_a Discretization	μ_{S_a} for Bridge			S_a Discretization	μ_{S_a} for Bridge
		A	B			A	B			A	B
○	$S_a < 0.05$	0.03	0.02	◐	$0.10 < S_a < 0.25$	0.22	0.19	●	$S_a > 0.50$	0.68	0.62
◔	$0.05 < S_a < 0.10$	0.08	0.07	◕	$0.25 < S_a < 0.50$	0.38	0.37				

Fig. 13. Effect of D_{5-95} on *Rot50CDR* using subduction GMs on: a) Bridge A (Formulation *VI*), b) Bridge B (Formulation *VI*), c) Bridge A (Formulation *VIII*), and d) Bridge B (Formulation *VIII*).

CRediT authorship contribution statement

Jawad Fayaz: Methodology, Investigation, Software, Writing - original draft. **Miguel Medalla:** Investigation, Writing - original draft. **Farzin Zareian:** Conceptualization, Writing - review & editing, Supervision.

Uncited references

[29].

Declaration of Competing Interest

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

Acknowledgments

Data Availability Statement

Some or all data, models, or codes that support the findings of this study are available from the corresponding author upon reasonable request.

Appendix A. Supplementary data

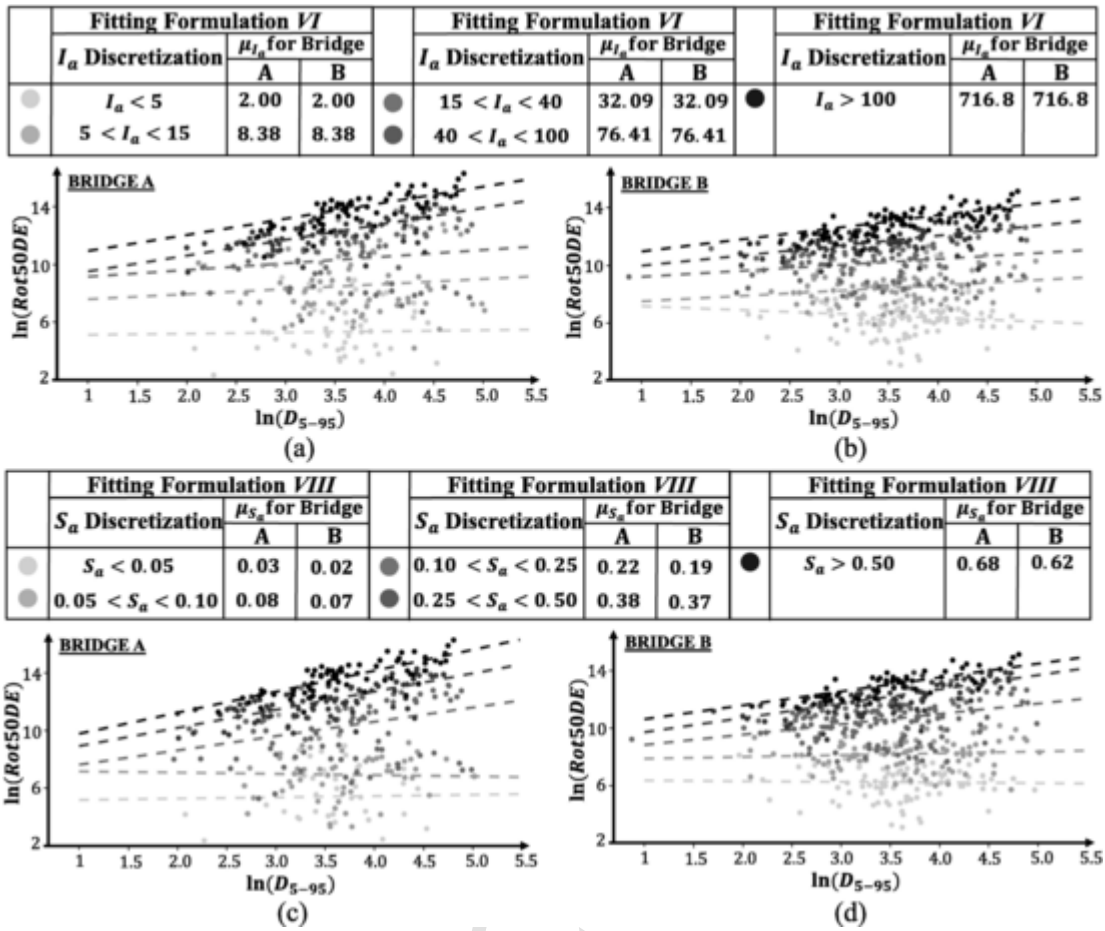Supplementary data to this article can be found online at https://doi.org/10.1016/j.engstruct.2020.110845.

Fitting Formulation *VI*				Fitting Formulation *VI*				Fitting Formulation *VI*		
	I_a Discretization	μ_{I_a} for Bridge			I_a Discretization	μ_{I_a} for Bridge			I_a Discretization	μ_{I_a} for Bridge
		A	B			A	B			A	B
○	$I_a < 5$	2.00	2.00	◔	$15 < I_a < 40$	32.09	32.09	●	$I_a > 100$	716.8	716.8
◑	$5 < I_a < 15$	8.38	8.38	◕	$40 < I_a < 100$	76.41	76.41				

(a)   (b)

Fitting Formulation *VIII*				Fitting Formulation *VIII*				Fitting Formulation *VIII*		
	S_a Discretization	μ_{S_a} for Bridge			S_a Discretization	μ_{S_a} for Bridge			S_a Discretization	μ_{S_a} for Bridge
		A	B			A	B			A	B
○	$S_a < 0.05$	0.03	0.02	◔	$0.10 < S_a < 0.25$	0.22	0.19	●	$S_a > 0.50$	0.68	0.62
◑	$0.05 < S_a < 0.10$	0.08	0.07	◕	$0.25 < S_a < 0.50$	0.38	0.37				

(c)   (d)

Fig. 14. Effect of D_{5-95} on *Rot50DE* using subduction GMs on: a) Bridge A (Formulation *VI*), b) Bridge B (Formulation *VI*), c) Bridge A (Formulation *VIII*), and d) Bridge B (Formulation *VIII*).
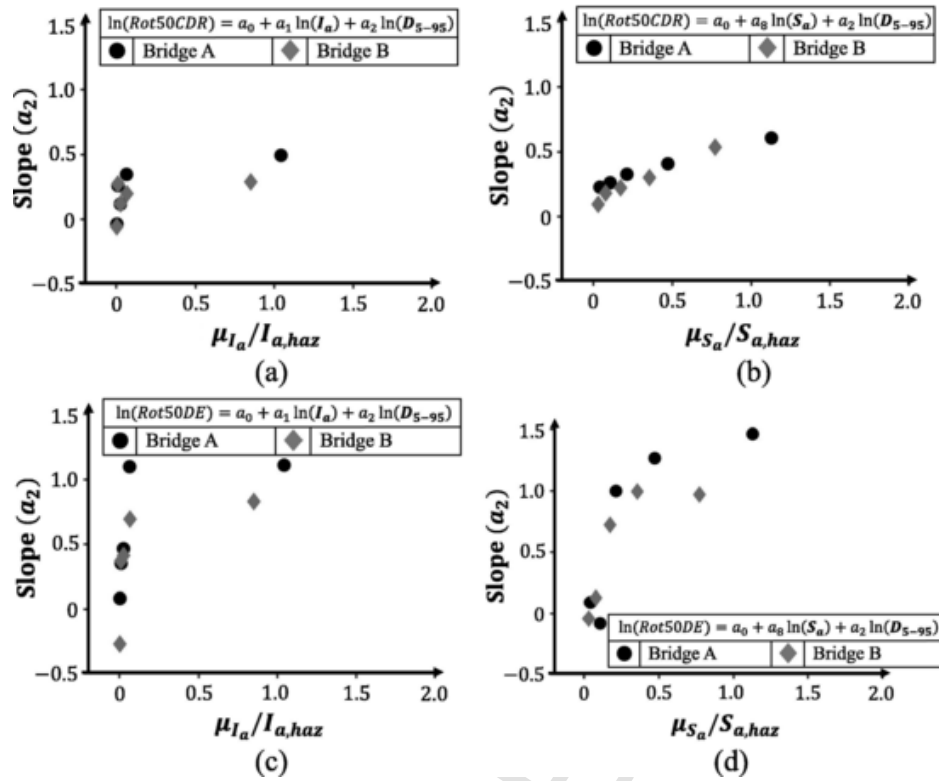
Fig. 15. Regression coefficient (a_2) of D_{5-95} representing the slope for subduction GMs for EDP: a) *Rot50CDR* (using Formulation *VI*), b) *Rot50CDR* (using Formulation *VIII*), c) *Rot50DE* (using Formulation *VI*), and d) *Rot50DE* (using Formulation *VIII*).

References

[1] AASHTO. Guide Specifications for LRFD Seismic Bridge Design. 2nd edition; 2011.

[2] A Arias A measure of earthquake intensity. In: R J Hansen, editor. Seismic Design for Nuclear Power Plants. Cambridge, MA: MIT Press; 1970. p. 438–483.

[3] Akaike H. Information theory and an extension of the maximum likelihood principle. In: 2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Akadémiai Kiadó, p. 267–81.

[4] S Aksoy, R Haralick Feature normalization and likelihood-based similarity measures for image retrieval. Special Issue on Image and Video Retrieval: Pattern Recognition. Lett; 2000.

[5] A Bozorgzadeh, S Megally, J Restrepo, S A Ashford Capacity evaluation of exterior sacrificial shear keys of bridge abutments. Bridge Eng 2006;11(5):555–565.

[6] Kazantzi AK, Vamvatsikos D. A study on the correlation between dissipated hysteretic energy and seismic performance. In: 15th World Conference on Earthquake Engineering (15WCEE), Lisbon, Portugal; 2012.

[7] G Benzoti, T Ohtaki, M J N Pristley, F Seible Seismic performance of circular reinforced concrete columns under varying axial load. Rep. 96/04, Division of Structural Engineering. CA: Univ. of California, San Diego; 1996.

[8] Bolt B. A. (1973). Duration of strong ground motion. 5th World Conference on Earthquake Engineering, Rome, Italy, 1304–1313.

[9] J J Bommer, G Magenes, J Hancock, P Penazzo The Influence of Strong-Motion Duration on the Seismic Response of Masonry Structures. Bull Earthq Eng 2004;2(1):1–26. doi:10.1023/B: BEEE.0000038948. 95616.bf.

[10] Y N Bozorgnia, K W Campbell Ground Motion Model for the Vertical-to-Horizontal (V/H) Ratios of PGA, PGV, and Response Spectra. Earthq Spectra 2016;32(2):951–978.

[11] Caltrans. Seismic Design Criteria. Version 1.7, California Department of Transportation, Sacramento, CA; 2013.

[12] Caltrans. Memo to Designers. California Dept. of Transportation, Sacramento, CA; 2013.

[13] Chandramohan R. Duration of earthquake ground motion: influence on structural collapse risk and integration in design and assessment practice. Ph.D. Thesis, Department of Civil and Environmental Engineering, Stanford University, Stanford, CA; 2016.

[14] E Choi Seismic analysis and retrofit of mid-America bridges Ph.D. Thesis. Atlanta (GA): Department of Civil and Environmental Engineering, Georgia Institute of Technology; 2002.

[15] R W Clough, K L Benuska, E L Wilson Inelastic Earthquake Response of Tall Buildings. II. Auckland, New: Zealand; 1965. p. 68–89.

[16] E Cosenza, G Manfredi In: H Krawinkler, P Fajfar, editors. Seismic Design Methodologies for the Next Generation of Codes. Rotterdam, Netherlands: Balkema; 1997. p. 119–130.

[17] M Dabaghi, A Der Kiureghian Stochastic model for simulation of near-fault ground motions. Earthq Eng Struct Dyn 2017;46(6):963–984.

[18] H Y Fang Foundation Engineering Handbook. second ed. New York: Van Nostrand Reinhold; 1991.

[19] Goldberger J, Hinton G, Roweis S, Salakhutdinov R. Neighbourhood Components Analysis. Adv Neural Inform Process Syst. 17;2005:513–20.

[20] Kottari A. Design and capacity assessment of external shear keys in bridge abutments. Ph.D. Thesis, Department of Structural Engineering, UC, San Diego, CA; 2016.

[21] S K Kunnath, E Erduran, Y H Chai, M Yashinsky Effect of near-fault vertical ground motions on seismic response of highway overcrossings. ASCE J Bridge Eng 2008;13(3):282–290.

[22] Mahin SA. Effects of duration and aftershocks on inelastic design earthquakes. In: 7th World Conference on Earthquake Engineering. Istanbul, Turkey, p. 677–80.

[23] Mantawy A, Anderson J. Assessment of Low-Cycle Fatigue Damage in R.C. Frame Buildings Under Long-Duration Earthquakes. In: SECED 2015 Conference: Earthquake Risk and Engineering towards a Resilient World. Cambridge, UK; 2015.

[24] F McKenna, M H Scott, G L Fenves Nonlinear finite element analysis software architecture using object composition. J Comput Civ Eng 2010;24(1):95–107.

[26] R Omrani, B Mobasher, X Liang, S Günay, K Mosalam, F Zareian, et al. Guidelines for Nonlinear Seismic Analysis of Ordinary Bridges: Version 2.0. Caltrans Final Report 2015;No. 15–65A0454.

[27] Oyarzo-Vera C, Chouw N. Effect of earthquake duration and sequences of ground motions on structural responses. In: 10th International Symposium on Structural Engineering for Young Experts. Changsha, China; 2008.

[28] M Raghunandan, A Liel, N Luco Collapse risk of buildings in the Pacific northwest region due to subduction earthquakes. Earthq Spectra 2015;31:2087–2115.

[29] K Ramanathan Next generation seismic fragility curves for California bridges incorporating the evolution in Seismic design philosophy Ph.D. Thesis. Atlanta (GA): Department of Civil and Environmental Engineering, Georgia Institute of Technology; 2012.

[30] Ruiz-García J, Miranda E. Performance-Based Assessment of Existing Structures Accounting for Residual Displacements. Tech Rep 153. Stanford, CA: John A. Blume Earthq Eng Res Center; 2005.

[31] K Ryan Influence of Vertical Ground Shaking on Design of Bridges Isolated with Friction Pendulum Bearings. Richmond, CA: PEER Researchers' Workshop; 2018.

[32] Sarada K Sarma, N N Ambraseys The Response of Earth Dams to Strong Earthquakes. Géotechnique 1967;17(3):181–213. doi:10.1680/ geot.1967.17.3.181.

[33] Gideon E Schwarz Estimating the dimension of a model. Ann Stat 1978;6(2):461–464. doi:10.1214/aos/1176344136, MR 0468014.

[34] S K Shahi, J W Baker An efficient algorithm to identify strong velocity pulses in multi-component ground motions. Bull Seismol Soc Am 2014;104(5):2456–2466.

[35] Shamsabadi A, Kapuskar M. Nonlinear seismic soil-abutment-structure interaction analysis of skewed bridges. In: Proc 5th National Seismic conference on bridges and highways, San Francisco, CA; 2006.

[36] Shumway and Stoffer Time Series Analysis and its applications, with examples in R. 4th ed. Springer; 2017.

[38] Timothy DA, Robert BD, Jonathan PS, Emel S, Walter JS, Brian S, et al., NGA-West2 Database. Earthq Spectra. August 2014;30(3):989–1005.

[39] M D Trifunac, A G Brady A study on the duration of strong earthquake ground motion. Bull Seismol Soc Am 1975;65(3):581–626.

[40] Veletsos AM, Newmark NM. Effect of inelastic behavior on the response of simple systems to earthquake motions. In: 2nd World Conference on Earthquake Engineering, 2, p. 895–912, Tokyo, Japan, 1960.

[41] W Yang, W Kuanquan, Z Wangmeng Neighborhood Component Feature Selection for High-Dimensional Data. J Comput 2012;7(1).

[42] Scharge I. Anchoring of Bearing by Friction, Publication SP- American Concrete Institute; 1981 p. 197–215.

[43] C M Uang, V V Bertero Evaluation of seismic energy in structures. Earthq Eng Struct Dyn 1990;19(1):77–90.

[44] R Zhu, D Zeng, M R Kosorok Reinforcement Learning Trees. J Am Stat Assoc 2015;110(512):1770–1784.

[45] Caltrans. Bridge Design Specifications Manual," California Department of Transportation;1990.