Structures 32 (2021) 329–341

Contents lists available at ScienceDirect

Structures

journal homepage: www.elsevier.com/locate/structures

Analytical seismic performance and sensitivity evaluation of bridges based on random decision forest framework

Farahnaz Soleimani

School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

ARTICLE INFO

Keywords:
Sensitivity analysis
Seismic demand
Bridge
Seismic performance
Random forest
Machine learning

ABSTRACT

Various bridge portfolios and modeling parameters will influence the seismic response of bridges differently. These features are typically fixed prior to modeling bridges, while there is inherent uncertainty associated with choosing them. Sensitivity analysis of analytical seismic demands with respect to the changing bridge attributes helps to improve estimated seismic demand models which are eventually used in the reliability assessment of bridges. To this end, the current study implements statistical approaches such as analysis of covariance to evaluate the impact of common bridge portfolios such as abutment types on the primary engineering demand parameters such as deck displacement. Moreover, this paper proposes a machine learning algorithm, Random Forest ensemble learning method, to assess the level of importance of modeling parameters on estimating seismic demands. The framework is presented for analyzing concrete box-girder bridges with tall piers that are typically constructed in response to the complex topography of the construction site such as mountain or valley regions. However, the proposed framework is applicable to other types of bridges. Furthermore, although previous research revealed distinctive seismic performance for bridges with tall piers compared to the bridges with ordinary configurations, there is still a lack of understanding of the variability of their seismic demands. Thereby, the findings of this study provide a better understanding of the seismic performance of this class of bridge.

1. Introduction

This study seeks to address two main research gaps: I) investigate how sensitive the bridge seismic demands are to propagating of modeling and analysis features using novel machine learning algorithms, II) evaluate the seismic performance of bridges with various height levels. Although several researchers (e.g., [7,15,21,30,31,34,36,37]) previously investigated the effect of uncertain parameters on the seismic performance of typical bridges, the seismic evaluation of bridges with tall piers is yet to be determined.

Over the recent years, a few studies [3,6,9,28] concentrated on the seismic performance of this particular class of bridges. For example, Ceravolo et al. [6] evaluated the efficiency of design criteria adopted in European code. In this regard, they assessed the seismic responses of tall pier bridges and indicated the formation of multiple plastic hinges along the pier height. In fact, this finding was in contrast with the prior assumption of the plastic hinge formation around the pier base. Likewise, Chen et al. [4] proposed design suggestions for bridges with tall piers based on their findings from shake table tests. Guan et al. [9] investigated the nonlinear behavior of tall piers with lumped plastic hinges. They found a negligible correlation between the column displacement and plastic rotation of piers. Hence, they reported that the displacement at the pier top was an inappropriate index of structural damage for tall piers. On the other hand, Chen et al. [3] found that section curvature ductility is a reliable damage index in the seismic design of tall piers, based on their incremental dynamic analysis. Moreover, Soleimani [28] proposed an optimized framework for developing probabilistic seismic demand models (PSDMs). The optimized models added an additional input independent parameter in the model caring for the most common irregularity parameters existing in the bridge configuration. To produce the PSDM of a bridge, the geometric irregularity parameter such as the column height ratio was implemented in the formulation of PSDM derived for a bridge with similar characteristics and normal column height. However, the sensitivity of the produced PSDMs to the variations of modeling and analysis features has not yet been evaluated.

Previous studies (e.g., [21,32]) revealed that pier height has a significant effect on the bridge performance such as the bridge fragility and

E-mail address: soleimani@gatech.edu.

https://doi.org/10.1016/j.istruc.2021.02.049
Received 29 November 2020; Received in revised form 16 February 2021; Accepted 17 February 2021
Available online 22 March 2021

the isolation performance of the bridge. More recent studies (e.g., [5,29]) demonstrated a noticeable seismic risk of bridges with tall piers as compared to the bridges that have piers of normal height, particularly because of their irregular configuration and the associated increased uncertainties.

To expand the knowledge on the behavior of tall bridges, there is a particular need to explore how the bridge responses vary as a result of implementing different types of bridge attributes in creating finite element models. Several aspects may be considered in the seismic sensitivity investigation of bridges. One aspect is focusing on the numerical parameters that are commonly used to design bridges. For instance, Mackie and Stojadinovic [12] developed probabilistic seismic demand models for 108 samples of California highway bridges subjected to 80 ground motions. The demand sensitivity to the considered design parameters has been analyzed. The sensitivity investigation revealed that the column curvature ductility elevates as the ratio of span to pier height increases, and this column response decreases as the ratio of column diameter to superstructure depth increases. Although under weaker excitations, the column demand was found insensitive to the longitudinal reinforcement ratio, a reduction in ductility demand was observed caused by using higher ratios once stronger motions were applied. Additionally, the effect of adding springs to incorporate abutments into the bridge model versus using roller supports was examined, and the results confirmed the role of abutments in improving bridge strength and reducing demands. Although the variation in the studied design parameters caused the alteration in demand values, the dispersions remain pretty constant.

Likewise, Dukes et al. [7] assessed the effect of design parameters including ratios of the column reinforcement, column height to column dimension, superstructure depth to column dimension, and span length to column height on the columns, abutments, and bearings demands. They found that for the studied two-span integral concrete box girder bridge, all of the tested design parameters excluding the transverse reinforcement ratio contributed significantly to the resulting demands. Rogers and Seo [24] conducted experimental designs to explore the effect of certain design parameters on the sensitivity of produced fragility values of curved I-girder bridges. It was observed that for curved superstructures, shorter spans decrease vulnerability and the sensitivity of bridges to the applied motions. Multi-span bridges were more fragile than single-span bridges. The first and second detected significant parameters were the ratio of span length to the radius of curvature and the bridge width. Bridges with longer deck width or stiffer bents with either shorter heights or larger diameters had higher system fragility values.

The other aspect is the assessment of modeling considerations. As an example, Kunnath et al. [11] investigated the influence that modeling considerations may have on the simulated engineering demands, and approximated degree of damage, and decision variable corresponding to a certain viaduct bridge. They found that the following modeling choices and assumptions have a statistically significant impact on the results: scaling and transformation of applied excitations, soil-structure interaction model, and foundation flexibility.

Another aspect is the investigation of the impact of modeling parameters on the predicted demands. For instance, Padgett and DesRoches [22] determined the modeling parameters of a multi-span simply supported steel girder bridge with the highest influence on the seismic responses. Considered bridges have been retrofitted with restrained cables, elastomeric bearings, steel jackets, and shear keys. Besides, the relative weight of uncertainty associated with the identified parameters has been assessed. The examined parameters include concrete and steel strengths, stiffness of foundation, abutment, and features associated with retrofit measures. The detected most important parameters from regular modeling characteristics were the direction of loading and gross geometry and from retrofit parameters were blocks, stiffness of elastomeric bearing and foundation in addition to the gap between the bearing and keeper plate. Wang et al. [36] carried out fragility-based sensitivity analyses to find out how seismic performance of pile-supported bridges located in liquefiable areas vary based on the changes in fourteen structural and soil parameters. Bridge models were subjected to 40 recorded motions and bearing deformation, column curvature, and pile curvature were monitored as the seismic demand candidates. The axial compressive ratio, column height, and concrete strength were detected as the three most influential parameters for the bearing demand. However, column characteristics including the longitudinal reinforcement, diameter, and height were highlighted in the results because of their significant effect on the column demand. As for the pile foundation demand, the pile diameter influenced results the most. The results indicated steel strength and the reinforcement ratio of column and pile as the least important features for estimating the bearing demand, while regarding the column and pile demands, they identified modulus of elasticity of steel as the least effective parameter. Towards the risk-informed assessment, Zhong et al. [38] proposed an optimization method based on a Gaussian process surrogate model to determine the optimal parameters of the isolation devices leading to minimize risk. To show the feasibility of their proposed framework in the sensitivity analysis of seismic isolation devices, they conducted probabilistic seismic demand analysis [37] of a cable-stayed bridge model.

Several researchers attempted to examine the sensitivity of estimated seismic demands to various structural modeling and design parameters. However, no study considered the impact of complexity in the seismic performance of bridges with tall piers using machine learning algorithms. Besides, the previous related works concentrated on a selected subset of modeling parameters and rarely considered the effect of different bridge portfolios. Although the previous works typically focused on the response of a couple of bridge components, this work attempts to assess a variety of demands corresponding to key bridge components and aims to extend insights toward the sensitivity of bridge responses to the uncertain factors in the bridge portfolio used in simulations. In this regard, it first evaluates the variability in the estimated seismic demands to varying types of the abutment, girder, column cross-section, footing, excitation direction, soil, and spans. Further, the analyzed bridges are classified into separate categories with the guidance of the observed patterns in the variation and similarities in the estimated seismic demands. Moreover, compared to the previous works, machine learning algorithms are applied in this study to perform a more robust sensitivity study on the modeling parameters. Applying random decision forest has many advantages such as providing high accuracy, low-biased model, and preventing overfitted models. In addition, this approach can capture the nonlinear relationship between the variables and responses, while it does not assume any particular distribution for either the response or the considered variables.

Machine learning approaches have been extensively implemented to solve engineering problems particularly because of their robust prediction power and the well-established algorithms for generalization. However, as mentioned earlier, these approaches have rarely been introduced for the sensitivity analysis of bridge seismic demands. The proposed approach has an excellent generalization capability since it can capture the nonlinearity in the data. However, previous literature on sensitivity analysis typically assumed a predefined simple linear relationship between the response and input variables and often performed a sensitivity analysis by evaluating the standard deviation of the regression model while changing the other individual parameters iteratively. The sensitivity analysis procedure and the application of the random forest algorithm are described in the following sections of the paper.

2. Analytical modeling and seismic analysis

2.1. Modeling procedure and characteristics

The class of bridges picked in this study is the multi-span concrete box girder bridge with tall piers. To simplify the presentation and discussion of the results, a set of nomenclatures are used hereafter based on the descriptions in Table 1. Tall bridges are categorized into three different classes based on their level of column height ratio. The ratio is calculated as the average column height of a tall bridge divided by the average column height of a regular bridge [28]. This ratio varies in the following spans: ($min = 1.5, max = 2.5$) for the first class with the low column height ratio, ($min = 2.5, max = 3.5$) for the second class with the medium column height ratio, and ($min = 3.5, max = 4.5$) for the third class with the high column height ratio.

Each bridge listed in Table 1 is modeled three-dimensionally in Opensees [16], which is renowned to be able to represent the nonlinearities in structural and material components in the finite element analyses. This powerful program in the seismic vulnerability assessments follows a "spine" modeling approach by creating the bridge superstructure as a single beam-column frame and making the substructure as a combination of frame and spring elements. Although a comprehensive explanation of the analytical modeling of individual typical bridge components can be found elsewhere [20,23,28], the general approach is illustrated herein.

As shown in Fig. 1, the bridge columns are modeled using fiber cross-section beam-column elements that consist of confined concrete, unconfined concrete, and steel rebar [8,13,17,33]. Nonlinear displacement-based elements are utilized to simulate the nonlinear behavior of the columns once subjected to strong motions. The behavior of the footing in different directions is simulated by adding translational and rotational springs to the columns' base. Rigid components are connecting the columns to the superstructure elastic deck elements. While a set of spring elements [26] represents the behavior of abutment, contact elements [19] simulate the pounding effect between the superstructure and the abutments. Spread and pile foundations are randomly assigned to an equal number of simulations for both piers and abutments. The shear keys and bearings are modeled using the zero-length elements, and Rayleigh damping with a critical damping ratio of 5% is employed. To capture the various sources of uncertainties associated with the bridge modeling, distribution of the bridge characteristics corresponding to the superstructure, foundation, and substructure in addition to the general modeling parameters is provided in Table 2.

Bridges with two different types of rigid-diaphragm and seat abutments are considered in this study. Rigid diaphragm abutment is commonly found in older bridges, while the seat abutment is often used in more recently designed bridges. When subjected to excitation, the rigid abutment quickly engages the backfill soil since the abutment is integrally connected to the superstructure. As a result of the occurring energy dissipation, unseating of the bridge deck rarely happens to this bridge type. In contrast, the seat abutment allows the superstructure to move independently while providing bearing support. Unseating may be observed in this bridge type as a result of the movement of the supersaturate.

Table 1
Description of the nomenclature assigned to the considered tall bridges.

Nomenclature	Bridge type	Abutment type	H_{ratio}	Type of analysis
MSTB-DLD	Multi-span tall bridge	Rigid-diaphragm	Low	Deterministic
MSTB-DLP	Multi-span tall bridge	Rigid-diaphragm	Low	Probabilistic
MSTB-SLD	Multi-span tall bridge	Seat	Low	Deterministic
MSTB-SLP	Multi-span tall bridge	Seat	Low	Probabilistic
MSTB-DMD	Multi-span tall bridge	Rigid-diaphragm	Medium	Deterministic
MSTB-DMP	Multi-span tall bridge	Rigid-diaphragm	Medium	Probabilistic
MSTB-SMD	Multi-span tall bridge	Seat	Medium	Deterministic
MSTB-SMP	Multi-span tall bridge	Seat	Medium	Probabilistic
MSTB-DHD	Multi-span tall bridge	Rigid-diaphragm	High	Deterministic
MSTB-DHP	Multi-span tall bridge	Rigid-diaphragm	High	Probabilistic
MSTB-SHD	Multi-span tall bridge	Seat	High	Deterministic
MSTB-SHP	Multi-span tall bridge	Seat	High	Probabilistic

2.2. Seismic demands of bridge components

Representative bridge models are created by randomly sampling across the uncertain parameters, listed in Table 2, using the Latin Hypercube that is a stratified sampling technique. This technique is widely applied to reduce the number of simulations, particularly when dealing with problems having various levels of uncertainty. The procedure consists of four main steps as follows:

- Step 1. The range of each input random variable is partitioned into N intervals of equal probability based on the cumulative density function, where N is the number of required sample points.
- Step 2. For a specific random variable, each of the N partitions is sampled once to generate one representative random sample from each range.
- Step 3. Steps 1 and 2 are repeated for each input random variable.
- Step 4. The independently generated random samples of the input variables are randomly combined.

Two types of deterministic and probabilistic analyses are performed separately on the bridge samples to simulate seismic responses of various bridge components. In the probabilistic case, the whole set of variables are considered as random parameters to cover a variety of tall bridges. However, in the deterministic case, median values are assigned to the modeling variables except for the column height ratios and the ground motion variabilities which are treated as random variables. Baker's series of ground motions [1] including 160 excitations with two scaling factors are then applied to the generated bridge samples, while the model parameters and ground motions are systematically varied. Table 3 provides the list of monitored engineering demand parameters (EDPs) in this study that are commonly captured in the seismic analysis of bridges. Next, regression lines are fitted to find the relationship between the median captured peak values of EDPs (μ_{DM}) and the ground motion intensity measure (IM). This relationship is known as the probabilistic seismic demand model (PSDM) in the case of probabilistic analysis, as shown in Eq. (1) in a log-transformed space. The regression coefficients are represented by a and b. The process of generating bridge seismic demands is demonstrated in Fig. 2.

$$\ln(\mu_{DM}) = \ln(a) + b\ln(IM) \tag{1}$$

3. Statistical sensitivity analysis procedure

3.1. Analysis of variance and covariance

The significance of several bridge portfolios (Table 4) is assessed in this study by performing an analysis of variance and covariance (ANOVA [18,35] and ANCOVA [10,25]) and the Kruskal-Wallis test. First, for each feature (e.g., Number of spans), sensitivity analysis is conducted for individual engineering demand parameters and the results are interpreted by computing the p-values that specify the differences between the seismic demand models of different categories (e.g., three-spans versus four-spans). Along this comparison process, five different possibilities are tested: the fitted regression lines from one group to the next (1) are the same, (2) have the same means, (3) have separate means, (4) are parallel, and (5) are separate. The p-values derived by ANCOVA determine either the slope or the intercept of the fitted seismic demand models are significantly different. The p-values

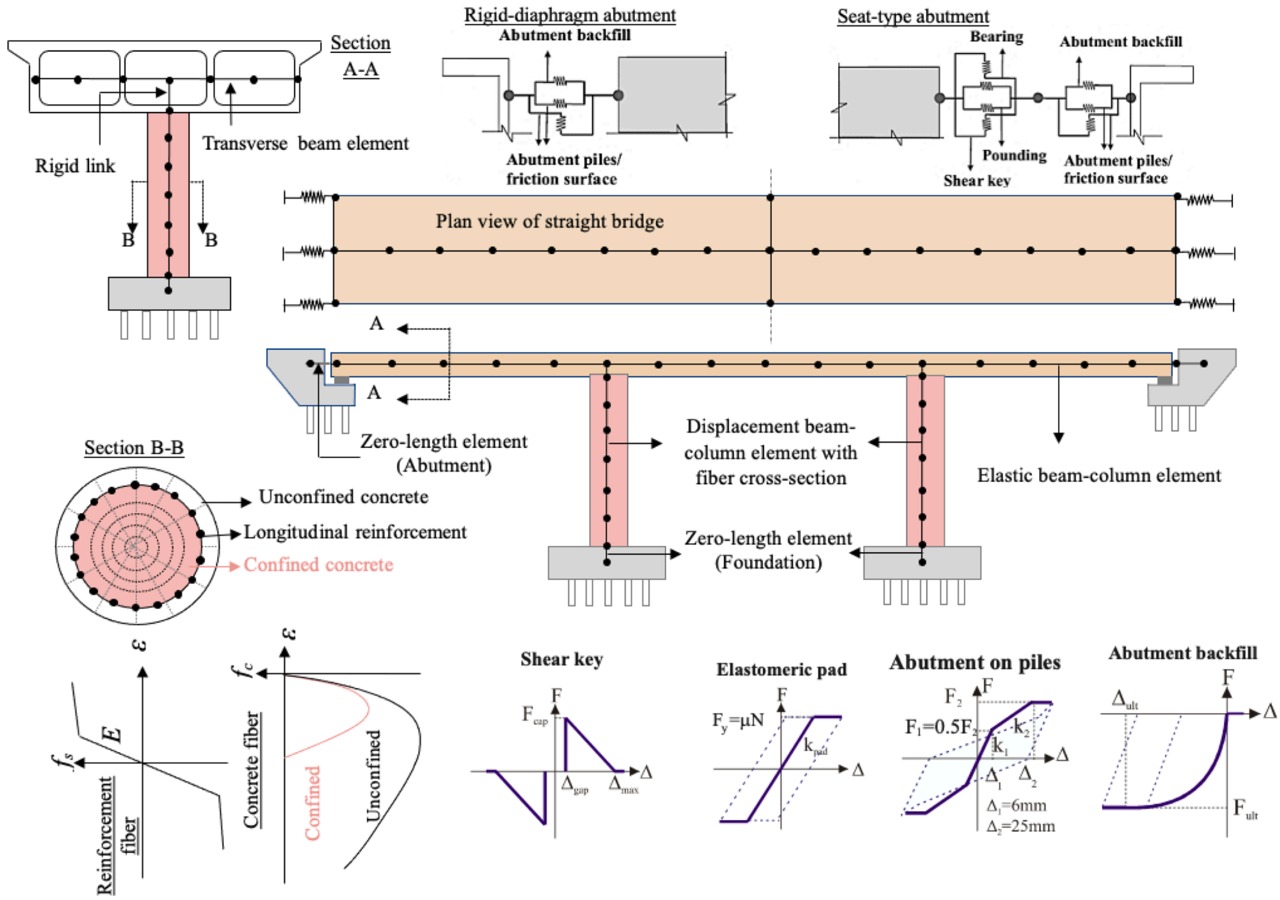

Fig. 1. Illustration of the analytical modeling of considered bridges.

smaller than the significance level of 0.05 express the sensitivity of the corresponding seismic responses to the changing bridge attribute. Conservatively, a bridge with three or more EDPs sensitive to a particular attribute is considered as a bridge with a high level of sensitivity.

Second, multiple comparisons are conducted to check the similarities across engineering demand parameters of various bridge classes. The insights gained from these comparative tests and the total variability in the responses help to investigate the possibility of classifying bridges with roughly similar demands. For this purpose, the ANOVA statistical technique reporting single p-values is performed to investigate differences between the means (μ) associated with two or more independent groups of data via testing the equality hypothesis described as $H_0 : \mu_1 = \mu_2 = \cdots = \mu_m$ that m is the total number of groups. The p-value, as the main criteria of this technique to decide about the acceptance or rejection of this hypothesis, is derived based on the F-statistics:

$$F = \frac{RSS_1 - RSS_2/(k_2 - k_1)}{RSS_2/(n - k_2)} \tag{2}$$

where in this formula RSS_1 and RSS_2 are the total residual sum of squares for the two groups. In this equation, k_1 and k_2 are the degrees of freedom of the residuals for the two groups while n is the total number of data samples. Similar to ANCOVA, if the p-value is less than the predefined significance level of 0.05 the null hypothesis is rejected. While ANOVA elaborates the variability due to the differences among the group means, the Kruskal-Wallis test investigates significant differences between the distribution of responses.

3.2. Random forest algorithm

3.2.1. General framework

Random forest algorithm [2] develops an ensemble of large decision or regression trees each grown from a different randomly sampled subspace of the predictor variables. The randomness helps to decrease overfitting. Regression trees are constructed by considering the whole training sample at the root node and then splitting the data into child nodes sequentially (Fig. 3). At each decision node, the random forest algorithm determines the best splitting feature from the corresponding subspace of variables using splitting criteria:

$$RSS = \sum_{left}(y_i - y_L^*)^2 + \sum_{right}(y_i - y_R^*)^2 \tag{3}$$

where y_L^* and y_R^* are the mean output values for the left and right nodes, respectively. This internally cross-validated approach uses bootstrap aggregation to combine unstable learners and random variable selection which leads to creating uncorrelated individual trees and improves generalization accuracy. The model prediction results are obtained by aggregating the results (i.e., $\theta(.)$) over the large ensemble of grown trees:

$$\hat{f}_{rf}^T(x) = \frac{1}{T}\sum_{t=1}^{T}\theta(x;\phi_t) \tag{4}$$

in which, ϕ_t represents characteristics of random forest t such as split variables and cut points. T denote the total number of grown trees and x represents the input variables.

3.2.2. Unbiased estimates of predictors importance

This study performs a sensitivity analysis based on machine learning

Table 2
Distribution of random variables used to create bridge models.

	Parameter	Unit	Distribution	Mean	Standard deviation
Superstructure	Span length	(m)	Normal	47.24	13.72
	Deck width	(m)	Normal	17.37	2.44
	Ratio of approach-span to main-span length	N/A	Normal	0.75	0.2
Foundation	Translational stiffness	(kN/mm)	Lognormal	175.1	0.44
	Transverse rotational stiffness	(GN-m/rad)	Lognormal	1.36	0.28
	Transverse/longitudinal rotational stiffness ratio	N/A	Lognormal	1.0	1.5
Substructure	Column cross section dimension	(cm)	Randomly assign equal portion of simulations to each		
	Circular		152, 168, 183, 213, 274		
	Rectangular		122x183, 122x244, 168x252, 183x274		
	Longitudinal reinforcement ratio	N/A	Uniform	0.02	0.006
	Confinement ratio	N/A	Uniform	0.009	0.003
	Column height	(m)	Lognormal	7.13	1.15
	Abutment backwall height	(m)	Lognormal	3.59	0.65
	Coefficient of friction of bearing pads	N/A	Normal	0.3	0.1
	Stiffness per deck width	(N/mm/mm)	Lognormal	0.630	0.299
	Transverse gap	(mm)	Uniform	19.1	11.0
	Longitudinal gap	(mm)	Lognormal	23.3	12.4
	Pile stiffness	(kN/mm)	Lognormal	0.124	0.045
General	Concrete compressive strength	(Mpa)	Normal	31.37	3.86
	Reinforcing steel yield strength	(Mpa)	Normal	475.7	37.9
	Mass factor	N/A	Uniform	1.05	0.06
	Damping	%	Normal	0.045	0.0125

*, ** Factors 1 and 2 are used to calculate the mean and standard deviation of normal, lognormal, and empirical distributions as well as the lower and upper bounds of uniform distribution.

Table 3
The list of recorded key engineering demand parameters.

Cases	Engineering demand parameters	Notation	Units
EDP#1	Column curvature ductility	D_{Col}	$1/mm$
EDP#2	Deck displacement	D_{Deck}	mm
EDP#3	Foundation translational displacement	D_{Fnd_trn}	mm
EDP#4	Foundation rotation	D_{Fnd_rot}	rad
EDP#5	Active abutment displacement	D_{Abut_act}	mm
EDP#6	Passive abutment displacement	D_{Abut_pass}	mm
EDP#7	Transverse abutment displacement	D_{Abut_trn}	mm

models of bridge seismic demands developed by the Random Forest algorithm. This sensitivity analysis technique assesses the influence of a given variable on a bridge response under scrutiny. In the initial step, a predictive model is developed using the full set of predictor variables and a selected training set of observations. In the next steps, each variable is permuted, and changes of the model predictive power are measured to investigate the influence of the tested variable on the studied response. In this study, variable importance in predicting the demands is calculated through an out-of-bag permutation approach, since there is a direct correlation between this index value and the model error. A subset of data not used for the initial model construction, the so-

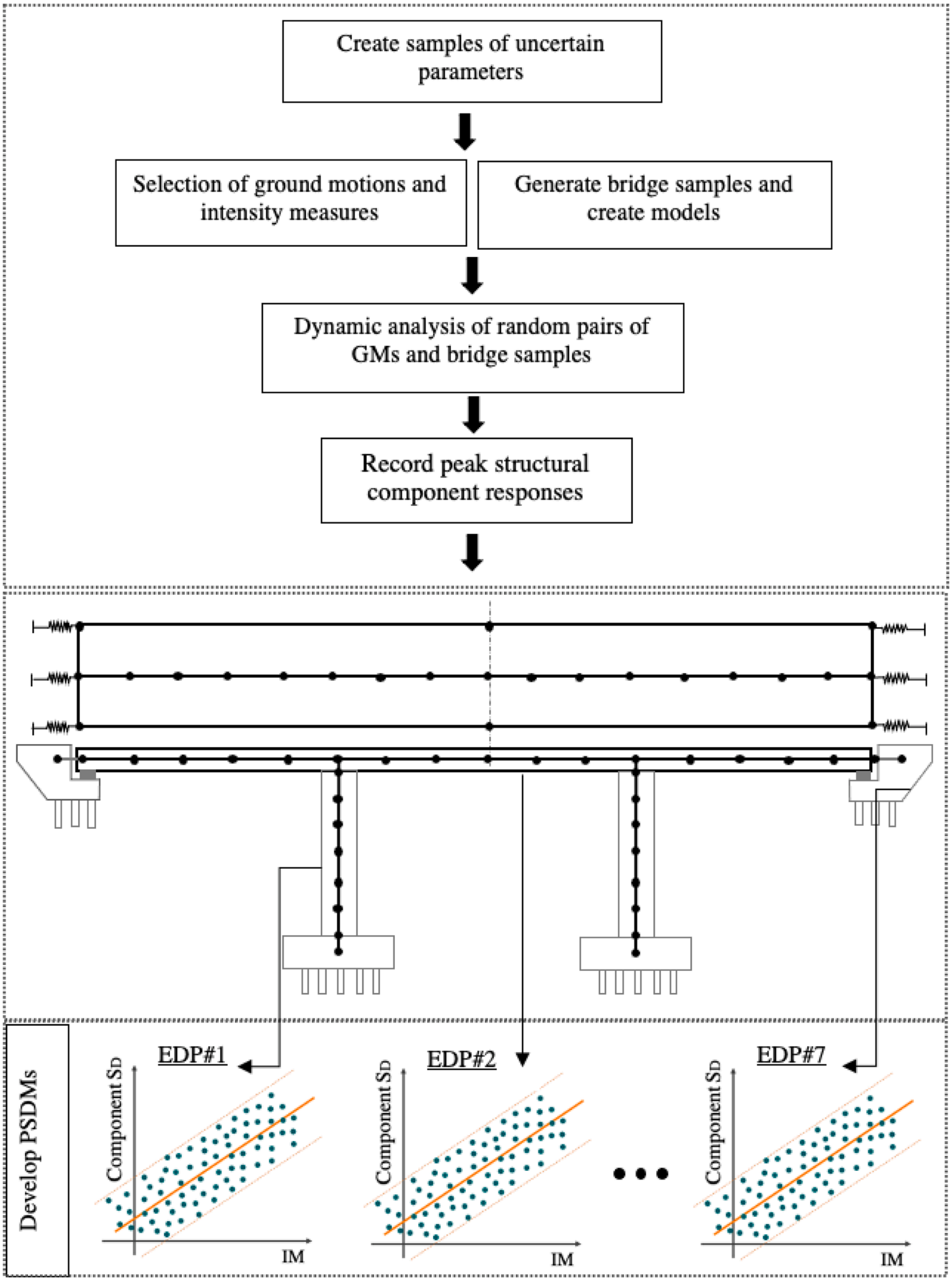

Fig. 2. Demonstration of bridge analysis procedure.

Table 4
The list of bridge modeling and analysis features for the sensitivity study.

Notation	Feature	Scenarios	
NSP	Number of spans	Three	Four
DIR	Direction of excitation	Longitudinal	Orthogonal
ABT	Abutment configuration	On spread footing	On piles
CSS	Column cross-section shape	Circular	Rectangular
FND	Pier foundation configuration	On spread footing	On piles
SCN	Superstructure concrete	Reinforced	Prestressed
ABS	Abutment backfill soil	Clay	Sand

called out-of-bag observations, are used to estimate an unbiased estimate of the predictor's importance. The level of importance is presented by a predictor importance index that its calculation process is described in the following permutation procedure:

- For a particular tree t in a random forest that consists of T learners and p number of predictor variables:
  1. estimate model prediction error (ε) using the out-of-bag data that corresponds to tree t;
  2. randomly permute a given variable $X_i, i = 1, \cdots, p$, while keeping all the other variables the same as step 1; The predictor variables are presented in Table 5;

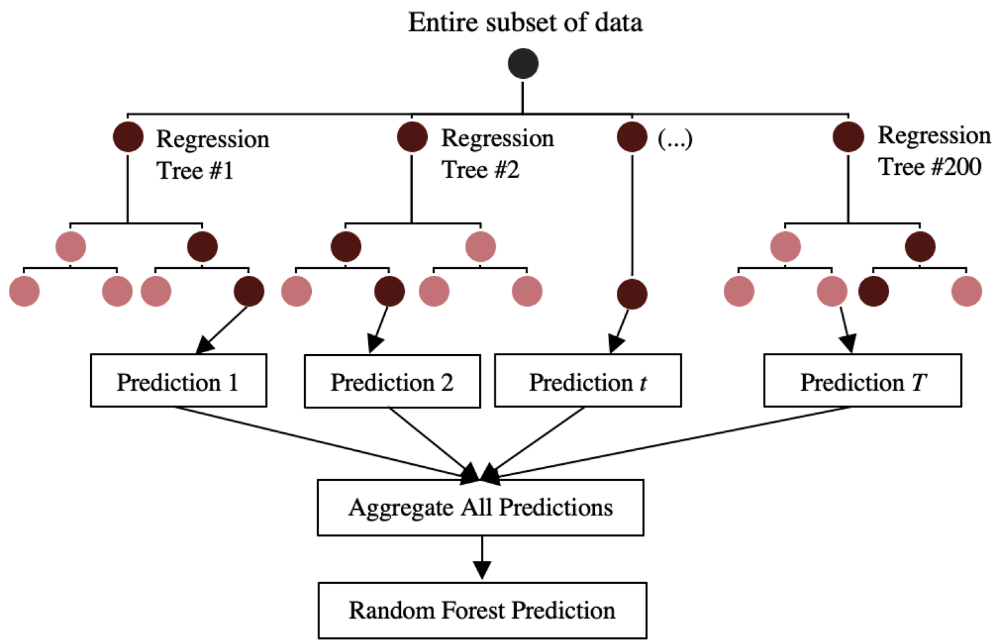

Fig. 3. Schematic procedure of the random decision forest algorithm.

Table 5
Definition of the considered predictor variables.

Common variables in bridges with rigid diaphragm and seat type abutments			
Predictor variable	Definition	Predictor variable	Definition
X1	Ground motion intensity	X11	Foundation rotational stiffness in transvers direction
X2	Span length	X12	Foundation rotational stiffness in longitudinal direction
X3	Column height	X13	Foundation translational stiffness
X4	Deck width	X14	Concrete compressive strength
X5	Number of cells	X15	Steel yield strength
X6	Girder space	X16	Mass factor
X7	Top flange thickness	X17	Damping ratio
X8	Superstructure depth	X18	dt_ground motion
X9	Reinforcement ratio	X19	Span ratio
X10	Abutment backwall height	X20	Column height ratio
Additional variables specific to bridges with seat type abutments			
Predictor variable	Definition	Predictor variable	Definition
X21	Shear key capacity	X23	Shear key gap
X22	Bearing stiffness	X24	Abutment gap

3. estimate the model prediction error (ε_i) using the data with permuted X_i;
4. find the difference between the calculated prediction errors ($d_i = \varepsilon_i - \varepsilon$)
5. repeat steps 2, 3, and 4 for each variable;

- Repeat the previous step over all random decision trees;
- Find the mean and standard deviation of the error variations ($\overline{d}_i, \sigma_i$);
- Compute predictor importance index $= \overline{d}_i / \sigma_i$.

The average variation of the prediction error on the out-of-bag observations, because of permuting a given variable, is a measure of the relevance of that variable. Thereby, higher index values imply a higher level of importance of a variable.

4. Performance-based sensitivity evaluation

4.1. Seismic sensitivity to bridge portfolios

Table 6 presents an example of the calculated p-values based on varying soil types, and Fig. 4 provides a visualization of provided p-values. Each bar in the stacked bar chart (Fig. 4) depicts the p-values corresponding to a specific bridge type (e.g., MSTB-SHP). Each bar is split into a number of sub-bars, each one corresponding to a different EDP, that are represented by different colors. The length of sub-bars is based on the p-values and provides a comparison of p-values between EDPs of a particular bridge. For instance, for the bridge type MSTB-DLP, a p-value of 0.026 is reported for the influence of the number of spans (SPN) on the EDP#1 corresponding to the column curvature ductility, implying that EDP#1 is sensitive to the number of spans in the bridge model. However, applying the ground motions in different directions (DIR) does not make noticeable differences in the column demand of this bridge type as the p-value 0.498 is higher than the considered threshold of 0.05. In order to have a more reliable estimate of the seismic assessment of bridges and to represent the variability in the data points, separate models are recommended to be developed for bridges with a high level of sensitivity to a specific bridge attribute, as displayed in Fig. 5.

Based upon the results, bridges with seat-type abutments indicate less overall sensitivity than bridges with rigid-diaphragm abutments that could be related to the more degree of freedom existing in seat abutments. Bridge responses obtained via deterministic analysis of all bridges with rigid abutments are found sensitive to the number of spans, abutment configuration, and the column cross-section shape.

In the case of probabilistic analysis, all bridges with rigid abutments are sensitive to the backfill soil type. The number of spans, abutment configuration, and the column cross-section shape are identified as the important features in many bridges. Regardless of the analysis type, the footing configuration, and the superstructure concrete type showed a significant influence on a few cases. Besides, the bridge responses are rarely affected by the applied direction of ground motion.

Table 6
The p-values corresponding to the sensitivity analysis of EDPs with respect to the soil type.

Bridge types	Seismic demands						
	EDP#1	EDP#2	EDP#3	EDP#4	EDP#5	EDP#6	EDP#7
MSTB-DLD	0.101	0.436	0.035	0.004	0.001	0.001	0.333
MSTB-DLP	0.457	0.185	0.006	0.027	0.005	0.004	0.264
MSTB-DMD	0.422	0.023	0.003	0.000	0.000	0.000	0.053
MSTB-DMP	0.296	0.506	0.005	0.042	0.000	0.000	0.437
MSTB-DHD	0.541	0.846	0.064	0.066	0.005	0.004	0.573
MSTB-DHP	0.026	0.037	0.168	0.085	0.008	0.007	0.037
MSTB-SLD	0.669	0.767	0.426	0.340	0.604	0.621	0.282
MSTB-SLP	0.410	0.310	0.550	0.697	0.007	0.374	0.628
MSTB-SMD	0.382	0.468	0.644	0.779	0.952	0.614	0.522
MSTB-SMP	0.575	0.701	0.486	0.385	0.694	0.453	0.575
MSTB-SHD	0.461	0.601	0.686	0.863	0.703	0.389	0.878
MSTB-SHP	0.476	0.418	0.347	0.342	0.419	0.476	0.494

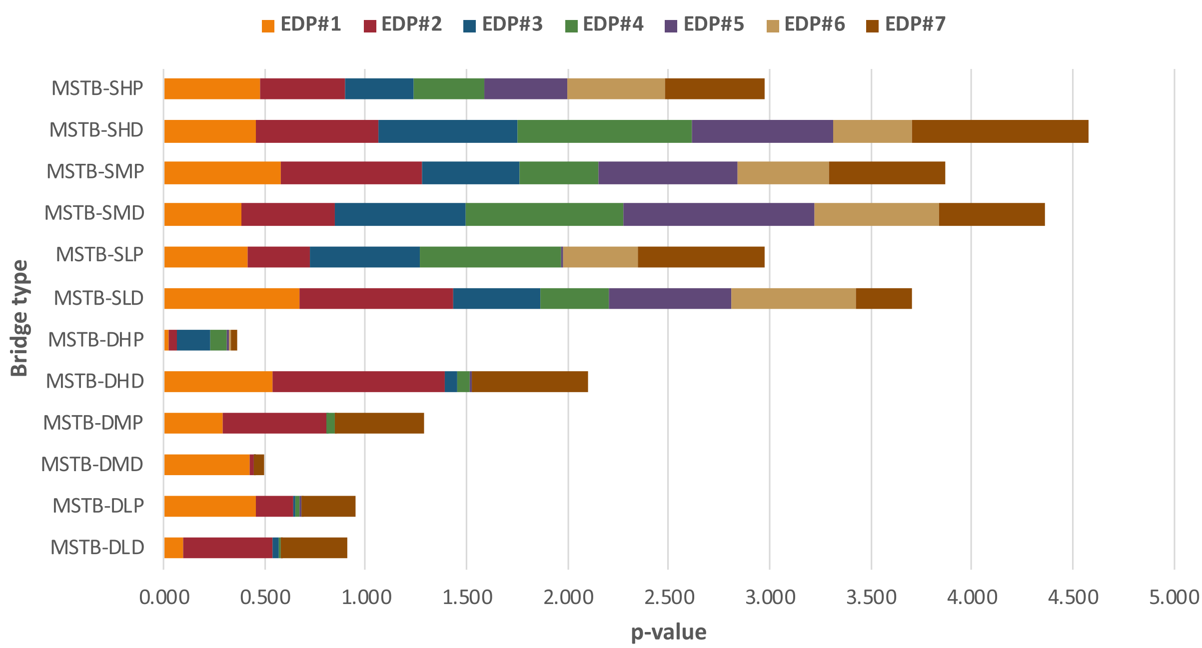

Fig. 4. Visualization of the p-values (reported in Table 6) obtained for the EDPs based on varying soil types.

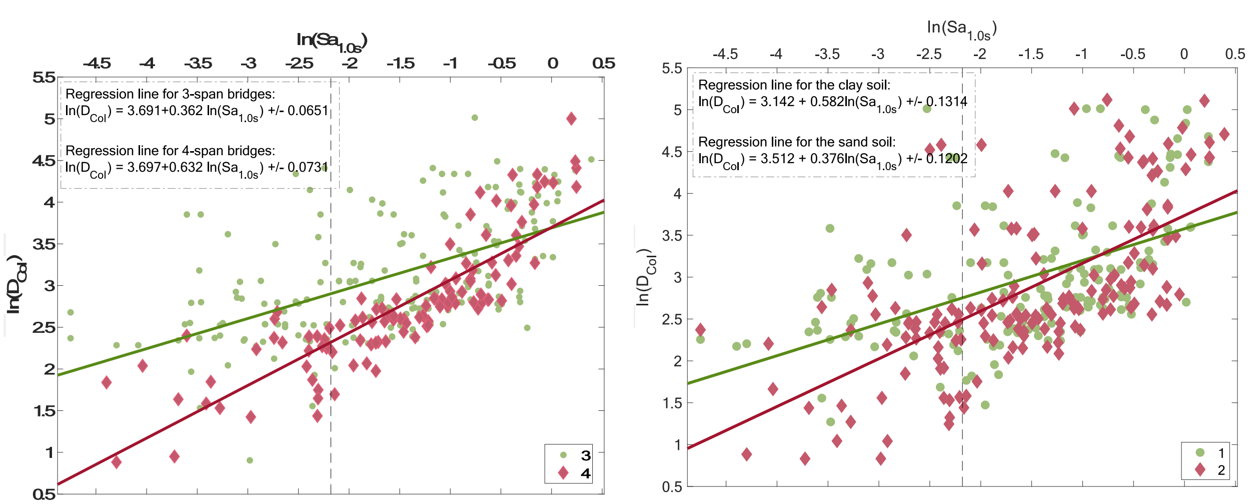

Fig. 5. Regression models developed for the relationship between the column curvature ductility and the ground motion intensity (spectral acceleration at 1 s, $Sa_{1.0s}$); (note: 1 represents clay and 2 sand soil; 3 and 4 represnet three and four number of spans).

Within the aspect of various column height ratios, almost all features except the direction of applied excitation cause significant variation in the monitored responses of tall bridges within the medium range of column heights. Although abutment configuration and column cross-section shape induce noticeable changes in the seismic response of several bridges with the low ratio of the column height, varying the number of spans and the soil type significantly changes EDPs of all bridges in that category. The superstructure concrete type rarely affects bridge responses. For bridges with high column height ratios, abutment configuration, and the column cross-section shape are determined as the most significant features. In some cases, the number of spans and the soil types make considerable changes in the response. However, the EDPs are not mainly influenced by either the foundation configuration or the superstructure concrete type.

4.2. Performance-based classification

A typical challenge in performing the seismic analysis is to balance between a suitable number of simulations and a proper level of uncertainty treatment while taking the computational efforts optimized. This issue is caused by the large variety of bridge attributes that their variation could alter the analysis results notably which eventually impede reliable inference and decision making.

Thereby, a prior understanding of the similarities and dissimilarities of the bridge seismic responses considering different modeling and analysis assumptions is essential. While the previous section improved a sense of how the seismic demand models developed for different bridge components are sensitive to various bridge attributes, this section presents a comparative assessment of the bridge responses and classify bridges with statistically similar response value.

It was noted that the results of the probabilistic analysis of bridges with the high ratio of column heights, and particularly bridges with seat-type abutments, are noticeably distinctive from the other bridges. In general, a common trend is not observed among all EDPs and the similarities and dissimilarities depend on the EDP of interest. Although similar mean values are observed in some cases, very small Kruskal-Wallis p-values are obtained from the evaluation of the response distributions, implying that the distributions of these responses are statistically different. Hence, the findings of the two tests of ANOVA and Kruskal-Wallis are combined to make general conclusions.

Fig. 6 illustrates an example of the results from the comparison of the column curvature ductility of analyzed bridges. According to the performance-based analysis, a potential classification is proposed in Fig. 7 for the studied concrete box-girder bridges with tall piers. The grouping scheme varies based on the EDPs of interest. The estimated column curvature ductility, based on deterministic analysis, of bridges with rigid abutments and various ranges of column heights are assigned to the first group G1. However, bridges with seat-type abutments are categorized under groups G2 and G3. Despite that, for probabilistic assessment of the same EDP, two different groups G4 and G5 are created for rigid and seat abutments.

With respect to the deck displacement in deterministic cases, each bridge is classified separately since their demands were found noticeably different. For the probabilistic cases, the grouping is similar to those of the column curvature ductility. In order to assess the translational displacement of the foundation, potential groupings are observed for the probabilistic analysis of bridges by distinguishing the seismic demand of bridges with a low ratio of the column height from those with the medium and high ratios. To evaluate the foundation rotation based on deterministic analysis, bridges with low and medium height ratios can be assigned to the same group, and bridges with the high ratio are classified individually.

A similar grouping scheme is found for the active and passive responses of the abutment. While bridges with medium and high column height ratios were noted to have a similar pattern in their responses obtained from the deterministic analysis, bridges with low ranges of the ratio displayed distinct performance. The transverse displacement of the abutment of each bridge is found significantly different. However, in probabilistic simulations, three different groups of G7, G8, and G9 are proposed.

The bridges classified into a particular group are observed to have a statistically similar performance during an earthquake. Traditional grouping of bridges was commonly performed by engineering judgment [14] since real-time data is not available for all bridges and yet the analytical seismic analyses of bridges are expensive procedures. Combined with the engineering judgment, statistical techniques yield a more reliable grouping scheme. This performance-based classification approach determines whether there are any significant differences between the means of seismic demands of considered bridges and whether the distribution of these demands is statistically different. The bridges with statistically close mean values and distribution of demands are assigned to a single group and those with statistically significant different responses are assigned to separate groups. Besides, although nonlinear time history analysis is known as the most rigorous approach to simulate the nonlinear performance of bridges, it is often computationally demanding and extremely time-consuming. The proposed grouping scheme has the potential to reduce the extensive simulations since instead of performing nonlinear time history analysis for each bridge, the analysis can be conducted on a representative bridge selected from each group. As a result, the produced PSDMs represent the probabilistic seismic demands of the bridges in the same group. The engineering application of the proposed grouping scheme is similar to the

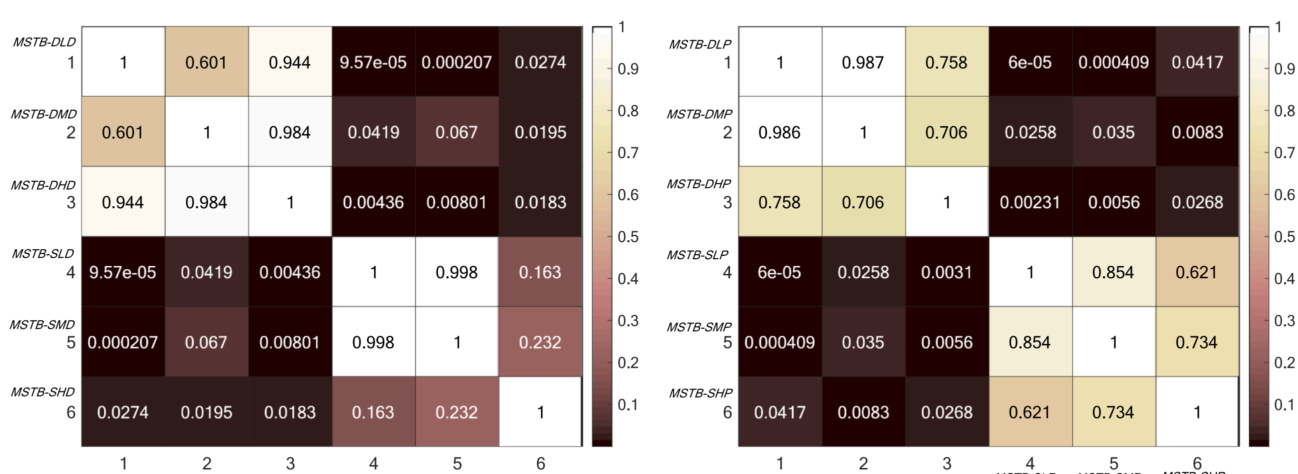

Fig. 6. Comparison of the p-values from the evaluation of the column curvature ductilities from (left) deterministic and (right) probabilistic analyses.

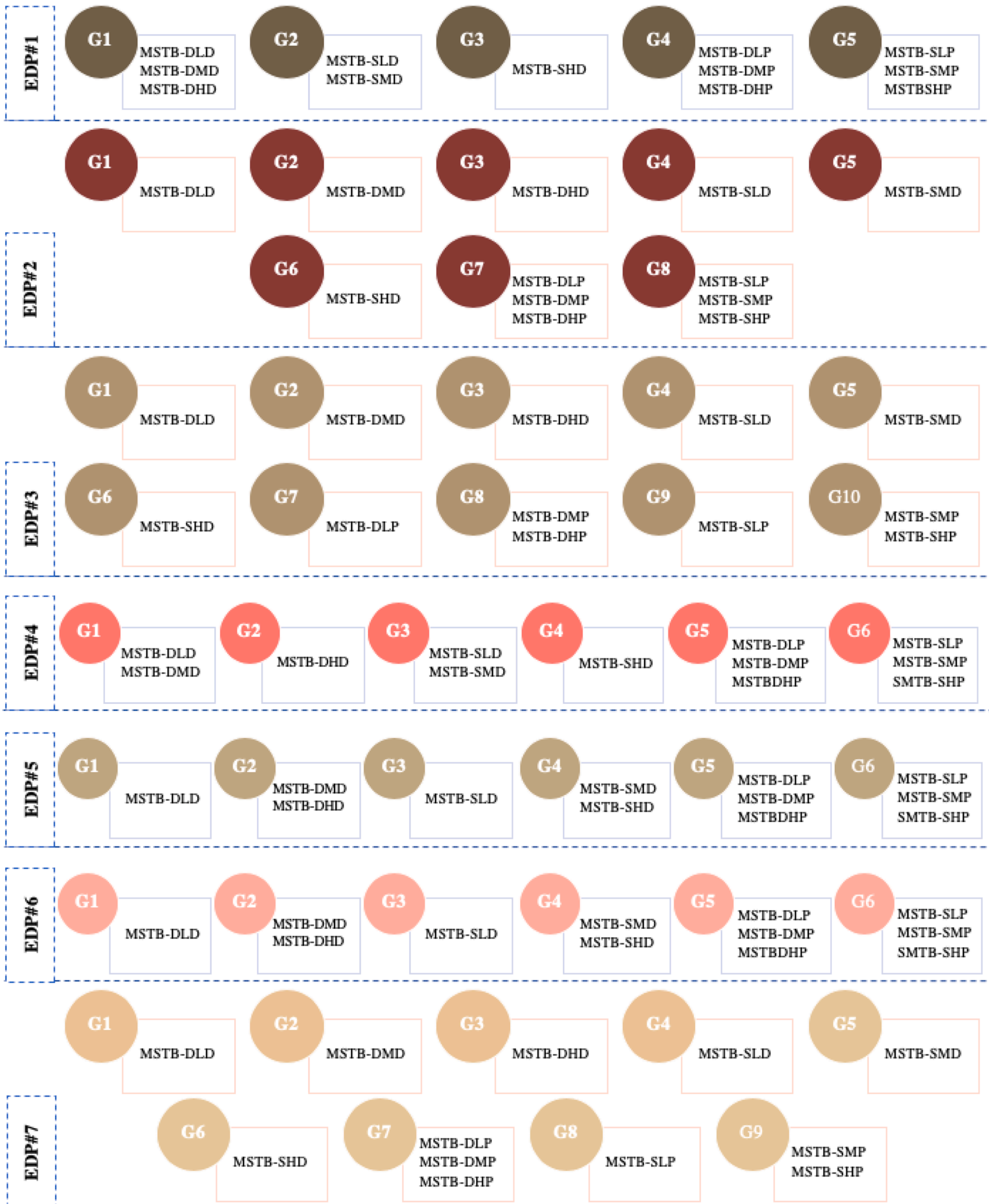

Fig. 7. Classification of concrete box-girder bridges with tall piers based on performance-based analysis.

traditional groupings, though the bridge classification in this study follows a performance-based strategy and is based on the statistical analysis of the simulated seismic demands of bridges.

4.3. Unbiased predictors importance

The aforementioned procedure is conducted for all the EPDs, and the results are presented in Figs. 8 and 9. It is noted that in all cases, Random Forest ranked the ground motion intensity measure as the first and most predictor variable to estimate the bridge seismic demand. That is in agreement with the commonly used probabilistic seismic demand model (PSDM) [12,27] that express the median seismic demand of a bridge component as a function of the ground motion intensity measure.

Furthermore, in the case of bridges with the rigid diaphragm, for the

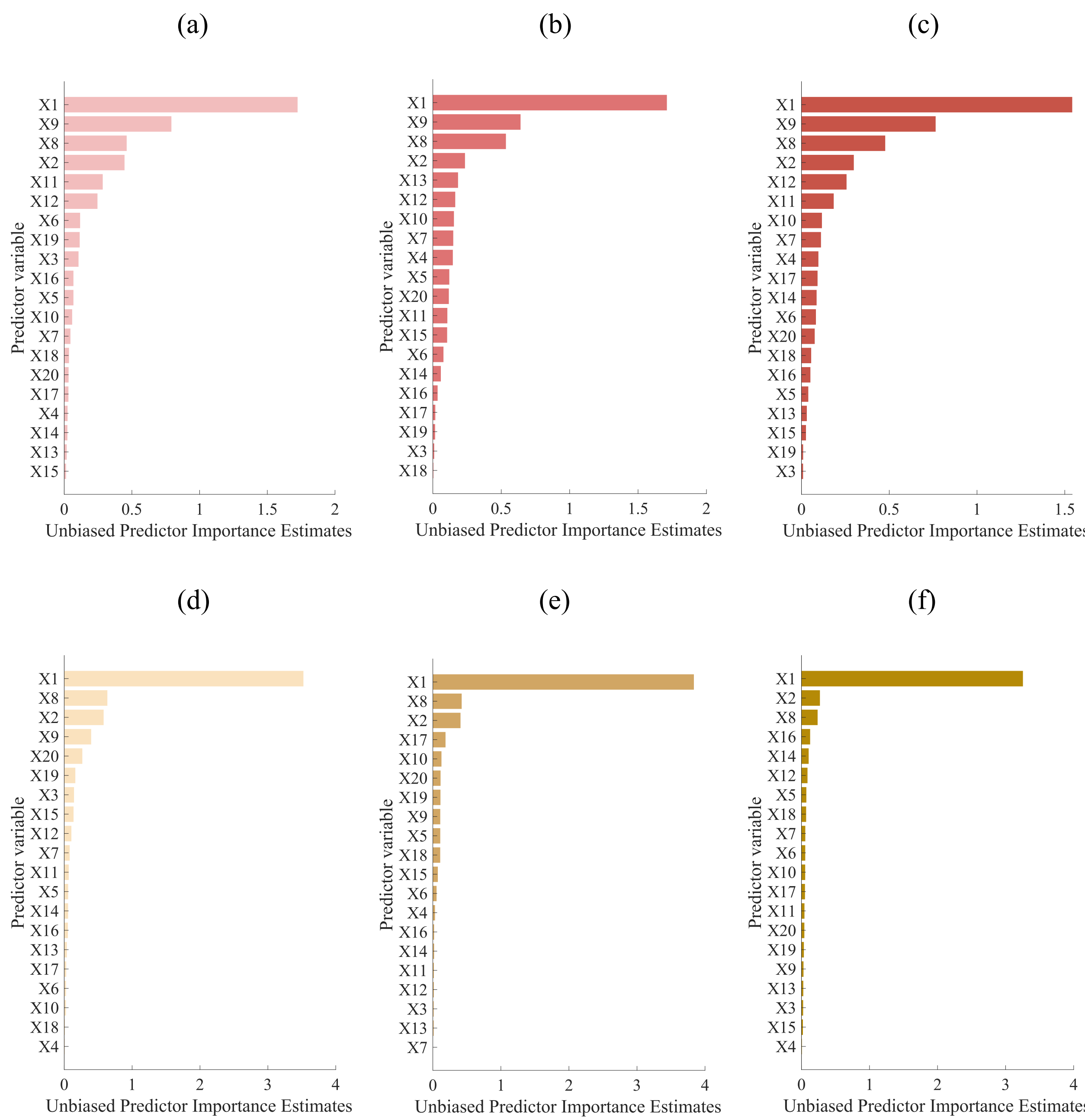

Fig. 8. The level of importance of predictor variables for EDP#1 (a, b, c) and EDP#2 (d, e, f) corresponding to the bridge types: (a) MSTB-DLP; (b) MSTB-DMP; (c) MSTB-DHP; (d) MSTB-DLP; (e) MSTB-DMP; (f) MSTB-DHP.

column curvature ductility (EDP#1) (Fig. 8), common variables that are found in the first top 10 variables are span length, superstructure depth, reinforcement ratio, and foundation rotational stiffness in the longitudinal direction. Regarding the deck displacement (EDP#2), the topmost common variables are detected as span length and superstructure depth. For bridges with rigid diaphragm abutment, the span length was found as one of the top-ranked variables for all EDPs.

In terms of foundation demands (EDP#3 and EDP#4), the topmost variables are shown in Fig. 9. These include reinforcement ratio and foundation translational stiffness, for the translational displacement, and column height, reinforcement ratio, foundation rotational stiffness in longitudinal and transversal directions, for the rotational displacements. For the abutment demands (EDP#5, EDP#6, and EDP#7), the common topmost variables consist of superstructure depth, abutment backwall height, and foundation rotational stiffness.

For bridges with seat type abutments and in terms of the column demand, the common variables are similar to those found in the case of rigid diaphragm abutments. However, for the deck displacement, additional variables such as the column height, deck width, abutment backwall height, and shear key gap are added. The list of top-ranked variables for the foundation demands is somehow similar to those of the rigid abutment cases with the addition of deck width and superstructure depth. Contrary to the other EDPs, the results corresponding to the abutment demands are significantly different than the results of the rigid abutment bridges. In this regard, the most important variables are the deck width, girder space, span ratio, foundation stiffness, shear key capacity, and column height ratio.

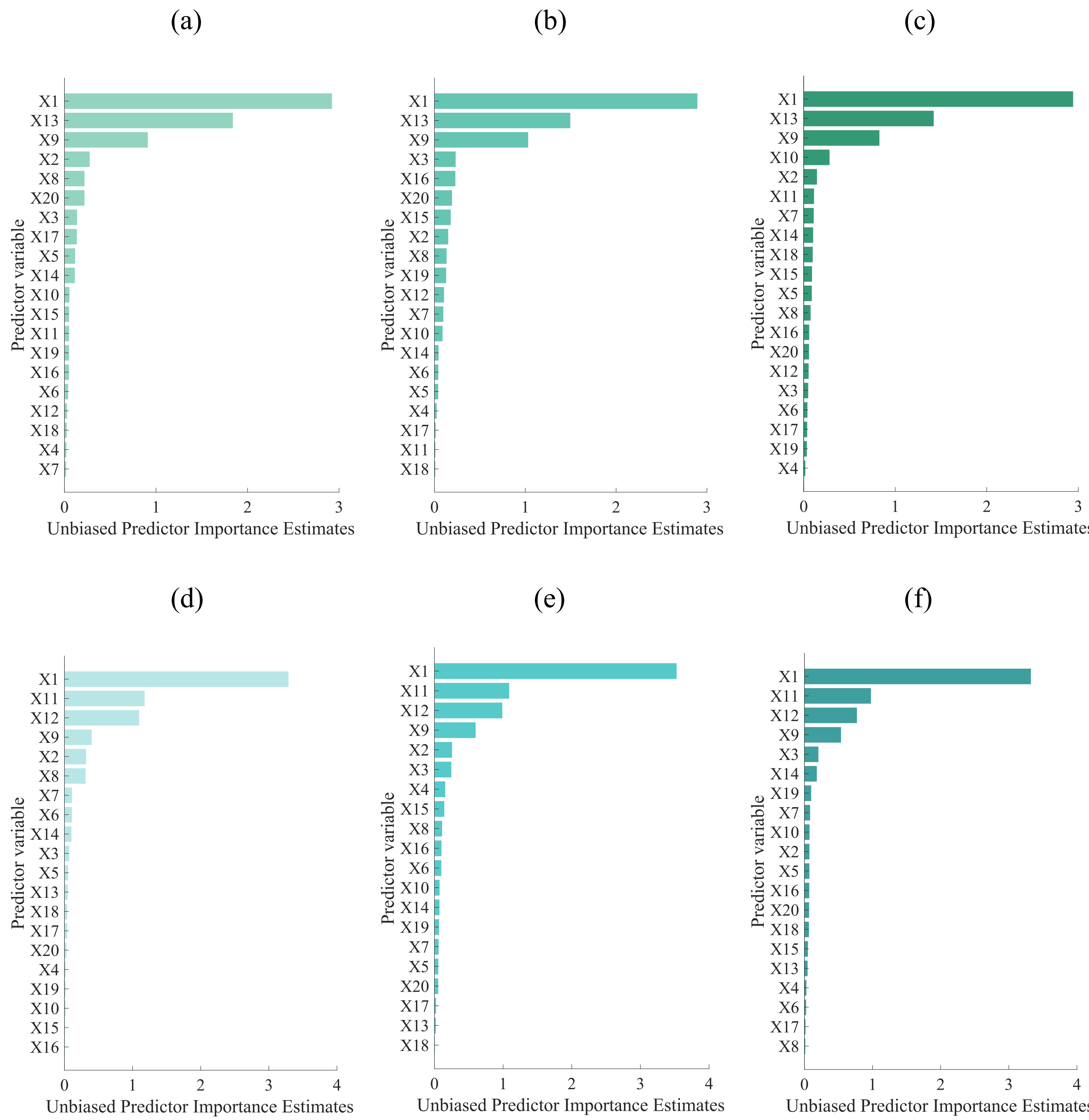

Fig. 9. The level of importance of predictor variables for EDP#3 (a, b, c) and EDP#4 (d, e, f) corresponding to the bridge types: (a) MSTB-DLP; (b) MSTB-DMP; (c) MSTB-DHP; (d) MSTB-DLP; (e) MSTB-DMP; (f) MSTB-DHP.

5. Conclusions

This study presents a sensitivity study via statistical algorithms to identify which modeling and analysis features significantly impact the seismic response estimation of various bridge components. Deterministic and probabilistic seismic analyses of multi-span concrete box-girder bridges with rigid-diaphragm and seat-type abutments are conducted. The primary engineering demand parameters are recorded by performing nonlinear time history analysis of considered bridges. Then, statistical approached including the analysis of covariance are implemented on the monitored bridge responses to test the fitness of single or multiple regression models to the pairs of seismic demands and intensity measures of the applied ground motions. Through this analysis, insight is provided on understanding the sensitivity of seismic demands of bridges with tall piers to a variety of attributes, which to date has not been thoroughly assessed.

The sensitivity analysis results indicated that bridges with seat-type abutments display less overall sensitivity than bridges with rigid-diaphragm abutments. Bridge responses obtained via deterministic analysis of all bridges with rigid abutments are found sensitive to the number of spans, abutment configuration, and the column cross-section shape. It is concluded that bridge responses are rarely affected by the applied direction of ground motion, foundation configuration, and superstructure concrete type. The effect of varying other considered features on the variability in the estimated demands depends on the EDPs. In the case of probabilistic analysis, all bridges with rigid abutments are found sensitive to the backfill soil type. The number of spans, abutment configuration, and the column cross-section shape is identified as the important features in many bridges. In terms of column height ratios, almost all features except the direction of applied excitation cause significant variation in the monitored responses of tall bridges that their column heights are within the medium range. For bridges with high column height ratios, abutment configuration, and the column cross-section shape are determined as the most significant features.

Moreover, potential classification for the investigated bridge classes is also proposed based on the results of the seismic performance assessment and the findings of the two tests of ANOVA and Kruskal-Wallis. The classification provides insight regarding the statistical similarities and dissimilarities of responses. The grouping scheme varies based on the EDPs of interest. For example, the estimated column curvature ductility, based on deterministic analysis, of bridges with rigid abutments and various ranges of column heights are assigned to one group. Besides, the classification depends on the column height ratios. As an example, considering the foundation rotation based on deterministic analysis, bridges with low and medium height ratios can be assigned to the same group, and bridges with the high ratio are classified separately.

This study aims to assess a variety of demands corresponding to key bridge components towards extending insights on the sensitivity of bridge responses to various bridge attributes. An efficient machine learning framework, an ensemble of large decision or regression trees, is proposed in this study to perform a more robust sensitivity study on the modeling parameters. By applying the random forest algorithm, the influence of a given variable on a bridge response is provided through a sequential process by permuting each variable and monitoring the variation of the predictive power. The computed predictor importance measures indicated the ground motion intensity measure as the first and most predictor variable to estimate the bridge seismic demands. Furthermore, for bridges with rigid diaphragm abutment, the span length was found as one of the top-ranked variables for all EDPs. For bridges with seat type abutments and in terms of the column, deck, and foundation demands, the common variables are somehow similar to those found in the case of bridges with rigid diaphragm abutments, with the addition of few variables such as deck width. Contrary to the other EDPs, the results corresponding to the abutment demands are found significantly different than the results of the rigid abutment bridges.

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.

References

[1] Baker JW, Lin T, Shahi SK, Jayaram N. New ground motion selection procedures and selected motions for the PEER transportation research program. Pacific Earthquake Engineering Research Center; 2011.
[2] Breiman L. Random forests. Mach Learn 2001;45(1):5–32.
[3] Chen X, Li J, Liu X. Seismic performance of tall piers influenced by higher-mode effects of piers. J Tongji Univ (Nat Sci) 2017;45(2):159–66.
[4] Chen Xu, Guan Z, Li J, Spencer BF. Shake table tests of tall-pier bridges to evaluate seismic performance. J Bridge Eng 2018;23(9):04018058. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001264.
[5] Chen Xu, Li J, Guan Z. Fragility analysis of tall pier bridges subjected to near-fault pulse-like ground motions. Struct Infrastruct Eng 2020;16(8):1082–95.
[6] Ceravolo R, Demarie GV, Giordano L, Mancini G, Sabia D. Problems in applying code-specified capacity design procedures to seismic design of tall piers. Eng Struct 2009;31(8):1811–21.
[7] Dukes J, DesRoches R, Padgett J. Sensitivity study of design parameters used to develop bridge specific fragility curves. Proc 15th World Conf Earthquake Eng 2012;720.
[8] Filippou FC, Popov EP, Bertero VV. Effects of bond deterioration on hysteretic behavior of reinforced concrete joints. Washington DC: SAC Joint Venture; 1983.
[9] Guan Z, Li J, Xu Y, Lu H. Higher-order mode effects on the seismic performance of tall piers. Front Architect Civil Eng China 2011;5(4):496–502.
[10] Huitema B. The analysis of covariance and alternatives: Statistical methods for experiments, quasi-experiments, and single-case studies, Vol. 608. John Wiley & Sons; 2011.
[11] Kunnath SK, Larson L, Miranda E. Modelling considerations in probabilistic performance-based seismic evaluation: case study of the I-880 viaduct. Earthquake Eng Struct Dyn 2006;35(1):57–75.
[12] Mackie K, Stojadinović B. Probabilistic seismic demand model for California highway bridges. J Bridge Eng 2001;6(6):468–81.
[13] Mander JB, Priestley MJN, Park R. Theoretical stress-strain model for confined concrete. J Struct Eng ASCE 1988;114(8):1804–26.
[14] Mangalathu S, Soleimani F, Jeon J-S. Bridge classes for regional seismic risk assessment: improving HAZUS models. Eng Struct 2017;148:755–66.
[15] Mangalathu S, Soleimani F, Jiang J, DesRoches R, Padgett JE. Sensitivity of fragility curves to parameter uncertainty using Lasso regression. In: Proceedings of the 16th world conference on earthquake engineering, Santiago de Chile, Chile, Paper (Vol. 135); 2017b.
[16] Mazzoni S, McKenna F, Scott MH, Fenves GL. OpenSees command language manual. Pacific Earthquake Engineering Research (PEER) Center; 2006.
[17] Mengotto M, Pinto P. Method of analysis for cyclically loaded RC frames including changes in geometry and non-elastic behaviour of elements under combined normal force and bending. IABSE congress reports of the working commission. 1973.
[18] Miller Jr RG. Beyond ANOVA: basics of applied statistics. CRC Press; 1997.
[19] Muthukumar S, DesRoches R. A Hertz contact model with non-linear damping for pounding simulation. Earthquake Eng Struct Dyn 2006;35(7):811–28.
[20] Nielson B. Analytical fragility curves for highway bridges in moderate seismic zones. Doctoral dissertation. Georgia Institute of Technology; 2005.
[21] Noori HZ, Amiri GG, Nekooei M, Zakeri B. Seismic fragility assessment of skewed MSSS-I girder concrete bridges with unequal height columns. J Earthquake Tsunami 2016;10(01):1550013. https://doi.org/10.1142/S179343111550013X.
[22] Padgett JE, DesRoches R. Sensitivity of seismic response and fragility to parameter uncertainty. J Struct Eng 2007;133(12):1710–8.
[23] Ramanathan K. Next generation seismic fragility curves for California bridges incorporating the evolution in seismic design philosophy. Doctoral dissertation. Georgia Institute of Technology; 2012.
[24] Rogers LP, Seo J. Vulnerability sensitivity of curved precast-concrete I-girder bridges with various configurations subjected to multiple ground motions. J Bridge Eng 2017;22(2):04016118. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000973.
[25] Rutherford A. ANOVA and ANCOVA: a GLM approach. John Wiley & Sons; 2011.
[26] Shamsabadi A, Yan L. Closed-form force-displacement backbone curves for bridge abutment-backfill systems. Proceedings of the geotechnical earthquake engineering and soil dynamics IV congress. 2008.
[27] Shome N, Cornell CA, Bazzurro P, Carballo JE. Earthquakes, records, and nonlinear responses. Earthquake Spectra 1998;14(3):469–500.
[28] Soleimani F. Fragility of California bridges-development of modification factors. Doctoral dissertation. Georgia Institute of Technology; 2017.
[29] Soleimani F. Propagation and quantification of uncertainty in the vulnerability estimation of tall concrete bridges. Eng Struct 2020;202:109812. https://doi.org/10.1016/j.engstruct.2019.109812.
[30] Soleimani F, Mangalathu S, DesRoches R. Seismic resilience of concrete bridges with flared columns. Procedia Eng 2017;199:3065–70. https://doi.org/10.1016/j.proeng.2017.09.417.
[31] Soleimani F, Mangalathu S, DesRoches R. A comparative analytical study on the fragility assessment of box-girder bridges with various column shapes. Eng Struct 2017;153:460–78. https://doi.org/10.1016/j.engstruct.2017.10.036.
[32] Soleimani F, Vidakovic B, DesRoches R, Padgett J. Identification of the significant uncertain parameters in the seismic response of irregular bridges. Eng Struct 2017;141:356–72.
[33] Soleimani F, McKay M, Yang CW, Kurtis KE, DesRoches R, Kahn L. Cyclic testing and assessment of columns containing recycled concrete debris. ACI Struct J 2016;113(5):1009. https://doi.org/10.14359/51689024.
[34] Soleimani F, Yang CSW, DesRoches R. The Effect of Superstructure Curvature on the Seismic Performance of Box-Girder Bridges with In-Span Hinges. Structures Congress 2017 2017:469–80.
[35] Tabachnick BG, Fidell LS. Experimental designs using ANOVA. Belmont, CA: Thomson/Brooks/Cole; 2007. p. 724.
[36] Wang X, Ye A, Ji B. Fragility-based sensitivity analysis on the seismic performance of pile-group-supported bridges in liquefiable ground undergoing scour potentials. Eng Struct 2019;198:109427. https://doi.org/10.1016/j.engstruct.2019.109427.
[37] Zhong J, Jiang L, Pang Y, Yuan W. Near-fault seismic risk assessment of simply supported bridges. Earthquake Spectra 2020;36(4):1645–69.
[38] Zhong J, Wan H-P, Yuan W, He M, Ren W-X. Risk-informed sensitivity analysis and optimization of seismic mitigation strategy using Gaussian process surrogate model. Soil Dyn Earthquake Eng 2020;138:106284. https://doi.org/10.1016/j.soildyn.2020.106284.