16th World Conference on Earthquake, 16WCEE 2017

Santiago Chile, January 9th to 13th 2017

Paper N° 135

Registration Code: S-Z1462212795

SENSITIVITY OF FRAGILITY CURVES TO PARAMETER UNCERTAINTY USING LASSO REGRESSION

S. Mangalathu[(1)], F. Soleimani[(2)], J. Jiang[(3)], R. DesRoches[(4)], J.E. Padgett[(5)]

(1) Graduate student, School of Civil & Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA, sujithmss@gatech.edu
(2) Graduate student, School of Civil & Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA, farahnazsoleimani@gmail.com.
(3) Associate Professor, Department of Civil Engineering, Zhejiang University City College, Hangzhou, Zhejiang, P.R. China, jiangjq@zucc.edu.cn
(4) Karen and John Huff School Chair and Professor, School of Civil & Environmental Engineering, Georgia Institute of Technology, Atlanta, USA, reginald.desroches@ce.gatech.edu.
(5) Associate Professor, Department of Civil and Environmental Engineering, Rice University, Houston, Texas, USA, jamie.padgett@rice.edu.

Abstract

Fragility curves are one of the essential tools in seismic risk assessment. Various source of uncertainties such as material, geometric and ground motion have to be considered in the generation of fragility curves for regional risk assessment of bridges. The identification of parameters that have a significant effect on the seismic demand or fragility curves often requires computationally expensive and time-consuming screening methods. This paper proposes an approach to identify the significant variables using an advanced regression technique called Lasso regression. The proposed approach helps to identify and remove the less significant parameters during the generation of fragility curves without any reiteration. The proposed approach is demonstrated in this paper through the case study of a two span box girder bridge with rigid diaphragms. Although the approach is illustrated with one specific bridge type, the methodology and approach is relevant to other structural systems.

Keywords: fragility curves, sensitivity, Lasso regression, box-girder bridges.

1. Introduction

Recent earthquakes have attracted considerable attention amongst the researchers for the seismic vulnerability assessment of bridges. An effective and increasingly popular technique to determine effects of ground motions on various bridge system components is fragility analysis. A fragility curve is defined as a conditional probability that gives the likelihood that a structure or component will meet or exceed a certain level of damage for a given ground motion intensity (IM).

Uncertainties such as geometric, material or component response parameters exist in a bridge port-folio due to structure-to-structure variation [1]. The source of uncertainties can be either due to lack of knowledge (epistemic) or due to inherent randomness (aleatoric). The number of parameters which are potentially variable in the bridge portfolios for regional risk assessment is very high [2,3]. Hence, the high number of uncertain parameters necessitates a sensitivity study to identify the influence of various modeling parameters and uncertainties on the seismic response of bridges. Such a study will provide insight in quantifying whether the variation of uncertain parameters should be treated explicitly or to be neglected.

There have been a number of studies in the past to evaluate the seismic sensitivity of bridge components (mainly for columns and bearings) to input parameter variation [4, 5] through design of experiments. However, such studies are often challenged with selecting a prudent level of uncertainty treatment while balancing the simulation and computational effort [1]. One of the objectives of the current study is to identify critical modeling parameters whose variability has a significant influence on the seismic bridge fragilities during the generation of seismic demand models or fragility curves with less computational effort. The identification of critical modeling parameters is achieved through the use of an advanced regression technique called Lasso regression [6, 7]. Lasso regression has the advantage of setting the non-significant coefficients to zero during the regression analysis, i.e., it performs variable selection. Also, it is more stable and computationally feasible for high-dimensional data [8].

The current study follows the generation of fragility curves through multi-parameter demand models [9-11]. The multi-parameter demand model has the ability to address the effect of uncertainty on the fragility curves and to incorporate field instrumentation data [10, 12]. The current study differs from the previous studies in using Lasso regression in the generation of fragility curves. The proposed approach is demonstrated in the current study through the case study of a two span box girder bridges with rigid diaphragms. Although the method is illustrated with one specific bridge type (concrete box-girder bridge), the proposed method is relevant and applicable to other bridges or structural systems.

2. Lasso Regression

Although an in-depth discussion on the Lasso regression algorithm and its advantages can be found elsewhere [6, 7], a brief summary is presented herein. Consider a set of data $(x_{i1}, x_{i2}, ..., x_{ip}, y_i)$, $i = 1, ..., \mathrm{n}$, where the x_{ij}'s are the regressors and the y_i is the response variable of the i^{th} observation. Lasso regression minimizes the residual sum of squares subject to the sum of absolute values of the coefficient being less than a constant, i.e.,

$$\hat{\beta}^{L} = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^{N}\left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 \qquad \text{subject to} \sum_{j=1}^{p}\left|\hat{\beta}^{L}\right| \le t \tag{1}$$

The bound 't' is a tuning parameter and Lasso algorithm will yield the same estimates as ordinary least square regression when $\sum_{j=1}^{p}\left|\hat{\beta}^{L}\right| \ge t$. For, $0 < t < \sum_{j=1}^{p}\left|\hat{\beta}^{L}\right|$, then Eq. (1) can be reformulated as

$$\hat{\beta}^{L} = \underset{\beta}{\operatorname{argmin}}\left(\sum_{i=1}^{N}\left(y_i-\beta_0-\sum_{j=1}^{p}x_{ij}\beta_j\right)^2+\lambda\sum_{j}\left|\beta_j\right|\right) \tag{2}$$

Eq. (2) will force some regression coefficients to be zero due to the nature of constraints. It is noteworthy to mention that Lasso sacrifices a little bias to reduce the variance of the predicted valued and improve the overall prediction accuracy.

3. Proposed approach

Recently, previous studies [9, 10, 11] developed and used parameterized component and system fragility curves of highway bridges using multi-parameter demand models in conjunction with logistic regression techniques. The major modification of the proposed approach lies in using the Lasso regression instead of least square regressions (or surrogate models) for the generation of fragility curves. As mentioned before, Lasso regression identifies the non-significant input parameter and set their corresponding regression coefficients to zero. A brief outline of this method is given in line with the input parameters ($p_1,\ldots,p_n$, IM) and the output measures ($k_1,\ldots,k_m$) used in the current study.

Step 1: Evaluate the linear regression coefficients (β_{li}) by performing Lasso regression analysis for each component (k_i, $i = 1,\ldots,m$) with the input parameters ($p_1,\ldots,p_n$, S_a). The entire predictor variables are assumed to be statistically independent. This step helps to identify the variables which have a significant effect on the seismic demand.

Step 2: Generate a large number of demand estimates (N, 1 million in this study) for each component, k_i, using their respective regression model by generating N values of randomly generated input parameters based on their probabilistic distribution (Table 1, Latin hypercube sampling technique).

Step 3: Generate N capacity values for a specific damage state for each bridge component, k_i, based on the assumed distribution of the limit states (Table 2).

Step 4: Obtain the binary survive-failure ($N \times 1$) vector by comparing the capacity values (step 3) with the demand values (step 2). If a demand value is less than the associated limit state value (or capacity value), the associated element of the vector is zero (survival); otherwise, it is a unity (failure).

Step 5: Conduct a Lasso logistic regression on the survive-failure vector to determine the k^{th} component probability model, conditioned on the input parameters as

$$PF_{k|S_a, p_1, p_2, \ldots p_{n_l}} = \frac{e^{\theta_{k,0}+\theta_{k,S_a}\ln(S_a)+\sum_{j=1}^{n_l}\theta_{kj}\ln(p_j)}}{1+e^{\theta_{k,0}+\theta_{k,S_a}\ln(S_a)+\sum_{j=1}^{n_l}\theta_{kj}\ln(p_j)}}, \quad n_l \le n \tag{3}$$

where, $\theta_{k,0}$, $\theta_{k,sa}$, and $\theta_{k,j}$'s ($j = 1,\ldots,n_l$) are the Lasso logistic regression coefficient's of the k^{th} bridge component. This step helps to identify the sensitivity of bridge component fragility curves to the uncertain input parameters, which are not identified in step 1.

Step 6: Assuming that the bridge failure is a series system (the system fails if one or more components fail), estimate the binary survive-failure vector. The system level failure probability can be obtained by the Lasso logistic regression analysis for the system level binary-survive failure vectors. This step helps to identify the sensitivity of bridge system fragility to the uncertain input parameters

$$PF_{SYS|S_a,p_1,p_2,..p_{n_s}} = \frac{e^{\theta_{SYS,0}+\theta_{SYS,S_a}\ln(Sa)+\sum_{j=1}^{n_s}\theta_{SYS,j}\ln(p_j)}}{1+e^{\theta_{SYS,0}+\theta_{SYS,S_a}\ln(Sa)+\sum_{j=1}^{n_s}\theta_{SYS,j}\ln(p_j)}}, \quad n_s \le n \tag{4}$$

where, $\theta_{SYS,0}$, $\theta_{SYS,sa}$, and $\theta_{SYS,j}$'s (j = 1,…, n_s) are the Lasso logistic regression coefficient's for the system failure.

Step 7: For a particular bridge with input parameters, $p_1,...,p_{n_s}$, the classical one-dimensional fragility curves can be obtained as

$$PF_{SYS|S_a} = \int_{p_1}\int_{p_2}...\int_{p_{n_s}} \frac{e^{\theta_{SYS,0}+\theta_{SYS,S_a}\ln(S_a)+\sum_{j=1}^{n_s}\theta_{SYS,j}\ln(p_j)}}{1+e^{\theta_{SYS,0}+\theta_{SYS,S_a}\ln(S_a)+\sum_{j=1}^{n_s}\theta_{SYS,j}\ln(p_j)}} f(p_1)...f(p_{n_s})\,dp_1...dp_{n_s} \tag{5}$$

where $f(p_1)$,…, $f(p_9)$ are the probability density parameters for parameters, $p_1,...,p_{n_s}$.

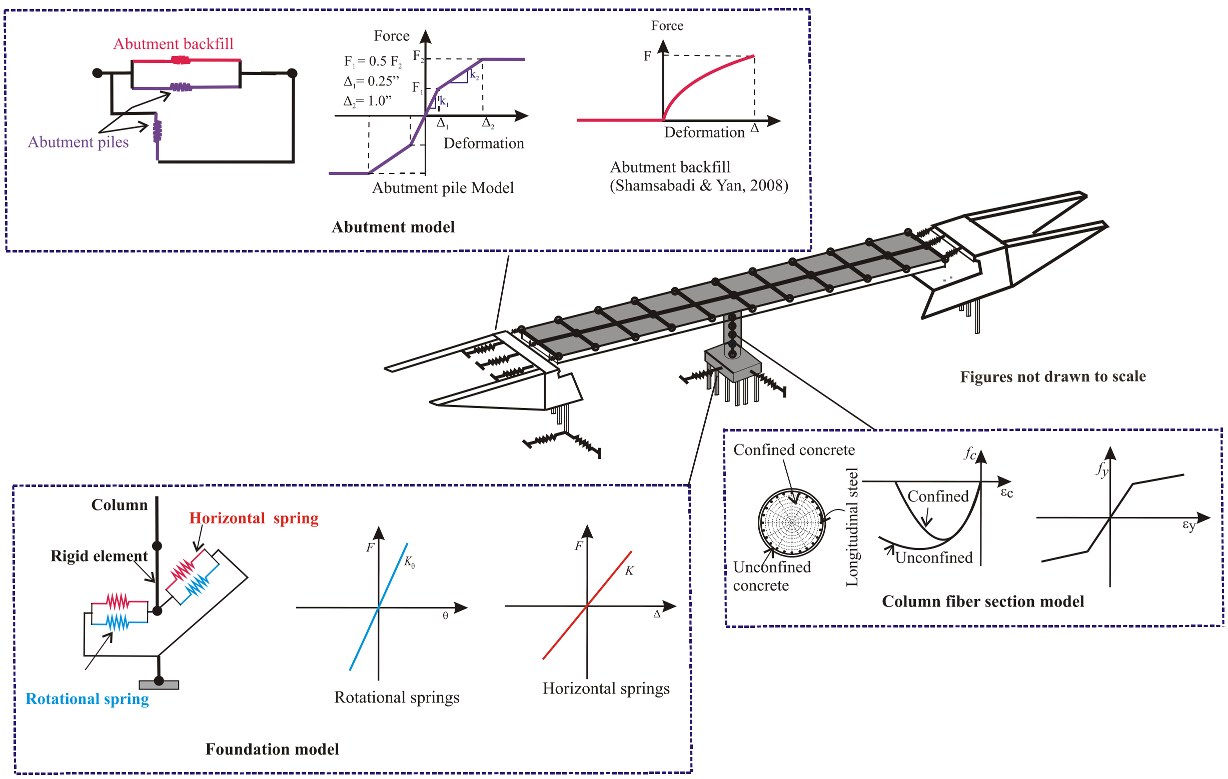

Fig. 1– Numerical model for the selected bridge

4. Sensitivity study with example bridge class

The application of the proposed approach is demonstrated in this study through the case study of two span box girder bridges with rigid abutments. The adopted bridge is one of the most common bridge types in California [13]. Although a detailed explanation of the analytical modeling of the bridge can be found elsewhere [2, 13], the general approach is presented herein. Three dimensional numerical modeling is carried out with the help of the finite element package OpenSees [14] incorporating both geometric and material nonlinearities. Longitudinal deck elements are modeled using elastic beam-column elements as they typically remain elastic during a seismic event. The columns and bent caps are modeled using fiber elements and foundations with rotational and translational springs. The contact element developed by Muthukumar and DesRoches [15] is used to model the pounding between the decks. The passive response of the abutment backwall is simulated using the hyperbolic soil model proposed by Shamsabadi and Yan [16]. Trilinear springs stemming from the recommendations of Mangalathu et al. [13] are used to model the piles. The typical configuration of the selected box-girder bridge and associated numerical model of various bridge components are shown in Fig. 1.

4.1 Uncertainty in bridge parameters

Mangalathu et al. [13] have identified the likely ranges in the modeling parameters and is used in the current study. The sources of uncertainty evaluated in this study can generally be classified as ground motion, gross geometry and modeling parameter uncertainty. The ground motions developed by Baker et al. [17] are used to assess the uncertainty in ground motions. It consists of 120 ground motions associated with moderate-to-strong earthquakes at small distances and 40 ground motions with strong velocity pulses characteristics of sites experiencing near-fault directivity effects. The spectral acceleration at 1.0 sec ($S_{a\text{-}1.0s}$) is considered as the IM in the current study. The geometric configuration can differ from bridge to bridge and the span length, deck width, column height and abutment wall height is considered as the uncertain geometric parameters. There are a large number of analytical modeling parameters which are potentially variable and are given in Table 1. The abutment soil backfill (ST), girder type (GT) and earthquake direction (ED) follows the Bernoulli distribution. For convenience, ST = e (where e is the Euler's number) if the backfill is sand. Otherwise, ST = e^2 (ln(e) = 1 and ln(e^2) = 2). In similar fashion, Bernoulli variables are assigned to GT (e for reinforced and e^2 for pre stressed) and ED (e for fault normal component applied to global X axis and e^2 otherwise).

Having identified the uncertain parameters for the regional risk assessment, statistically significant yet nominally identical 3-D bridge models are generated by sampling across the range of parameters using Latin Hypercube Sampling. One hundred and sixty analytical bridge models are generated consistent with the number of ground motions and are paired randomly. Non-linear time history analysis (NLTHA) is carried out on each bridge model and the peak component responses are noted to determine the relationship between the peak demands and the input parameters.

4.2 Engineering demand parameters and associated limit states

Five engineering demand parameters such as column curvature ductility (μ_ϕ), deck displacement (δ_d), abutment displacement in active (δ_a), passive (δ_p) and transverse direction (δ_p) are considered in the current study and the associated limit states are shown in Table 2. The capacity models are described by a two-parameter lognormal distribution with median, S_c and dispersion, β_c (β_c is assigned as 0.35 in a subjective manner due to lack of sufficient information and adopted as same across the components and the respective damage states).

Table 1 – Uncertainty (modeling) parameters and their probability distribution [13]

Parameter	Probability distribution		
	Type §	Parameters†	
		α	β_1
Material properties			
Concrete compressive strength, f_c (ksi)	N	3.90	0.48
Rebar yield strength, f_y (ksi)	LN	4.21	0.08
Superstructure			
Span length, L (ft)	LN	4.61	0.27
Deck width D_{w} (ft)	LN	3.45	0.20
Girder type, *GT* (Reinforced vs. Prestressed)	B	–	–
Interior bent (single-column)			
Column clear height, H_c (mm)	LN	10.33	0.13
Column longitudinal reinforcement ratio, ρ_c	U	0.01	0.04
Column diameter D (60 vs 72)	B	–	–
Translational stiffness of a pile group, K_{ft} ($\times 10^3$ kip/in.)	N	1.70	0.80
Rotational stiffness of a pile group, K_{fr} ($\times 10^7$ kip-in./rad)	N	4.10	1.20
Exterior bent (diaphragm abutment on piles)			
Abutment height, H_a (ft)	LN	2.35	0.15
Backfill type, *ST* (sand vs. clay)	B	–	–
Pile stiffness, k_p (kip/ft)	LN	1.79	0.35
Other parameters			
Mass factor, m_f	U	1.1	1.4
Damping, ξ	N	0.045	0.01
Earthquake direction (fault normal FN vs. parallel FP), *ED*	B	–	–

§ N = normal, LN = lognormal, U = uniform, and B = Bernoulli distribution.

† α and β_1 represent parameters of the respective distribution. These denote mean and standard deviation for a normal distribution, lower and upper bound in the case of uniform distribution and mean and standard deviation of the associated normal distribution in the case of a lognormal distribution

Table 2 – Limit state models for EDPs of bridge components [18]

Component		LS_1 (slight)		LS_2 (moderate)		LS_3 (extensive)		LS_4 (complete)	
		S_C	β_C	S_C	β_C	S_C	β_C	S_C	β_C
Colum curvature ductility, μ_ϕ		0.8	0.35	0.9	0.35	1.0	0.35	1.2	0.35
Deck displacement, δ_d (in.)		4.0	0.35	12.0	0.35	–	–	–	–
Abutment displacement (in.)	Passive action, δ_p	3.0	0.35	10.0	0.35	–	–	–	–
	Active action, δ_a	1.5	0.35	4.0	0.35	–	–	–	–
	Tangential action, δ_t	1.0	0.35	4.0	0.35	–	–	–	–

4.3 Identification of significant predictor variable

The predictor variables that have a significant effect on the seismic demand and the fragility curves are identified through the above mentioned approach (section 3). The advantage of the proposed approach is that it helps to identify the significant variables during the generation of fragility curves, i.e, only one iteration is needed for the generation of the fragility curves and the identification of the significant predictor variables. The proposed approach also helps to identify the significant parameters depending upon the limit state under consideration. It is seen from Table 3 that out of the seventeen input parameters with uncertainty, the parameters which has a significant effect on the system or component fragilities are ground motion intensity measure (*IM*), span length (*L*), soil type (*ST*), girder type (*GT*), column diameter (*D*), reinforcement ratio (ρ_c), pile stiffness (k_p), mass factor (m_f) and foundation translation stiffness (K_{ft}). *IM*, and *L* seem to have a significant effect on all the EDPs in all the considered limit states. Column vulnerability is significantly affected by the parameters *D*, *ST*, ρ_c and f_c. It is also noted that the seismic demand and the seismic fragilities are less affected by the *ED*, f_y, H_a, D_w, H_c and ξ.

Table 3 – Significant parameters for demand models

EDPs	Significant input parameters			
	Slight	Moderate	Extensive	Collapse
μ_φ	*IM, ST, L, D,* ρ_c, f_c	*IM, ST, L, D,* ρ_c, f_c	*IM, ST, L, D,* ρ_c, f_c	*IM, ST, L, D,* ρ_c, f_c
δ_d	*IM, ST, L,* k_p, K_{ft}	*IM, ST, GT, L,* k_p, K_{ft}, m_f	–	–
δ_p	*IM, ST, L, GT*	*IM, ST, L, GT*	–	–
δ_a	*IM,, L,* k_p, K_{ft}	*IM,, L*	–	–
δ_t	*IM,, L,* k_p	*IM,,ST, L,* k_p, K_{ft}	–	–

4.4. Multi-dimensional fragility curves

The multi-dimensional fragility curves for various components are generated through the approach outlined in section 3 (steps 1 - 5) based on the survive-failure vector using the identified significant parameters. The multi-dimensional system fragility is derived based on a series system assumption (step 6, section 3). The one-dimensional fragility curves (connection IM and probability of damage exceeding a specified limit state) can be generated through the integration of multi-dimensional fragility through the domain of uncertain input parameters (Eq. (5)). The integrated fragility curves can be used for the regional risk assessment of bridge inventory. Fig. 2 shows the comparison of the system fragility curves considering all the uncertain input parameters and the significant input parameters. It is seen that the fragility curves generated by considering only the significant parameters are fairly in good agreement with the fragility curves generated by considering all the uncertain input parameters.

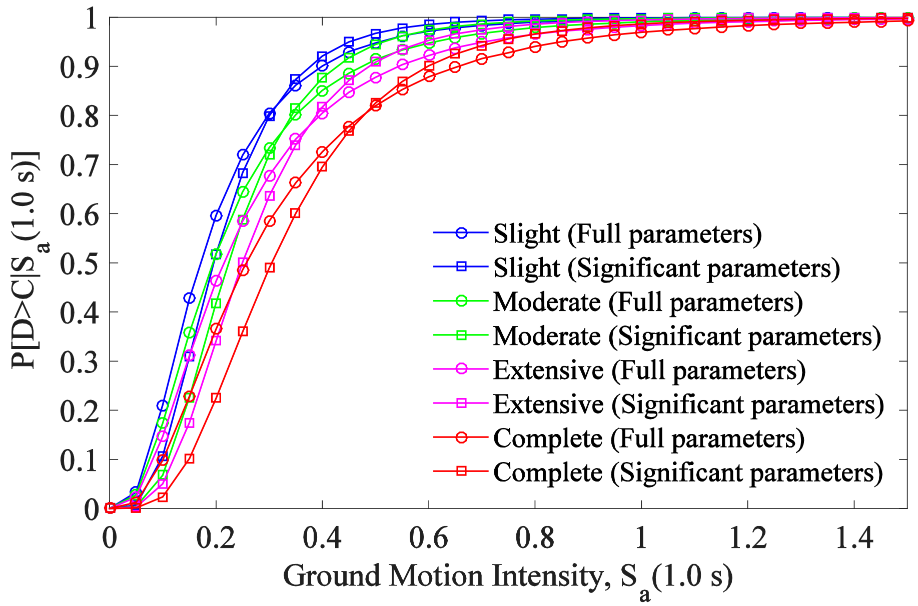

Fig. 2 – Comparison of system fragility curves for various damage states

5. Conclusions

The current study suggests an approach to identify the uncertain input parameters which have a significant influence on the seismic demand as well as the fragility curves. Such an approach helps to gain insight in quantifying whether the uncertain parameter variations should be treated explicitly or to be neglected. The proposed approach identifies the significant parameters during the generation of the fragility curves. The approach also helps to identify the significant parameters depending upon the limit state under consideration.

The proposed approach is demonstrated in this paper through the case study of a two span box girder bridge with diaphragm abutments. Three dimensional finite element models are developed in OpenSees platform accounting for the material, geometric and modeling parameter uncertainties, and are paired randomly with the ground motions. Non-linear time history analysis is carried out for each model and the peak responses are noted. The significant input parameters are identified through Lasso regression in the generation of multi-dimensional fragility curves. The parameters which have a significant effect on the system or component fragilities are ground motion intensity measure, span length, soil type, girder type, column diameter, reinforcement ratio, pile stiffness, mass factor and foundation translation stiffness. Ground motion direction, steel yield strength, height of column, deck width, abutment height, and damping ratio have less influence on the seismic demand and the fragilities. The seismic demands, as well as the fragilities, were found to be particularly sensitive to the uncertainty in the geometric parameters. Although shown here with a specific bridge, the proposed approach is relevant and applicable to other structural systems

6. References

[1] Padgett JE, DesRoches R (2007): Sensitivity of Seismic Response and Fragility to Parameter Uncertainty. *Journal of Structural Engineering*, **133**, 1710-1718.

[2] Mangalathu S, Jeon JS, DesRoches R, Padgett JE (2015): Analysis of Covariance to Capture the Importance of Bridge Attributes on the Probabilistic Seismic Demand Model. *10th Pacific Conference on Earthquake Engineering. Sydney*, Australia.

[3] Ramanathan K, Padgett JE, DesRoches R (2015): Temporal evolution of seismic fragility curves for concrete box-girder bridges in California. *Engineering Structures*, **97**, 29-46.

[4] Jangid R (2004): Seismic response of isolated bridges. *Journal of Bridge Engineering*, **9**, 156-166.

[5] Nielson BG, DesRoches R (2006): Influence of modeling assumptions on the seismic response of multi-span simply supported steel girder bridges in moderate seismic zones. *Engineering Structures*, **28**, 1083-1092.

[6] Tibshirani R (1996): Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267-288.

[7] Friedman J, Hastie T, Tibshirani R (2001): The elements of statistical learning. Springer series in statistics, *Springer*, Berlin.

[8] Zhang CH, Huang J (2008): The sparsity and bias of the Lasso selection in high-dimensional linear regression. *The Annals of Statistics*, 1567-1594.

[9] Dukes, J D (2013): Application of bridge specific fragility analysis in the seismic design process of bridges in California. *Ph.D Thesis*, Civil and Environmental Engineering. Georgia Tech

[10] Ghosh J, Padgett JE, Dueñas-Osorio L (2013): Surrogate modeling and failure surface visualization for efficient seismic vulnerability assessment of highway bridges. *Probabilistic Engineering Mechanics*, **34**, 189-199.

[11] Mangalathu S, Jeon JS, Soleimani F, DesRoches R, Padgett JE, Jiang J (2015): Seismic vulnerability of multi-span bridges: An analytical perspective. *10th Pacific Conference on Earthquake Engineering*, Sydney, Australia.

[12] Jeon JS, Mangalathu S, Song J, DesRoches R (2015): Parameterized seismic fragility curves for curved multi-frame concrete box-girder bridges using Bayesian parameter estimation. *Journal of Earthquake Engineering*, In Review.

[13] Mangalathu S, Jeon JS, DesRoches R, Padgett JE (2016): ANCOVA-based grouping of bridge classes for seismic fragility assessment. *Engineering Structures*, In Review.

[14] Mazzoni, S., F. McKenna, M. H. Scott & G. L. Fenves (2006): OpenSees command language manual. *Pacific Earthquake Engineering Research (PEER) Center*.

[15] Muthukumar S, DesRoches R (2006): A Hertz contact model with non-linear damping for pounding simulation. Earthquake *Engineering & Structural Dynamics*, **35**, 811-828.

[16] Shamsabadi A, Yan L (2008): Closed-form force-displacement backbone curves for bridge abutment-backfill systems. *Proceedings of the Geotechnical Earthquake Engineering and Soil Dynamics IV Congress, American Society of Civil Engineers*.

[17] Baker JW, Lin, T, Shahi. SK, Jayaram N (2011): New ground motion selection procedures and selected motions for the *PEER transportation research program. Pacific Earthquake Engineering Research Center*.

[18] Ramanathan KN (2012): Next generation seismic fragility curves for California bridges incorporating the evolution in seismic design philosophy. *Ph.D Thesis*, School of Civil & Environmental Engineering. Georgia Institute of Technology. Atlanta, Georgia.