


The promise of implementing machine learning in earthquake engineering: A state-of-the-art review¹

Earthquake Spectra
2020, Vol. 36(4) 1769–1801
© The Author(s) 2020

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/8755293020919419
journals.sagepub.com/home/eqs



Yazhou Xie, M.EERI¹, Majid Ebad Sichani² ,²
Jamie E Padgett, M.EERI², and Reginald
DesRoches, M.EERI²

Abstract³

Machine learning (ML) has evolved rapidly over recent years with the promise to substantially alter and enhance the role of data science in a variety of disciplines. Compared with traditional approaches, ML offers advantages to handle complex problems, provide computational efficiency, propagate and treat uncertainties, and facilitate decision making. Also, the maturing of ML has led to significant advances in not only the main-stream artificial intelligence (AI) research but also other science and engineering fields, such as material science, bioengineering, construction management, and transportation engineering. This study conducts a comprehensive review of the progress and challenges of implementing ML in the earthquake engineering domain. A hierarchical attribute matrix is adopted to categorize the existing literature based on four traits identified in the field, such as ML method, topic area, data resource, and scale of analysis. The state-of-the-art review indicates to what extent ML has been applied in four topic areas of earthquake engineering, including seismic hazard analysis, system identification and damage detection, seismic fragility assessment, and structural control for earthquake mitigation. Moreover, research challenges and the associated future research needs are discussed, which include embracing the next generation of data sharing and sensor technologies, implementing more advanced ML techniques, and developing physics-guided ML models.

Keywords⁵

Machine learning, earthquake engineering, seismic hazard analysis, system identification and damage detection, structural control, seismic fragility assessment⁶

Date received: 19 September 2019; accepted: 22 February 2020⁷

¹Department of Civil Engineering and Applied Mechanics, McGill University, Montreal, QC, Canada

²Department of Civil and Environmental Engineering, Rice University, Houston, TX, USA

Corresponding author:

Yazhou Xie, Department of Civil Engineering and Applied Mechanics, McGill University, Montreal, QC H3A 0C3, Canada.

Email: tim.xie@mcgill.ca

Introduction ¹

Machine learning (ML) has played a pivotal role in many areas of science, finance, and engineering. By definition (Samuel, 1959), ML is a field of study that gives computers the ability to learn without being explicitly programmed. ML algorithms can be divided into two main types: supervised learning and unsupervised learning. Supervised learning uses prior knowledge of the labeled data set to learn a function that best approximates the relationship between input and labeled output in the data. In contrast, unsupervised learning aims to infer the natural structure present with a set of data points that have no target labels. Depending on the data characteristics (i.e. discrete or continuous) and task goals, supervised learning can be further subdivided into classification and regression, while unsupervised learning comprises clustering and dimensionality reduction. Figure 1 summarizes the two types of ML and some commonly used ML algorithms (Kong et al., 2018).

Previous studies have surveyed the applications of ML and other advanced soft computing tools in civil engineering. Adeli (2001) reviewed the works on the integration of artificial neural networks (ANNs) with different computing paradigms, such as genetic algorithm (GA), fuzzy logic, and wavelet analysis, to enhance the application of ANNs in civil/structural engineering. An extensive literature study of evolutionary computation in the context of structural design, such as topological optimum design, has been conducted by Kicinger et al. (2005). In a separate review, Liao et al. (2011) demonstrated the soundness of employing metaheuristics in addressing optimization problems in project and construction management, where metaheuristics involve high-level procedures designed to find, generate, or select heuristics for optimization. Thereafter, a similar study by Saka and Geem (2013) has reviewed different mathematical modeling schemes on the optimum design of steel frames. A general review has been made by Lu et al. (2012) to examine various artificial intelligence (AI) algorithms in solving civil engineering issues. The stochastic nature of soil behavior in geotechnical engineering has also attracted the significant use of AI techniques, where review studies have been conducted on both shallow foundations

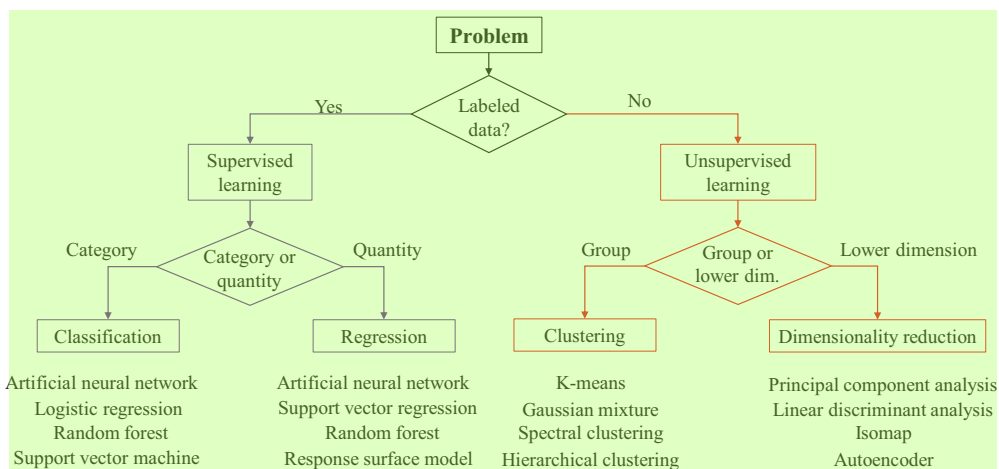


Figure 1. ML types and some commonly used ML algorithms (Kong et al., 2018). ⁵
ML: machine learning.

(Shahin, 2014) and pile foundations (Shahin, 2016). Following Adeli's (2001) study, a review of the recent applications of ANNs on civil infrastructure has been performed by Amezcuita-Sanchez et al. (2016), where several topics have been covered, including structural system identification, structural health monitoring, structural vibration control, structural design and optimization, prediction applications, construction engineering, and geotechnical engineering. Recently, Salehi and Burgueño (2018) summarized the involvement of ML, pattern recognition (PR), and deep learning (DL) in structural engineering, whereas an overview of the ML applications in seismology was provided by Kong et al. (2018) who discussed the benefits of practicing ML algorithms in earthquake detection and phase picking, earthquake early warning, ground motion prediction, seismic tomography, and earthquake geodesy.

Although some of the review works have touched on a few of the ML applications in earthquake engineering, a comprehensive review is lacking in this area. As such, it is unclear to what extent ML has permeated the earthquake engineering domain, enabling and advancing research or supporting decision makers to reduce the effects of seismic hazards on civil structures. Earthquake engineering is an interdisciplinary engineering branch that describes earthquake hazard at the source, characterizes site effects and structural response, evaluates seismic risk and vulnerability, as well as assesses seismic protection measures. Within each subfield, ML algorithms indeed have been implemented in numerous circumstances. Prominent academic databases, such as Web of Science, Engineering Village, and Wiley Online Library, were used to search and select the publications that have titles or keywords consisting of ML algorithms in an earthquake engineering context. The search results indicate that nearly 200 relevant publications are now available, with a clear exponential growth in publications that intersect these two fields. A survey of this work is of interest to systematically present theoretical background of the commonly used ML methods, historical context on their use, state-of-the-art of research developments, potential challenges, as well as future promises.

The body of knowledge of the reviewed publications can be categorized according to four traits in the field, such as ML method, topic area, data resource, and scale of analysis. The available research is designated to different sets of attributes grouped by these four traits. Table 1 provides a portion of the attribute matrix for some example studies, while the complete list of studies and the description of the column headings (e.g. 1a) can be found in the Appendix (provided as a supplementary file). As illustrated in Table 1, a hierarchical structure is formed to explore the literature related to each attribute and its relation with other attributes. Such a structure makes clear the linkages across all studies that can facilitate implementation of research outcomes; improve data generation, sharing, and collection; provide reference work for model validation and comparison; and assist in identifying long-term future work.

This article focuses on reviewing the literature using the first two traits, namely, ML method and topic area. As shown in Figure 2, seven classes of ML methods have found the most use to date in four topic areas in earthquake engineering. The seven classes of ML methods are ANN, support vector machine (SVM), response surface model (RSM), logistic regression (LR), decision tree (DT) and random forest (RF), hybrid methods that couple two or more soft computing algorithms, and all other methods (e.g. evolutionary computing (EC) and genetic expression programming (GEP)) that are not significant in number of applications. The four topic areas include (1) seismic hazard analysis, (2) system identification and damage detection, (3) seismic fragility assessment, and (4) structural control for earthquake mitigation. These four topics are defined to be extensive enough to

Table 1. Matrix of attributes for sample studies reviewed in this article 1

Reference	ML method							Topic area				Data resource			Scale of analysis		
	1a	1b	1c	1d	1e	1f	1g	2a	2b	2c	2d	3a	3b	3c	4a	4b	4c
Liu and Tesfamariam (2012)																	
Javadi et al. (2006)																	
Kaya (2016)																	
Mitra et al. (2011)																	
Jeon et al. (2014)																	
Yaseen et al. (2018)																	
Mangalathu and Jeon (2018)																	

ML: machine learning. 3

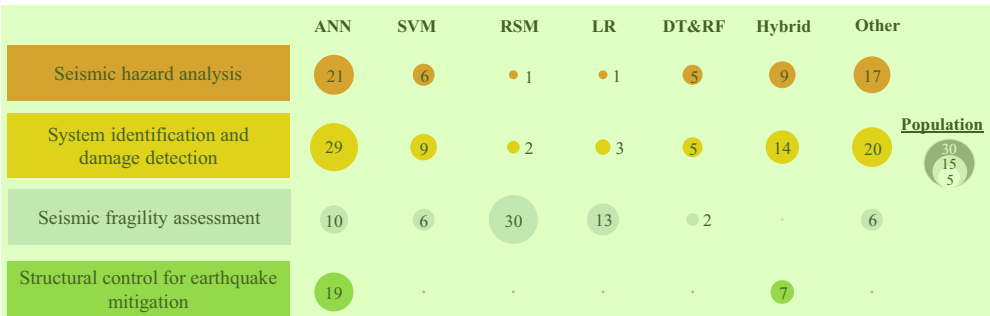


Figure 2. Applications of seven classes of ML methods in four topic areas in earthquake engineering. 5
ML: machine learning; ANN: artificial neural network; SVM: support vector machine; RSM: response surface model; LR: logistic regression; DT: decision tree; RF: random forest.

capture the full range of ML applications in earthquake engineering, but simple enough such that each topic represents a generally distinct area of research. First, seismic hazard analysis consists of the studies that predict the level of ground shaking and its associated uncertainty at a given site or location. Also, the use of ML tools to evaluate soil liquefaction potential and predict the liquefaction-induced lateral spread displacement is considered within this area. The second topic is a twofold area where system identification comprises a collection of studies that utilize ML to emulate a structural system and predict its deterministic seismic response, and damage detection is broadly defined as the use of ML models to recognize, classify, and assess seismic damage to civil structures. Third, seismic fragility assessment that incorporate various sources of uncertainties has shown to be a promising and popular field to practice ML techniques, where ML methods have been employed to develop probabilistic seismic demand models (PSDMs) and parameterized fragility functions. The fourth topic area lies in ML-equipped active and semi-active control of structures that mitigate the adverse effects of earthquake hazards. As is depicted in Figure 2, ANN gains the highest total number of applications, reflecting its robustness of tackling various types of problems in earthquake engineering. Conversely, the most popular method in seismic fragility assessment turns out to be RSM, which has been more

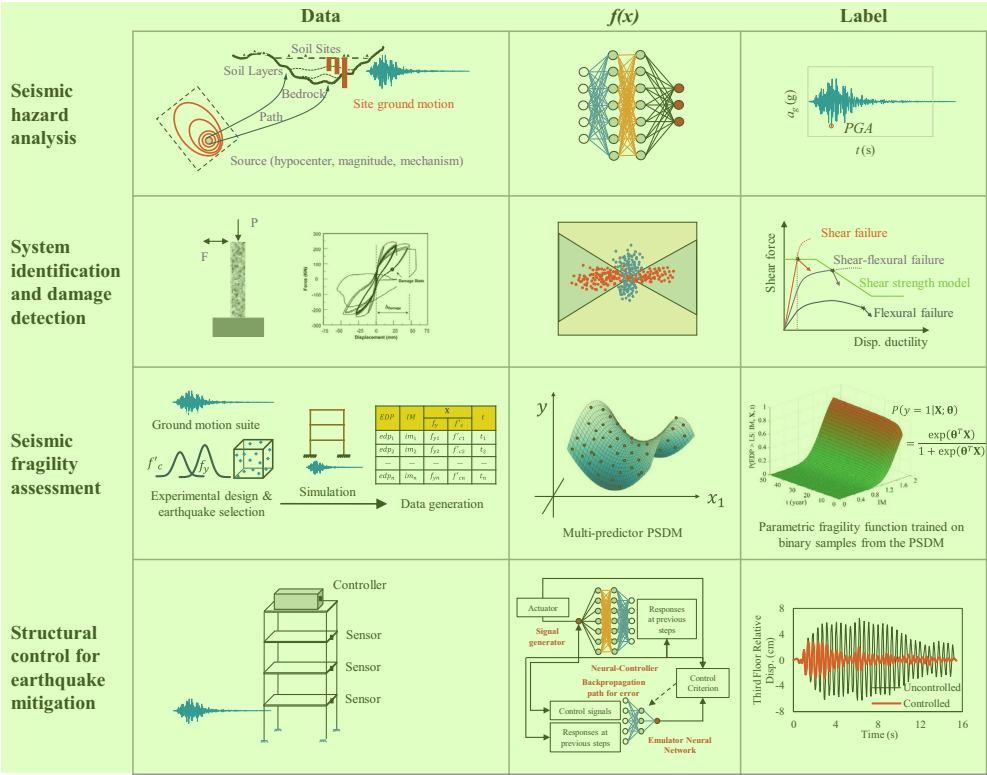


Figure 3. Example ML applications in earthquake engineering: The top row shows training an ANN to predict the PGA of ground motions using strong motion databases; the second row shows predicting the column failure mode by applying a quadratic discriminant analysis approach to laboratory data; the third row shows the development of multi-predictor PSDM and multidimensional fragility model through RSM and LR; the fourth row shows the results from coupling ANNs with active control to mitigate seismic response of the building. ML: machine learning; ANN: artificial neural network; PGA: peak ground acceleration; PSDM: probabilistic seismic demand model; RSM: response surface model; LR: logistic regression.

frequently implemented than other ML methods. Examples of ML applications in these four topic areas are provided in Figure 3.

This article serves as the first comprehensive survey on the applications of ML in earthquake engineering, attempting to be comprehensive of the primary themes in the literature to date. The common classes of ML methods and their four areas of applications are investigated in detail in order to understand the current state of the field, elucidate connections across existing literature, and pave the path to promote broader and more fundamental ML advances in solving related research issues in earthquake engineering. In the next section, this article presents the technical background for commonly used ML methods to provide context and foundation for discussing their application in earthquake engineering. Subsequently, the history and state-of-the-art applications of various ML algorithms in each of four topic areas within earthquake engineering are reviewed, shedding light on what subfields of earthquake engineering have benefited (or are benefiting) from the promise of ML methods. Moreover, a discussion of existing challenges and future opportunities

concludes the article in anticipation of the future growth of ML applications in the earthquake engineering domain. ¹

ML methods used in earthquake engineering ²

The theoretical background of ML methods is discussed in this section. As shown in Figure 2, representative ML methods are considered as ANN, SVM, RSM, LR, DT&RF, and Hybrid methods that couple ANN with other soft computing algorithms, such as fuzzy logic and wavelet analysis. It is worth mentioning that there exist other ML methods that have a limited number of applications. These ML methods, which are listed in the “Other” category in Figure 2, are not discussed in this section because of the limited space. ³

Artificial neural network ⁴

First developed by McCulloch and coworkers (Perlovsky, 2000), a “shallow” ANN typically consists of three layers: the input layer, hidden layer, and output layer. In particular, model variables in the input layer are weighted and fed into the hidden layer that consists of a series of nonlinear relationships such as sigmoidal functions, which are further weighted and fed into the output layer to provide a regression or classification model. Connection weights are learned in the forward propagation and updated through a training process that minimizes the prediction error, which is typically propagated in the backward direction. The classic ANN can be extended to be a deep neural network (DNN), or DL, that incorporates multiple hidden layers (LeCun et al., 2015). The deep architecture in a DNN consists of multiple processing layers and nonlinear transformations, enabling superior model performance without requiring well-selected features as inputs. The structure of ANNs is especially suited for modeling the behavior of ill-posed problems with nonlinear and intricate patterns. However, finding the optimal structure and tuning the model parameters could be challenging, potentially producing over-fitted models if not trained carefully. Additional information on ANN is available in the literature, such as Bishop (2006) and Murphy (2012). ⁵

Support vector machine ⁶

SVM is a binary classification algorithm that uses kernel functions to enable an implicit mapping of the data into a high-dimensional feature space. For separable data shown in Figure 4(a), an optimum margin classifier is carried out to construct a separating hyperplane that maximizes the margin between the hyperplane and the support vectors, which comprise the data points that lie closest to the hyperplane. Note that the incorporation of slack variables ξ_i and ξ_j in Figure 4(a) allows some sample cases to be misclassified. ⁷

SVM can also be used as a regression method, named as support vector regression (SVR) in this study, by maintaining the main features that characterize the algorithm (e.g. the maximal margin concept) and applying a few minor changes. In analogy with SVM, SVR is realized through a loss function that defines an ε -insensitive region around the true responses of the training points (Vapnik, 1998). As shown in Figure 4(b), the errors are considered zero for all points that are inside the ε -insensitive region. Through nonlinear kernel mapping, SVR performs linear regression in the high-dimensional feature space, while slack variables are incorporated to measure the deviation of training samples outside the ε -insensitive zone (Figure 4(b)) (Bishop, 2006; Murphy, 2012). ⁸

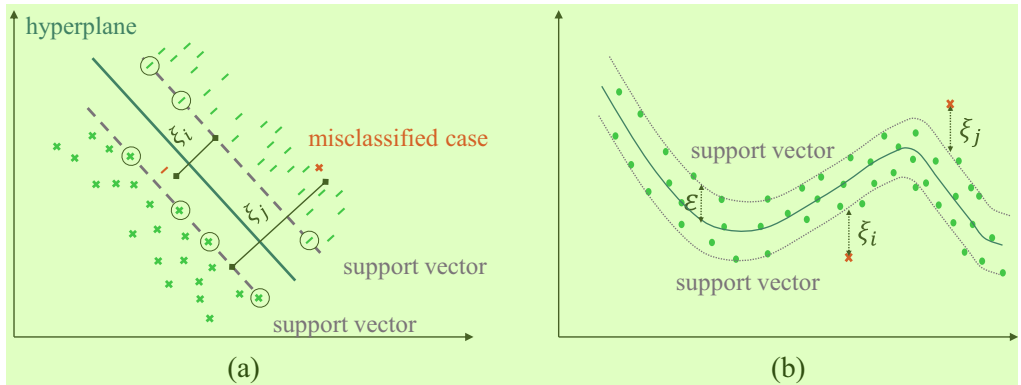


Figure 4. SVM for (a) classification and (b) regression. 2

SVM: support vector machine.

Response surface model³

Originally developed as a statistical method that explores the relationship of explanatory variables of a system and its responses (Box and Hunter, 1957), RSM has been commonly applied in various areas because of its simplicity, transparency, and transferability features. RSM estimates the responses using a series of basis functions. While high-order polynomials can be used as the basis functions for an RSM, first- and second-order polynomials are preferable in most of the cases. Note that the prediction error of an RSM is assumed to hold a normal distribution with mean zero and a variance value.

Two types of statistical methods can be used to improve an RSM. First, feature selection techniques can be applied to identify the most influential variables because not all of them may act as good predictors. For instance, stepwise algorithms determine the best fitting subset of the predictor variables by sequentially adding and removing terms from the proposed RSM based on a given statistical criterion, such as F-statistics or goodness of fit (Rawlings et al., 1998). Second, regularization techniques can be used to add a penalty term in the loss function to control the model size. Relevant methods include the least absolute shrinkage and selection operator (LASSO) algorithm (Tibshirani, 1996) and the elastic net (Zou and Hastie, 2005).

Logistic regression⁶

Although many more complex extensions exist, LR in its basic form measures the relationship between the categorical dependent variable and one or more independent variables through a logistic distribution function. Namely, the sigmoid function is used to estimate the probability that a new data point belongs to one of the two classes. Additional information about LR is available in the literature (Hosmer and Lemeshow, 2000).

Decision tree and random forest⁸

DT, also named as classification and regression tree (CART), is an ML algorithm that recursively partitions the input space and defines a local model in each resulting region of the input space. As shown in Figure 5(a), a simple regression model can be fit to each subspace in case of regression, while a class can be assigned to each subspace in case of

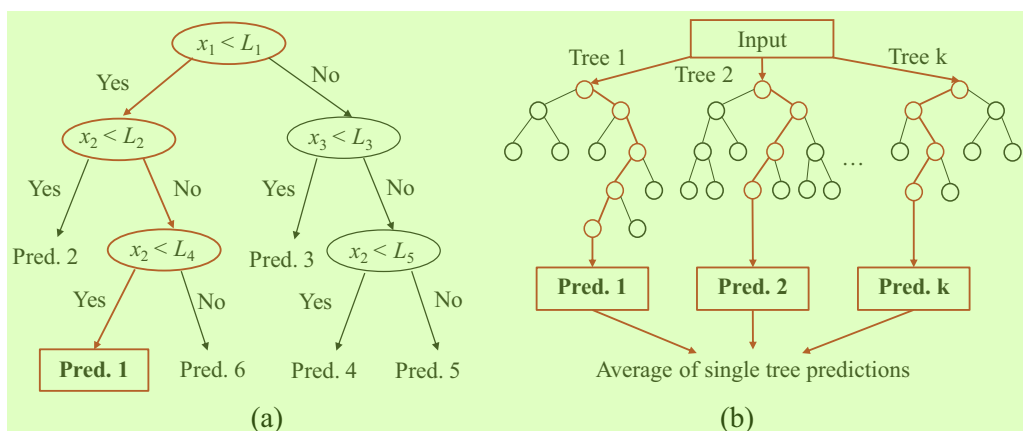


Figure 5. Illustration of (a) binary DT and (b) RF. 2
DT: decision tree; RF: random forest.

classification. A cost function is applied together with a greedy construction procedure to find the optimal partitioning of the data. To overcome the potential overfitting and instability issues associated with a single tree, RF constructs a multitude of DTs at training and outputs the mean predictions of individual trees (Ho, 1995) (Figure 5(b)). The training algorithm generates RFs by bootstrap aggregating or bagging. In order to keep uncorrelated predictors in the final model, RF uses a randomly selected subset of the problem variables in the DTs (Murphy, 2012).

Hybrid methods 4

ML methods, particularly ANNs, can be combined with other soft computing techniques to deal with the complexity and ambiguity in earthquake engineering problems. Such a combination provides a promising integrated system to overcome the inherent limitations of individual techniques. ANNs can be coupled with fuzzy logic algorithms to take advantage of the computational capabilities from both algorithms in a complimentary way. For instance, imprecise data information, such as linguistic statements, can be transformed into numerical data through a fuzzy logic interface for an ANN to train and make decisions (Alvanitopoulos et al., 2010). GAs can interact with an ANN in various steps to improve its performance (Adeli, 2001). Moreover, wavelet transform can analyze the different frequency components of an earthquake signal with varying levels of details, which has emerged as an effective technique that can be embedded in an ANN for system identification and damage detection (Adeli and Jiang, 2006; Hung et al., 2003). Technical background of these soft computing techniques and their combinations with ANNs can be found in relevant studies (Fuller, 1995; Wallen, 2004; Whitley et al., 1990).

Seismic hazard analysis 6

Ground motion prediction and generation 7

Ground motion prediction and generation form an essential part in seismic hazard analysis of civil structures. Conventional empirical methods relied on regression analyses to derive attenuation equations for different intensity measures of ground motions as

functions of the source, path, and site parameters (Boore and Atkinson, 2008; Boore et al., 2014; Douglas, 2003). A typical ground motion prediction equation (GMPE) for predicting peak ground acceleration (PGA) has the following form:

$$\log(\text{PGA}) = F_M + F_D + F_S + \varepsilon, \quad (1)$$

where F_M , F_D , and F_S represent functions to quantify the influences from magnitude, distance, and site effects, respectively, and ε is the error term. Selection of a proper functional form for F_M , F_D , and F_S is not straightforward as it requires not only appropriate identification and inclusion of significant independent variables but also a reasonable parametric quantification of the relationship between these input variables and the output. In this regard, the recent implementation of ML methods can eliminate the constraints from this predefined mathematical structure, where either parametric or nonparametric models can be derived.

ML studies in predicting GMPEs benefited from the availability of the strong motion database in Taiwan, Turkey (Gülkan and Kalkan, 2002), Iran (Amiri et al., 2010), Europe (Akkar et al., 2014), western United States (i.e. the NGA and NGA-West2 database (Ancheta et al., 2014; Chiou et al., 2008)), and central America for induced earthquakes (Khosravikia et al., 2018). A group of three to five independent parameters was typically identified as significant predictors to predict time-domain intensity measures, such as PGA, peak ground velocity (PGV), and peak ground displacement (PGD), and frequency-domain measures such as pseudo-spectral acceleration (PSA). The identified significant predictors are moment magnitude of the earthquake, source-to-site distance, the average shear-wave velocity of the site, faulting mechanism, and focal depth. The ML tools utilized in GMPEs include the ANN (the top row in Figure 3) (Bakhshi et al., 2014; Derras et al., 2014; Dhanya and Raghukanth, 2018; Güllü and Ergelebi, 2007; Kerh and Ting, 2005; Khosravikia et al., 2019), genetic programming (GP) (Cabalar and Cevik, 2009), multi-expression programming (MEP) (Alavi et al., 2011), SVR (Tezcan and Cheng, 2012; Thomas et al., 2017), GEP (Güllü, 2012; Javan-Emrooz et al., 2018), Lagrange equation discovery (ED) system (Markič and Stankovski, 2013), conic multivariate adaptive regression splines (CMARS) (Yerlikaya-Ozkurt et al., 2014), randomized adaptive neuro-fuzzy inference system (RANFIS) (Thomas et al., 2016), M5' model tree and CART (Hamze-Ziabari and Bakhshpoori, 2018; Kaveh et al., 2016), DNN (Derakhshani and Foruzan, 2019), and hybrid methods such as the coupling of GP and orthogonal least squares (OLS) (Gandomi et al., 2011), the combination of ANN and simulated annealing (SA) (Alavi and Gandomi, 2011), the coupling of GP and SA (Mohammadnejad et al., 2012), and the coupling of GA, ANN, and regression analysis (RA) (Akhani et al., 2019).

Significant improvements have been achieved in various aspects to enhance the accuracy and the generalization capability of the ML models in GMPEs. For instance, earlier studies have developed almost “perfect” ANN models for the whole data set (with close-to-one R^2 values) but failed to carry out necessary procedures to avoid overfitting. Some studies lack a well-established testing step, while others have shown a substantial accuracy mismatch between the training and testing data sets. These issues have been tackled when a comprehensive ground motion database became available, namely, the NGA project strong motion database (Chiou et al., 2008), which contains 3551 recorded motion data from 173 shallow crustal earthquakes, ranging in magnitude from 4.2 to 7.9. The following enhancements have been made using the NGA motion database: (1) a well-constructed learning, validation, and testing structure to avoid overfitting (Alavi et al., 2011; Alavi and Gandomi, 2011); (2) a separate

testing procedure on a different motion database to verify the generalization capability of the model (Gandomi et al., 2011; Thomas et al., 2016, 2017); (3) a hybrid framework that couples two or three ML tools to significantly improve the model performance (Akhani et al., 2019; Alavi and Gandomi, 2011; Gandomi et al., 2011; Hamze-Ziabari and Bakhshpoori, 2018; Mohammadnejad et al., 2012; Thomas et al., 2016). Moreover, the newly compiled NGA-West2 strong motion database (Ancheta et al., 2014), consisting of 21,336 recordings from 599 shallow crustal earthquakes, has provided a promising resource for developing more complex soft computing models. To this end, Derakhshani and Foruzan (2019) employed a DNN approach to construct more reliable predictive models for estimating the site PGA, PGV, and PGD based on earthquake magnitude, rake angle, source to site distance, and soil shear-wave velocity. Their DNN models, with correlation coefficients of around 0.9, have shown to be able to outperform other existing models. 1

However, there exist additional factors that are known, in a physical sense, to affect ground motion estimations. Factors such as stress drop and directivity effects have been neglected in most of the studies mentioned above, resulting in large variance or bias embedded in the derived models. This issue has been partially recognized and addressed by Trugman and Shearer (2018) who applied an RF algorithm to correlate earthquake stress drop with PGA for earthquakes in the San Francisco Bay Area. Their study has verified that ML techniques can produce more precise ground motion estimates if new independent features are included in the predictors. 2

Other than GMPEs, relatively few studies have focused on using ML algorithms for probabilistic seismic hazard analysis. Alimoradi and Beck (2015) used principal component analysis (Jolliffe, 2002) to extract a set of orthonormal basis vectors that capture the predominant variations in the time histories of earthquake waves. A Gaussian process procedure was subsequently used in their study for regression to generate a probabilistic target hazard spectrum under a given scenario event. The optimal linear combination of the orthonormal basis vectors to match the target spectrum was then identified through a genetic optimization process. Their study avoided the typical questions associated with the selection and scaling of recorded ground motions in the seismic analysis and design of civil structures. 3

Soil liquefaction potential and liquefaction-induced lateral spread 4

Historical earthquake events have witnessed significant structural damage caused by soil liquefaction (Bird et al., 2006; Cubrinovski et al., 2014). The application of ML tools in soil liquefaction research is twofold: (1) the use of PR tools to classify the occurrence versus non-occurrence of soil liquefaction (i.e. the assessment of liquefaction potential), and (2) the use of regression procedures to predict the liquefaction-induced lateral spread over a free face or gentle-slope condition. 5

Based on Cone Penetration Test (CPT) and Standard Penetration Test (SPT) databases, empirical approaches utilized two-dimensional graphs to visually identify the boundary curve that separates liquefaction and non-liquefaction (Boulanger and Idriss, 2014; Idriss and Boulanger, 2006; Moss et al., 2006). To this end, ML techniques outperform these empirical studies in more objectively capturing the nonlinear and multidimensional relationship between the critical inputs and the triggering of soil liquefaction. Significant studies in this area include the implementations of SVM (Goh and Goh, 2007; Pal, 2006), ANN (Goh, 1994, 1996; Hanna et al., 2007; Juang and Chen, 1999; Ramakrishnan et al., 2008; Ülgen and Engin, 2007), a combination of kernel Fisher discriminant analysis (KFDA) with SVM (Hoang and Bui, 2018), a combination of ANN and RSM (Pirhadi 6

et al., 2018), RF (Kohestani and Ardakani, 2015), stochastic gradient boosting (SGB) (Zhou et al., 2019), generalized linear model (GLM) (Zhang et al., 2013), and evolutionary polynomial regression (EPR) (Rezania et al., 2010, 2011).

Prediction of the amount of lateral displacement under soil liquefaction is a complex problem. Liquefaction-induced lateral spreads involve a large number of influential factors, including earthquake magnitude, fault-to-site distance, and local soil profile information such as the slope of the ground and the fine content and particle sizes of liquefiable sediments (Bartlett and Youd, 1995). Pioneered by the studies from Youd and his coauthors (Bartlett and Youd, 1995; Youd et al., 2002), who used multilinear regression (MLR) to develop predictive equations against a case history database, various ML models have been developed to improve the lateral displacement prediction. These models consist of the use of the ANN (Baziar and Ghorbani, 2005; Chiru-Danzer et al., 2001; Wang and Rahman, 1999), a hybrid neuro-fuzzy procedure (García et al., 2008), SVR (Oommen and Baise, 2010), multivariate adaptive regression splines (MARS) (Goh and Zhang, 2014), RF (Liu and Tesfamariam, 2012), GP and EC (Javadi et al., 2006; Rezania et al., 2011), and multilayer perceptrons (MLPs) and the adaptive neuro-fuzzy inference system (ANFIS) (Kaya, 2016).

These ML methods have provided feasible solutions to tackle soil liquefaction problems. However, most methods somewhat possess their inherent challenges for practical applications, including, for instance, the determination of the architecture for ANN and the hyperparameters for SVM. A majority of the studies failed to carry out the cross-validation (CV) process, which makes the overall model performance questionable. Moreover, the abovementioned studies are purely driven by case histories with discrete data points. As a result, it remains unclear whether the derived models will still be valid when interpolating within or extrapolating beyond the range of conditions constrained by the database. In particular, most of the lateral spread regression models have a prediction accuracy within 200%; namely, the predicted displacements vary from 50% to 200% of the observed values. Such a significant level of uncertainty is mainly caused by the imperfect quality of the case history data sets, which contain considerable subjective information that inevitably prevents an explicit mapping between the inputs and the lateral displacement outputs.

System identification and damage detection⁴

The area of system identification and damage detection consists of a wide array of studies that tackle a mixed collection of problems, ranging from classification of component failure modes using laboratory tests, to the detection of region-wide structural damage using satellite imageries. As such, the aforementioned attribute matrix table is utilized herein to disclose the overall literature structure and the interconnections among all relevant studies. Note that the literature in this topic area is selected to be strictly constrained by the keywords of “machine learning” and “earthquake engineering.” Resultantly, many other studies (e.g. Sen et al., 2019; Yang et al., 2017) that develop data-driven models to identify and infer the health conditions of civil structures agnostic to the triggering event are not included in this review.

System identification⁶

The topic area of system identification comprises a broad branch of studies that develop ML-based models to emulate a structural system and predict its downstream seismic

response. Table A3 in the Supplemental appendix lists all relevant studies and their associated attribute matrix. As is seen, the existing literature apply various ML methods (1a to 1g) on data sets from both laboratory tests (3a) and numerical simulations (3b) to identify the seismic behavior for both structural components (4a) and individual structural systems (4b). For clarification purpose, these studies are subdivided into two groups based on the trait of data resources.

First, laboratory tests on reinforced concrete (RC) structures have provided one source of data that enable ML methods to identify their failure modes, strength, capacities, and constitutive behaviors. In this regard, a handful of studies focused on predicting failure modes and shear strength for beam-column joints (Jeon et al., 2014; Mangalathu and Jeon, 2018; Mitra et al., 2011; Naderpour and Mirrashid, 2019; Yaseen et al., 2018). For instance, Mitra et al. (2011) used LR to categorize the non-ductile joint shear failure versus the ductile beam yielding failure for interior beam-column joints. Mangalathu and Jeon (2018) applied a group of ML tools, that is, LR, LASSO, discriminant analysis, K-nearest neighbors, naïve Bayes classification, SVM, DT, RF, and extreme learning machine (ELM) for classification, and stepwise, ridge, LASSO, elastic net, and RF for regression, to realize failure mode classification and shear strength prediction together. A similar group of ML tools have been utilized by Mangalathu and Jeon (2019a) to predict the flexural, flexural-shear, and shear failure modes of bridge columns (the second row in Figure 3). Computer vision algorithms that incorporate image segmentation, feature extraction, and nonlinear regression analysis were developed by Lattanzi et al. (2015) to estimate peak drift of bridge columns using lateral-load test data. Recently, a multi-output least-squares support vector machine (MLS-SVMR) algorithm was implemented by Luo and Paal (2018) to construct bilinear force-displacement constitutive relationships for RC columns. A similar study has been conducted by Luo and Paal (2019) to predict the drift capacity of RC columns using a locally weighted least square SVR approach. Also, a group of six ML algorithms has been utilized by Huang and Burton (2019) to classify the in-plane failure modes of RC frame structures with masonry infill panels, and 114 test results from infill frame specimens were investigated. Other than RC structures, Farfani et al. (2015) employed the centrifuge test data of a soil-pile-structure system to train, test, and validate the ANN and SVM models in predicting the dynamic characteristics and seismic response of the pile structure. Data sets from dynamic cyclic tests and shaking table tests have been used to develop ML models for an actively controlled frame (Bani-Hani et al., 1999a), a full-scale nonlinear viscous damper (Yun et al., 2008), and a five-floor structure (Zhang et al., 2008). Hybrid methods that couple ANN with wavelet analysis and fuzzy logic have also been examined to simulate the seismic behavior of building frames (Adeli and Jiang, 2006; Hung et al., 2003).

The second group of studies in system identification deals with the data sets from numerical simulations, where data inputs (structural parameters and ground motions) and outputs (structural responses) enable ML methods to build a relationship mapping that emulates a civil structure. To this end, ML methods, particularly ANNs, have been verified to be effective in replacing finite element modeling of civil structures because of their adaptability to approximate complex structures without being constrained by any specific forms. Starting from Conte et al. (1994) who used ANN to learn and simulate the linear elastic behavior of multi-story buildings, ANNs have been implemented in identifying a variety of structures, including building frames (Joghataie and Farrokh, 2008; Xu et al., 2004), concrete gravity dams (Karimi et al., 2010), a prestressed concrete bridge (Jeng and Mo, 2004), a typical embankment (Tsompanakis et al., 2009), and column splices in low-,

medium-, and high-rise steel moment frames (Akbas et al., 2011). Moreover, ANNs have been combined with other soft computing algorithms to minimize the prediction error, increase the training speed, and improve the generalization capability. For instance, a dynamic time-delay fuzzy wavelet ANN model was developed to (1) create a PR model that captures the time series data accurately and efficiently and (2) handle two types of imprecision in the measured data: fuzzy information and measurement uncertainties (Jiang and Adeli, 2005; Jiang et al., 2017). A multi-branch ANN model that separates the structural state variables and seismic inputs was developed to identify a frame structure (Li and Yang, 2007). An intelligent neural system, which combines competitive ANNs and radial basis function ANNs, was developed to enhance accuracy, generality, and speed (Gholizadeh et al., 2009). A functional link ANN that incorporates polynomial functions was also developed by Sahoo and Chakraverty (2018) for improved accuracy and efficiency. Recently, a DL approach, named as the long short-term memory (LSTM) network, was proposed by Zhang et al. (2019b) to model and predict the seismic responses of a nonlinear hysteretic system, a real-world building with field sensing data, and a steel moment resisting frame.

Damage detection²

The topic area of damage detection is broadly defined as a group of studies that develop ML models to recognize, classify, and assess seismic damage of civil structures. As shown in Table A4 in the Supplemental appendix, the existing literature in this area possess a large range of attributes in ML method (1a to 1g), data resource (3a to 3c), and analysis scale (4a to 4c). To be consistent, data resource is used as the main trait to subdivide the relevant studies herein.

First, several studies relied on post-earthquake linguistic or photographic records to predict seismic damage. A major challenge in this area lies in the addressment of damage information in linguistic forms. To this end, De Stefano et al. (1999) used ANN and Bayesian classification to predict seismic damage mechanisms of historic churches. Fuzzy logic models have been utilized in a collection of studies that transform physical descriptions of seismic damage into mathematical model parameters (Allali et al., 2018; Alvanitopoulos et al., 2010; Carreño et al., 2010; Demartinos and Dritsos, 2006; Elwood and Corotis, 2015; Silva and Garcia, 2001). Recently, the linguistic damage records from the 2014 South Napa earthquake have been used to develop a DL-based method that classifies the building damage (Mangalathu and Burton, 2019). On the contrary, the damaged RC column images collected after the 2010 Haiti earthquake have been used by German et al. (2012) to develop a procedure that automatically detects spalled regions on the column surface and measures the properties of the spalling. The multi-step procedure measures the area of spalled concrete, the area of the reinforcement, as well as the sizes of exposed reinforcing bars. This damage detection procedure was further incorporated into a comprehensive framework that links the column damage with the residual drift capacity and post-earthquake fragility curves of RC structures (German et al., 2013; Paal et al., 2014). Besides, ML methods have been implemented in dealing with satellite imageries and digital maps to detect and classify building damage (Gong et al., 2016; Peyk-Herfeh and Shahbahrami, 2014). Gao and Mosalam (2018) have also constructed an image database called “Structural ImageNet,” from which two DL technologies such as transfer learning (TL) and visual geometry group (VGGNet) were applied to recognize structural damage caused by earthquakes and other natural hazards.

A large part of the existing literature uses simulated and test data to detect the seismic damage of building structures. In particular, an ANN model is first trained with respect to the reference system in its undamaged state, whereas the response data from the damaged state of the same system are fed into the same model. As a result, the variation in the level of prediction error between the two states can serve as a reference to quantify the structural damage in a nonparametric manner (Huang et al., 2003; Nakamura et al., 1998). Similar studies have considered using innovative metrics such as Bayes factors, natural modes, and coefficients of autoregressive models for damage detection (De Lautour and Omenzetter, 2010; González and Zapico, 2008; Jiang and Adeli, 2007; Jiang and Mahadevan, 2008). Following the same logic, a couple of studies have improved the approach to enable parametric quantification of structural damage (e.g. damage quantified through the change of stiffness values) (Wu et al., 2002; Xu et al., 2005). In a broader context, ANN models have been developed to predict the seismic response for a variety of structures so as to infer their damage conditions. Related studies in this area include (1) quick earthquake damage estimation on ordinary wooden framed houses in Japan (Molas and Yamazaki, 1995); (2) seismic vulnerability assessment of chemical industrial plants with various topologies (Aoki et al., 2002); (3) damage index prediction of RC frames (De Lautour and Omenzetter, 2009; Morfidis and Kostinakis, 2017, 2018); (4) seismic damage evaluation of concrete shear walls (Vafaei et al., 2013) and cantilever structures (Vafaei et al., 2014); and (5) global damage classification of RC slab-column frames by combining ANN with SVM (Kia and Sensoy, 2014). ML has also been utilized by Burton and his coworkers (Burton et al., 2017; Zhang and Burton, 2019; Zhang et al., 2018) to link the seismic damage patterns of buildings to the residual structural capacity indices (i.e. the median capacity ratio between the intact and damaged buildings). Their proposed framework integrates seismic demand analysis, component damage simulation, and residual collapse capacity estimation on both intact and damaged structures. The applied ML algorithms involve CART and RF for safety classification, and LASSO and SVM for capacity index prediction.

Seismic fragility assessment²

Performance-based earthquake engineering (PBEE) is a multi-step framework that employs seismic hazard analysis, structural analysis, damage assessment, and risk assessment to provide scientific estimates of the structural performance in terms of metrics such as casualties, repair cost, and downtime (Porter, 2003). In this framework, PSDMs capture the relationships between the structural responses (quantified by engineering demand parameters (EDPs)) and the IMs of earthquake motions. By combining the demand models with capacity limit states, fragility functions are developed to estimate the damage probability of the structure given earthquake IM and other potential predictors of the system. Moreover, the rate of limit state exceedance at a site-specific structure can be calculated by convolving the fragility model with the seismic hazard curve at the site or region. These components in the PBEE have been utilized in numerous studies to investigate the seismic vulnerability and mitigation schemes for individual structures, where earthquake hazard is the primary source of uncertainty (Xie et al., 2019b; Xie and Zhang, 2016, 2018).

Fragility models are one of the critical components in PBEE frameworks. The first class of seismic fragility models provides an estimate of the probability of limit state exceedance in terms of a single variable, the earthquake IM (Cornell et al., 2002), often for structure specific analysis. When considering portfolios of structures with significant variations in the structural parameters, work has considered analysis of archetype structures or groups of representative structures to derive class fragilities. Recent work supporting fragility and

risk modeling of regional portfolios of structures considers PSDMs with multiple predictors that reflect the variation across a portfolio and consequently multidimensional fragility models (Ebad Sichani et al., 2018; Jeon et al., 2019; Kameshwar and Padgett, 2014; Mangalathu et al., 2018a; Xie et al., 2019a; Zhang et al., 2019a). Namely, the demand and fragility models can be tailored to structures across a region since they are dependent on not only earthquake IMs but also other significant uncertain parameters such as material, geometric, and aging parameters of the structure. As shown in the third row of Figure 3, procedures to develop multi-predictor PSDMs and multidimensional fragility models include (1) design an experiment that captures the variations of all uncertain parameters; (2) generate a group of motion-structure samples for seismic demand analysis; (3) develop multi-predictor PSDMs using surrogate metamodeling; and (4) apply Monte Carlo simulation or other applicable tools to generate multidimensional fragility models. In this process, ML methods have exhibited significant promises in facilitating the development of both PSDMs and fragility functions.

First, the high-dimensional nonlinear relationship between the predictors and the EDPs of concern can be efficiently quantified through ML methods. Regression-based surrogate metamodeling can be coupled with efficient statistical sampling to develop multi-predictor PSDMs, which can provide approximate estimates of EDP distributions. As a result, time and efforts can be saved since fewer numerical simulations are needed to cover the dispersion ranges for a large number of uncertain parameters. In this regard, multi-predictor PSDMs have been prevalently derived using RSMs, including the response surface with random block effects (Buratti et al., 2010), dual response surface (Perotti et al., 2013), and RSM with polynomial basis functions (De Felice and Giannini, 2010; De Grandis et al., 2009; Liel et al., 2009; Pan et al., 2007; Park and Towashiraporn, 2014; Rajeev and Tesfamariam, 2012; Ravi Kiran et al., 2019; Ricci et al., 2013; Saha et al., 2016; Seo et al., 2012; Seo and Linzell, 2010, 2012, 2013; Seo and Park, 2017; Verderame et al., 2014). Stepwise regression and other regularization algorithms have been used along with RSMs to identify the most informative predictors for developing PSDMs. For instance, Ebad Sichani et al. (2018) employed a stepwise RSM to establish a two-layer PSDM of concrete dry cask structures that are vulnerable to sliding, wobbling, and rocking in seismic events. Xie and DesRoches (2019) tested stepwise and LASSO regression models for probabilistic seismic demand analysis of California highway bridges. Other than RSMs, multi-predictor PSDMs have been developed using ANN (Calabrese and Lai, 2013; Lagaros et al., 2009; Lagaros and Fragiadakis, 2007; Liu and Zhang, 2018; Mitropoulou and Papadrakakis, 2011; Pang et al., 2014; Wang et al., 2018), bootstrapped ANN (Ferrario et al., 2017), SVM (Ghosh et al., 2018; Hariri-Ardebili and Pourkamali-Anaraki, 2018; Huang et al., 2017; Mahmoudi and Chouinard, 2016), kriging metamodeling (Gidaris et al., 2015), GLM (Xie et al., 2019a), MARS (Kameshwar and Padgett, 2014), K-nearest neighbor (Hariri-Ardebili and Pourkamali-Anaraki, 2018), naïve Bayes classifier (Hariri-Ardebili and Pourkamali-Anaraki, 2018), high-dimensional model representation (Sahu et al., 2019), and RF (Mangalathu and Jeon, 2019b).

Once the PSDMs are obtained, Monte Carlo simulations can be used to convolve the PSDMs with the capacity models to develop multidimensional fragility functions. Such fragility models typically have no explicit mathematical expression and cannot be easily reproduced. To address this issue, researchers have used an alternative approach to compare the demand versus capacity and generate binary survival-failure samples, from which additional models are trained (often adopting the LR model) to develop a parameterized fragility model. Relevant studies in this area have developed LR-based fragility models for

highway bridges (Dukes et al., 2018; Ghosh et al., 2013, 2014; Ghosh and Sood, 2016; Jeon et al., 2019; Kameshwar and Padgett, 2014, 2018; Mangalathu et al., 2018a, 2018b), single-degree-of-freedom structures on liquefiable sand deposit (Koutsourelakis, 2010), rigid blocks installed with safety devices (Contento et al., 2017), and RC shear walls (Yazdi et al., 2016). In particular, the LR-based fragility models feature logistic expressions that explicitly quantify the influences of the earthquake IM and any other significant structure parameters, enabling the fragility analysis across multiple scales. 1

Research advances can be further pursued on this portfolio of work to leverage the full capability and efficiency of available ML methods. For example, despite that one structure may have multiple EDPs of interest, most of the existing work develops separate PSDM for each EDP. To this end, multivariate PSDMs that incorporate all EDPs in a single model can not only save multiple rounds of model calibration but also capture the correlation among different EDPs in the same structure, resulting in more realistic demand models (Du and Padgett, 2019; Goda and Tesfamariam, 2015; Luco et al., 2005). Also, alternative ML methods should be explored to avoid using too many correlated predictors in one PSDM. Recent studies have shown promise in developing sparse PSDMs that consist of a small subset of uncorrelated predictors for better accuracy and generalization capability (e.g. Mangalathu et al., 2018b). Moreover, parameterized fragility models have been primarily developed through LR, while it remains unclear whether there exist other models that can provide significant advantages over existing LR models. 2

Structural control for earthquake mitigation 3

Previous studies have demonstrated the effectiveness of utilizing active control methods and semi-active control devices in attenuating earthquake-induced structural vibrations (Housner et al., 1997). In essence, structural control requires the identification and analysis of various components in the controlled system, including actuator dynamics, nonlinear structures, and real-time measurements. Compared with conventional control methods that are programmed to perform a specific task, the incorporation of ML in the control loop allows the controllers to learn the task, either in a supervised or unsupervised manner (Ghaboussi and Joghataie, 1995). Such adaptive learning mechanism has provided inherent capability to handle nonlinear mapping, incorporate delays, and recover from partial system failure (Wen et al., 1995), all of which make ML-based structural control more effective in mitigating the adverse effects of earthquake hazards. 4

Active control using ANN 5

Considerable studies have explored the soundness of engaging ANN in designing active control schemes to alleviate the seismic impacts on buildings. Developed by Ghaboussi and coworkers, the neural-controller (i.e. the actuator controlled by the neural network instead of an ad hoc control algorithm) is trained with the aid of the emulator neural network, which learns the transfer function between the actuator signal and the output of the sensors measuring the response of the structure (Bani-Hani et al., 1999a, 1999b; Bani-Hani and Ghaboussi, 1998; Ghaboussi and Joghataie, 1995; Wen et al., 1995). It is worth mentioning that the emulator not only learns the structural behavior but also incorporates the effects of the actuator dynamics and sampling period. The neural-controller was trained in an iterative process to make sure the controlled outputs are within a specified tolerance when compared with the structural responses from the emulator neural network (Bani-Hani et al., 1999b). As an example shown in the fourth row of Figure 3, the neuro-control 6

training scheme is composed of the following steps: (1) send a random control signal to the actuator when the structure is subjected to the base excitation; (2) collect the structural responses and send them to the control criterion box; (3) simultaneously feed the control signal to the emulator neural network and send the output to the control criterion box; (4) calculate the error between the two outputs and back-propagate it through the emulator neural network and the neuro-controller; and (5) modify the connection weights of the neuro-controller until the control criterion is satisfied (Bani-Hani and Ghaboussi, 1998). After the neural-controller is well trained, it can generate appropriate signals to the actuator based on the feedback signal from the sensors.

The neuro-controller concept was further developed in several studies that enhance the efficiency and robustness of the active control scheme (Brown and Yang, 2001; Khodabandolehlou et al., 2018; Rao and Datta, 2006; Subasri et al., 2014; Yakut and Alli, 2011). Improvements have been made in the following studies: (1) the use of a cost function to train the ANN (Kim et al., 2000); (2) the use of a sensitivity evaluation algorithm to replace the emulator neural network for saving the training time (Kim and Lee, 2001); (3) the development of a counter propagation network (CPN) to realize unsupervised learning (Madan, 2005); (4) the use of lattice forms in the training pattern to save the calculation efforts and accelerate the training process (Kim et al., 2008); (5) the utilization of an extended minimal resource allocation network to accomplish real-time online adaptation of the ANNs (Suresh et al., 2010). Recently, a fuzzy wavelet neuro-emulator model was developed by Jiang and Adeli (2008a) that can predict the nonlinear structural response in future time steps only from the immediate past structural response and actuator dynamics. This model was further combined with the floating point GA in a companion study by the authors (Jiang and Adeli, 2008b), which identified the optimum control forces at any time step.

Semi-active control using ANN ³

Semi-active control devices have the capability of adapting to the changes in earthquake loading conditions, similar to the fully active systems, yet without requiring access to large power supplies. In this regard, interest in the use of ANN has grown remarkably, particularly for structures that are designed with magnetorheological (MR) dampers. Contributions in relevant studies include (1) an ANN model to represent the nonlinear differential equations and simulate the dynamic behavior of the MR damper (Chang and Roschke, 1998); (2) an inverse optimal ANN model to predict the required voltage given the desired force of the MR damper (Xia, 2003); (3) the use of ANN to replicate the damper's dynamics and induce the MR damper in controlling the seismic shaking of non-isolated and isolated structures (Bani-Hani and Sheban, 2006; Xu et al., 2003). In a study by Lee et al. (2006), the neuro-controller was trained to minimize a cost function that includes both structural responses and control signals, and a clipped algorithm was employed to actuate the damper and generate the desired control force. This work has been modified to use modal coordinates as the neuro-controller inputs, which facilitated the controller design (Lee et al., 2008). Moreover, a hybrid neuro-fuzzy control strategy for MR dampers was proposed by Xu and Guo (2008), who tackled the time-delay issue using the ANN and reduced the responding time of the control currents through the fuzzy control. Control efficiency and robustness of implementing neuro-controllers for MR dampers were further verified by Bozorgvar and Zahrai (2019) by combining the fuzzy inference system with ANN and GA. Other than the MR damper, the ANN has been utilized in modeling the MR elastomer isolator (Fu et al., 2016) and actuating the variable

orifice fluid damper to control the response of a structure isolated by a device named as curved surface slider (Krishnamoorthy et al., 2017). 1

The existing literature focuses on developing ANN models in active and semi-active control, while more advanced learning algorithms, such as reinforcement learning (Khalatbarisoltani et al., 2019), should be explored to realize robust control in tackling various sources of uncertainties simultaneously, including seismic uncertainties and time delays. Meanwhile, the promise of implementing ML in structural control can be enhanced if other forms of control devices are investigated, such as active tuned mass damper, distributed actuators, and semi-active stiffness dampers (Fisco and Adeli, 2011). Although theoretical studies and experimental campaigns have verified its effectiveness, structural control is facing cost and reliability issues that limit its real-world applications (Casciati et al., 2012). In this regard, ML-based structural control should also provide viable solutions for easier implementation and maintenance of control systems. 2

Discussions and conclusion 3

This article reviewed the significance of various ML techniques for earthquake engineering applications. Four topic areas of ML implementations include seismic hazard analysis, system identification and damage detection, seismic fragility assessment, and structural control for earthquake mitigation. Although limited in number of data for some cases, data sets collected from laboratory and field tests, previous earthquake events, and numerical simulations have enabled researchers to practice a collection of advanced ML tools. Technical background of some commonly used ML methods was presented, whereas the adaptation and application of these tools in each of the four topic areas were discussed in detail. The study revealed that ML techniques have the capability to learn and infer complicated interrelations among the contributing parameters, and thus allow tackling diverse problems in earthquake engineering that are challenging, or not possible, to solve using traditional methods. As observed during the review process, the cross field of ML and earthquake engineering is a new, but increasingly dynamic area for high impact research, where a vast breadth and depth of topics can be investigated. For the purpose of further promoting ML applications in earthquake engineering, the potential challenges and associated research needs are discussed herein. 4

Data quantify and quality 5

ML typically requires large amounts of high-quality data for learning to be effective. Data availability has become less an issue in the field of GMPE, in which tens of thousands of input motion data sets are available in the newly developed NGA-West2 strong motion database (Ancheta et al., 2014). However, for some areas that require high-fidelity computational analyses or large-scale field tests, high-quality data points are often limited to hundreds or fewer. As is expected, ML in these areas is facing a strong challenge. For instance, numerous existing RC frame buildings and bridges in seismic active regions were constructed prior to the implementation of modern seismic design provisions, leaving the RC columns susceptible to brittle shear failures in these structures. Recently, ML has been utilized to predict the displacement capacity of RC columns, where input data that are available for shear-critical columns are much fewer than those for flexure-critical columns. Consequently, ML models for shear-critical columns do not perform as well as those for flexure failures (Luo and Paal, 2019). Note that such data quantity issues cannot be easily tackled by switching or developing a more advanced ML model (Domingos, 2012). Also, 6

low data quality often leads to inferior ML models in earthquake engineering. As previously mentioned, on a separate problem of lateral spread prediction under soil liquefaction, researchers tended to acknowledge the chaotic response of liquefying soil masses to earthquakes, and most of the ML models indicated that the predicted displacements lie between one-half and twice the measured counterparts (Goh and Zhang, 2014). Although this degree of accuracy is deemed to be comparable to other geotechnical predictions, such as estimated ground settlement from consolidation tests (Youd, 2018), ML would be expected to provide much more accurate predictions should higher quality data points become accessible. To be specific, a crucial aspect that causes the inaccuracies on lateral spread predictions is the subjective information contained in the raw database. As proof, the scattering caused by the vague and subjective information in the data has been smoothed by García et al. (2008) through a fuzzy clustering technique, after which a neural network procedure was applied that substantially improved the model accuracy.

To address this data issue, research efforts are suggested in the following directions. First, more transparent, accessible, and high-quality data are needed to be compiled in a computer-readable form. As more journals in earthquake engineering encourage researchers to make the data associated with their publications available, a widely accepted platform is required to store and share such data using a standard data structure. In ML applications, although significant expertise is required and effort directed at model tuning, training, validation and interpretation, generally, the time spent running ML is much less in comparison with the time to gather data, integrate it, clean it, and pre-process it (Domingos, 2012). In this regard, it is vital that in earthquake engineering, there exists a community-driven cyberinfrastructure embraced by the community that allows researchers to share and analyze data more effectively, integrate diverse data sets, and practice and develop ML tools. One such cyberinfrastructure platform is the DesignSafe (<https://www.designsafe-ci.org/>) (Rathje et al., 2017) that supports natural hazards engineering research, through which various recently generated data sets have been uploaded and shared (e.g. Brewick et al., 2019; Hutchinson et al., 2019; Kameshwar et al., 2019; Mosalam et al., 2019; Stewart et al., 2016). Also, since experimental laboratory or field derived databases are limited and initiating new large-scale campaigns can be time-consuming and cost-prohibitive, researchers should be encouraged to provide more simulation-based data, especially for those that have high quality, are physics-driven, and are well-validated against existing test results. While other communities of practice embrace sharing models and output from simulations, this culture is still in its infancy in the earthquake engineering community. However, many such high-fidelity models also have great merit in providing training data for ML applications. Moreover, researchers conducting ML in earthquake engineering are expected to increasingly deal with new sources of data generated from other cutting-edge technologies, such as wireless sensing, computer vision, Internet of things (IoT), smart cities, geographic information system (GIS), and quantum computing, and so on. Once earthquakes occur, these technologies can provide new forms of data on a completely different scale. For instance, the IoT for smart cities can use real-time connected sensors to generate a large amount of spatiotemporal data about bridge networks or pipeline systems for the entire region (Salehi and Burgueño, 2018). To this end, more ambitious problems in earthquake engineering can be tackled by integrating ML with these new technologies. Such problems can be high-fidelity seismic performance prediction of a structural component at the granular level or the seismic resilience assessment of inter-connected infrastructure systems on the regional scale.

Implementation and development of ML methods¹

Given potential rapid data growth due to abovementioned technologies, ML is expected to provide a tremendous opportunity to systematically advance the research and practice in earthquake engineering. However, the next generation of spatiotemporal data, which tend to be large-scale, high-dimensional, nonlinear, non-stationary, and heterogeneous, are expected to challenge the capabilities of existing ML methods often adopting in the earthquake engineering domain. To this end, more advanced ML techniques, such as active learning, reinforcement learning, and deep learning (Andriotis and Papakonstantinou, 2019; Blatman and Sudret, 2010; Cadini et al., 2014; Echard et al., 2011; Kong et al., 2020; Memarzadeh and Pozzi, 2019; Nguyen and Medjaher, 2019; Schöbi et al., 2017), are needed to (1) characterize the higher-order correlation and dependencies within the data, (2) perform efficient and reliable imputation and prediction for decision making, and (3) develop scalable learning models for large-scale and time-dependent problems.

Moreover, there exists an emerging trend for a paradigm shift that requires earthquake engineering researchers to consider how to best balance the use of physics-based approaches, which are transparent, interpretable, and somewhat predictable, with the use of data-driven ML models that are not unique and sometimes hardly interpretable. Distinct from physics-based approaches, ML algorithms learn a data science model using a set of algorithms, rules, and criteria that automatically extract the relationships between input and output variables. Such relationships are entirely data-driven and cannot alone explain the physical cause–effect mechanisms between variables. To be specific, although spurious relationships can be learned for a complex problem that look deceptively accurate on training and test sets, the model may perform much worse outside the available labeled data (Karpatne et al., 2017).

A rational path forward lies in the increasing incorporation of physical knowledge into ML-based earthquake engineering studies. If the ML model fails to deliver a physical understanding of the underlying process, it hardly can be used as a generalizable solution to solve other similar problems. As previously investigated in diverse disciplines (Faghmous and Kumar, 2014; Fischer et al., 2006; Wagner and Rondinelli, 2016), theory-guided ML can be developed through a variety of approaches (Karpatne et al., 2017). Opportunities exist to accelerate work along this path. First, domain experts can take the lead in converting raw data into a new feature space that reflects better the scientific nature of the underlying problem. For example, dimensionally consistent predictors and observables make more scientific sense to form an RSM than the raw data in their original units (Xie et al., 2019a). Second, an ensemble of different ML algorithms, or similar algorithms with different values for their internal parameters, should be examined to create a more robust overall model (Butler et al., 2018). Recent works in earthquake engineering have embraced this trend to use multiple ML algorithms in solving the same problem, where the algorithm with the best performance was identified (Luo and Paal, 2019; Mangalathu and Jeon, 2018; Sichani, 2018; Thomas et al., 2017). Third, physical understanding of a problem can be increasingly used to design and learn ML models. For example, ANN can be trained to solve supervised learning tasks while respecting the physical law described by general nonlinear partial differential equations (Raissi and Karniadakis, 2018).

In summary, despite the growing number of studies every year, the implementation of ML in earthquake engineering is still in its early stage when compared with other disciplines. However, supported by the next generation of diverse data sharing and sensor technologies, ML has a great promise to revolutionize the profession of earthquake

engineering. Furthermore, the earthquake engineering community has the opportunity to probe unexplored ML algorithms in various contexts, or inspire new ones driven by our application needs, while opening dialogue on best practices to integrate physics-based and data-driven methods to solve grand challenges in earthquake engineering.


Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Majid Ebad Sichani  <https://orcid.org/0000-0003-0494-3287>

Supplemental material

Supplemental material for this article is available online.

References

- Adeli H (2001) Neural networks in civil engineering: 1989-2000. *Computer-Aided Civil and Infrastructure Engineering* 16(2): 126–142.
- Adeli H and Jiang X (2006) Dynamic fuzzy wavelet neural network model for structural system identification. *Journal of Structural Engineering* 132(1): 102–111.
- Akbas B, Shen J and Sabol TA (2011) Estimation of seismic-induced demands on column splices with a neural network model. *Applied Soft Computing* 11(8): 4820–4829.
- Akhani M, Kashani AR, Mousavi M and Gandomi AH (2019) A hybrid computational intelligence approach to predict spectral acceleration. *Measurement* 138: 578–589.
- Akkar S, Sandikkaya MA, Şenyurt M, Sisi AA, Ay B, Traversa P, Douglas J, Cotton F, Luzi L, Hernandez B and Godey S (2014) Reference database for seismic ground-motion in Europe (RESORCE). *Bulletin of Earthquake Engineering* 12(1): 311–339.
- Alavi AH and Gandomi AH (2011) Prediction of principal ground-motion parameters using a hybrid method coupling artificial neural networks and simulated annealing. *Computers and Structures* 89: 2176–2194.
- Alavi AH, Gandomi AH, Modaresnezhad M and Mousavi M (2011) New ground-motion prediction equations using multi expression programming. *Journal of Earthquake Engineering* 15(4): 511–536.
- Alimoradi A and Beck JL (2015) Machine-learning methods for earthquake ground motion analysis and simulation. *Journal of Engineering Mechanics* 141(4): 04014147.
- Allali SA, Abed M and Mebarki A (2018) Post-earthquake assessment of buildings damage using fuzzy logic. *Engineering Structures* 166: 117–127.
- Alvanitopoulos PF, Andreadis I and Elenas A (2010) Neuro-fuzzy techniques for the classification of earthquake damages in buildings. *Measurement* 43(6): 797–809.
- Amezquita-Sanchez JP, Valtierra-Rodriguez M, Aldwaik M and Adeli H (2016) Neurocomputing in civil infrastructure. *Scientia Iranica* 23(6): 2417–2428.
- Amiri GG, Khorasani M, Hessabi RM and Amrei SAR (2010) Ground-motion prediction equations of spectral ordinates and arias intensity for Iran. *Journal of Earthquake Engineering* 14(1): 1–29.
- Ancheta TD, Darragh RB, Stewart JP, Seyhan E, Silva WJ, Chiou BSJ, Wooddell KE, Graves RW, Kottke A, Boore DM, Kishida T and Donahue J (2014) NGA-West2 database. *Earthquake Spectra* 30(3): 989–1005.

- Andriotis CP and Papakonstantinou KG (2019) Managing engineering systems with large state and action spaces through deep reinforcement learning. *Reliability Engineering and System Safety* 191: 106483.
- Aoki T, Ceravolo R, De Stefano A, Genovese C and Sabia D (2002) Seismic vulnerability assessment of chemical plants through probabilistic neural networks. *Reliability Engineering and System Safety* 77(3): 263–268.
- Bakhshi H, Bagheri A, Ghodrati Amiri G and Barkhordari MA (2014) Estimation of spectral acceleration based on neural networks. *Proceedings of the Institution of Civil Engineers—Structures and Buildings* 167(8): 457–468.
- Bani-Hani K and Ghaboussi J (1998) Neural networks for structural control of a benchmark problem, active tendon system. *Earthquake Engineering and Structural Dynamics* 27(11): 1225–1245.
- Bani-Hani K, Ghaboussi J and Schneider SP (1999a) Experimental study of identification and control of structures using neural network. Part 1: Identification. *Earthquake Engineering and Structural Dynamics* 28(9): 995–1018.
- Bani-Hani K, Ghaboussi J and Schneider SP (1999b) Experimental study of identification and control of structures using neural network. Part 2: Control. *Earthquake Engineering and Structural Dynamics* 28(9): 1019–1039.
- Bani-Hani KA and Sheban MA (2006) Semi-active neuro-control for base-isolation system using magnetorheological (MR) dampers. *Earthquake Engineering and Structural Dynamics* 35(9): 1119–1144.
- Bartlett SF and Youd TL (1995) Empirical prediction of liquefaction-induced lateral spread. *Journal of Geotechnical Engineering* 121(4): 316–329.
- Baziar MH and Ghorbani A (2005) Evaluation of lateral spreading using artificial neural networks. *Soil Dynamics and Earthquake Engineering* 25(1): 1–9.
- Bird JF, Bommer JJ, Crowley H and Pinho R (2006) Modelling liquefaction-induced building damage in earthquake loss estimation. *Soil Dynamics and Earthquake Engineering* 26(1): 15–30.
- Bishop CM (2006) *Pattern Recognition and Machine Learning* (ed M Jordan, J Kleinberg and B Scholkopf). New York: Springer.
- Blatman G and Sudret B (2010) An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. *Probabilistic Engineering Mechanics* 25(2): 183–197.
- Boore DM and Atkinson GM (2008) Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s. *Earthquake Spectra* 24(1): 99–138.
- Boore DM, Stewart JP, Seyhan E and Atkinson GM (2014) NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes. *Earthquake Spectra* 30(3): 1057–1085.
- Boulanger RW and Idriss IM (2014) *CPT and SPT based liquefaction triggering procedures*. Report No. UCD/CGM-14/01, April. Davis, CA: University of California at Davis.
- Box GEP and Hunter JS (1957) Multi-factor experimental designs for exploring response surfaces. *The Annals of Mathematical Statistics* 28: 195–241.
- Bozorgvar M and Zahrai SM (2019) Semi-active seismic control of buildings using MR damper and adaptive neural-fuzzy intelligent controller optimized with genetic algorithm. *Journal of Vibration and Control* 25(2): 273–285.
- Brewick P, Yu T, Johnson E, Sato E, Sasaki T and Christenson R (2019) *NEESR E-Defense Base Isolation 2013: System Identification*. DesignSafe-CI. DOI: 10.17603/ds2-3d7a-x318.
- Brown AS and Yang HTY (2001) Neural networks for multiobjective adaptive structural control. *Journal of Structural Engineering* 127(2): 203–210.
- Buratti N, Ferracuti B and Savoia M (2010) Response Surface with random factors for seismic fragility of reinforced concrete frames. *Structural Safety* 32(1): 42–51.
- Burton HV, Sreekumar S, Sharma M and Sun H (2017) Estimating aftershock collapse vulnerability using mainshock intensity, structural response and physical damage indicators. *Structural Safety* 68: 85–96.

- Butler KT, Davies DW, Cartwright H, Isayev O and Walsh A (2018) Machine learning for molecular and materials science. *Nature* 559(7715): 547–555.
- Cabalar AF and Cevik A (2009) Genetic programming-based attenuation relationship: An application of recent earthquakes in turkey. *Computers and Geosciences* 35(9): 1884–1896.
- Cadini F, Santos F and Zio E (2014) An improved adaptive kriging-based importance technique for sampling multiple failure regions of low probability. *Reliability Engineering and System Safety* 131: 109–117.
- Calabrese A and Lai CG (2013) Fragility functions of blockwork wharves using artificial neural networks. *Soil Dynamics and Earthquake Engineering* 52: 88–102.
- Carreño ML, Cardona OD and Barbat AH (2010) Computational tool for post-earthquake evaluation of damage in buildings. *Earthquake Spectra* 26(1): 63–86.
- Casciati F, Rodellar J and Yildirim U (2012) Active and semi-active control of structures-theory and applications: A review of recent advances. *Journal of Intelligent Material Systems and Structures* 23(11): 1181–1195.
- Chang C and Roschke P (1998) Neural network modeling of a magnetorheological damper. *Journal of Intelligent Material Systems and Structures* 9: 755–764.
- Chiou B, Darragh R, Gregor N and Silva W (2008) NGA project strong-motion database. *Earthquake Spectra* 24(1): 23–44.
- Chiru-Danzer M, Juang CH, Christopher RA and Suber J (2001) Estimation of liquefaction-induced horizontal displacements using artificial neural networks. *Canadian Geotechnical Journal* 38(1): 200–207.
- Conte JP, Durrani AJ and Shelton RO (1994) Seismic response modeling of multi-story buildings using neural networks. *Journal of Intelligent Material Systems and Structures* 5(3): 392–402.
- Contento A, Gardoni P, Egidio A and Di De Leo AM (2017) Probability models to assess the seismic safety of rigid block-like structures and the effectiveness of two safety devices. *Procedia Engineering* 199: 1164–1169.
- Cornell CA, Jalayer F, Hamburger RO and Foutch DA (2002) Probabilistic basis for 2000 SAC federal emergency management agency steel moment frame guidelines. *Journal of Structural Engineering* 128: 526–533.
- Cubrinovski M, Winkley A, Haskell J, Palermo A, Wotherspoon L, Robinson K, Bradley B, Brabhaharan P and Hughes M (2014) Spreading-induced damage to short-span bridges in Christchurch, New Zealand. *Earthquake Spectra* 30(1): 57–83.
- De Felice G and Giannini R (2010) An efficient approach for seismic fragility assessment with application to old reinforced concrete bridges. *Journal of Earthquake Engineering* 14(2): 231–251.
- De Grandis S, Domaneschi M and Perotti F (2009) A numerical procedure for computing the fragility of NPP components under random seismic excitation. *Nuclear Engineering and Design* 239(11): 2491–2499.
- De Lautour RO and Omenzetter P (2009) Prediction of seismic-induced structural damage using artificial neural networks. *Engineering Structures* 31(2): 600–606.
- De Lautour RO and Omenzetter P (2010) Damage classification and estimation in experimental structures using time series analysis and pattern recognition. *Mechanical Systems and Signal Processing* 24(5): 1556–1569.
- De Stefano A, Sabia D and Sabia L (1999) Probabilistic neural networks for seismic damage mechanisms prediction. *Earthquake Engineering & Structural Dynamics* 28(8): 807–821.
- Demartinos K and Dritsos S (2006) First-level pre-earthquake assessment of buildings using fuzzy logic. *Earthquake Spectra* 22(4): 865–885.
- Derakhshani A and Foruzan AH (2019) Predicting the principal strong ground motion parameters: A deep learning approach. *Applied Soft Computing* 80: 192–201.
- Derras B, Bard PY and Cotton F (2014) Towards fully data driven ground-motion prediction models for Europe. *Bulletin of Earthquake Engineering* 12(1): 495–516.
- Dhanya J and Raghukanth STG (2018) Ground motion prediction model using artificial neural network. *Pure and Applied Geophysics* 175(3): 1035–1064.
- Domingos P (2012) A few useful things to know about machine learning. *Communications of the ACM* 55(10): 78.

- 1 Douglas J (2003) Earthquake ground motion estimation using strong-motion records: A review of equations for the estimation of peak ground acceleration and response spectral ordinates. *Earth-Science Reviews* 61(1–2): 43–104.
- Du A and Padgett JE (2019) Multivariate surrogate demand modeling of highway bridge structures. In: *12th Canadian conference on earthquake engineering*, Quebec City, Canada, 17–20 June, p. 8.
- Dukes J, Mangalathu S, Padgett JE and DesRoches R (2018) Development of a bridge-specific fragility methodology to improve the seismic resilience of bridges. *Earthquake and Structures* 15(3): 253–261.
- Ebad Sichani M, Padgett JE and Bisadi V (2018) Probabilistic seismic analysis of concrete dry cask structures. *Structural Safety* 73: 87–98.
- Echard B, Gayton N and Lemaire M (2011) AK-MCS: An active learning reliability method combining kriging and Monte Carlo Simulation. *Structural Safety* 33(2): 145–154.
- Elwood E and Corotis RB (2015) Application of fuzzy pattern recognition of seismic damage to concrete structures. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* 1(4): 04015011.
- Faghmous JH and Kumar V (2014) A big data guide to understanding climate change: The case for theory-guided data science. *Big Data* 2(3): 155–163.
- Farfani HA, Behnamfar F and Fathollahi A (2015) Dynamic analysis of soil-structure interaction using the neural networks and the support vector machines. *Expert Systems with Applications* 42(22): 8971–8981.
- Ferrario E, Pedroni N, Zio E and Lopez-Caballero F (2017) Bootstrapped artificial neural networks for the seismic analysis of structural systems. *Structural Safety* 67: 70–84.
- Fischer CC, Tibbetts KJ, Morgan D and Ceder G (2006) Predicting crystal structure by merging data mining with quantum mechanics. *Nature Materials* 5(8): 641–646.
- Fisco NR and Adeli H (2011) Smart structures: Part I—Active and semi-active control. *Scientia Iranica* 18(3A): 275–284.
- Fu J, Liao G, Yu M, Li P and Lai J (2016) NARX neural network modeling and robustness analysis of magnetorheological elastomer isolator. *Smart Materials and Structures* 25(12): 1–17.
- Fuller R (1995) *Neural Fuzzy Systems*. CBE Life Sciences Education, vol. 9. Turku: Abo Akademi University.
- Gandomi AH, Alavi AH, Mousavi M and Tabatabaei SM (2011) A hybrid computational approach to derive new ground-motion prediction equations. *Engineering Applications of Artificial Intelligence* 24(4): 717–732.
- Gao Y and Mosalam KM (2018) Deep transfer learning for image-based structural damage recognition. *Computer-Aided Civil and Infrastructure Engineering* 33(9): 748–768.
- García SR, Romo MP and Botero E (2008) A neurofuzzy system to analyze liquefaction-induced lateral spread. *Soil Dynamics and Earthquake Engineering* 28(3): 169–180.
- German S, Brilakis I and Desroches R (2012) Rapid entropy-based detection and properties measurement of concrete spalling with machine vision for post-earthquake safety assessments. *Advanced Engineering Informatics* 26(4): 846–858.
- German S, Jeon J-S, Zhu Z, Bearman C, Brilakis I, DesRoches R and Lowes L (2013) Machine vision-enhanced postearthquake inspection. *Journal of Computing in Civil Engineering* 27(6): 622–634.
- Ghaboussi J and Joghataie A (1995) Active control of structures using neural networks. *Journal of Engineering Mechanics* 121(4): 555–567.
- Gholizadeh S, Salajegheh J and Salajegheh E (2009) An intelligent neural system for predicting structural response subject to earthquakes. *Advances in Engineering Software* 40(8): 630–639.
- Ghosh J and Sood P (2016) Consideration of time-evolving capacity distributions and improved degradation models for seismic fragility assessment of aging highway bridges. *Reliability Engineering and System Safety* 154: 197–218.
- Ghosh J, Padgett JE and Dueñas-Orsorio L (2013) Surrogate modeling and failure surface visualization for efficient seismic vulnerability assessment of highway bridges. *Probabilistic Engineering Mechanics* 34: 189–199.

- Ghosh J, Rokneddin K, Padgett JE and Dueñas-Orsorio L (2014) Seismic reliability assessment of aging highway bridge networks with field instrumentation data and correlated failures, I: Methodology. *Earthquake Spectra* 30(2): 795–817.
- Ghosh S, Roy A and Chakraborty S (2018) Support vector regression based metamodeling for seismic reliability analysis of structures. *Applied Mathematical Modelling* 64: 584–602.
- Gidasir I, Taflanidis AA and Mavroedidis GP (2015) Kriging metamodeling in seismic risk assessment based on stochastic ground motion models. *Earthquake Engineering & Structural Dynamics* 44(14): 2377–2399.
- Goda K and Tesfamariam S (2015) Multi-variate seismic demand modelling using copulas: Application to non-ductile reinforced concrete frame in Victoria, Canada. *Structural Safety* 56: 39–51.
- Goh ATC (1994) Seismic liquefaction potential assessed by neural networks. *Journal of Geotechnical Engineering* 120(9): 1467–1480.
- Goh ATC (1996) Neural-network modeling of CPT seismic liquefaction data. *Journal of Geotechnical Engineering* 122(1): 70–73.
- Goh ATC and Goh SH (2007) Support vector machines: Their use in geotechnical engineering as illustrated using seismic liquefaction data. *Computers and Geotechnics* 34(5): 410–421.
- Goh SH and Zhang WG (2014) An improvement to MLR model for predicting liquefaction-induced lateral spread using multivariate adaptive regression splines. *Engineering Geology* 170: 1–10.
- Gong L, Wang C, Wu F, Zhang J, Zhang H and Li Q (2016) Earthquake-induced building damage detection with post-event sub-meter VHR TerraSAR-X staring spotlight imagery. *Remote Sensing* 8(11): 1–21.
- González MP and Zapico JL (2008) Seismic damage identification in buildings using neural networks and modal data. *Computers and Structures* 86(3–5): 416–426.
- Gülkan P and Kalkan E (2002) Attenuation modeling of recent earthquakes in Turkey. *Journal of Seismology* 6(3): 397–409.
- Güllü H (2012) Prediction of peak ground acceleration by genetic expression programming and regression: A comparison using likelihood-based measure. *Engineering Geology* 141: 92–113.
- Güllü H and Ergelebi E (2007) A neural network approach for attenuation relationships: An application using strong ground motion data from Turkey. *Engineering Geology* 93(3–4): 65–81.
- Hamze-Ziabari SM and Bakhshpoori T (2018) Improving the prediction of ground motion parameters based on an efficient bagging ensemble model of M5' and CART algorithms. *Applied Soft Computing Journal* 68: 147–161.
- Hanna AM, Ural D and Saygili G (2007) Neural network model for liquefaction potential in soil deposits using Turkey and Taiwan earthquake data. *Soil Dynamics and Earthquake Engineering* 27(6): 521–540.
- Hariri-Ardebili MA and Pourkamali-Anaraki F (2018) Simplified reliability analysis of multi hazard risk in gravity dams via machine learning techniques. *Archives of Civil and Mechanical Engineering* 18(2): 592–610.
- Ho TK (1995) Random decision forests. In: *Proceedings of 3rd international conference on document analysis and recognition*, Montreal, QC, Canada, 14–16 August, pp. 278–282. New York: IEEE.
- Hoang N-D and Bui DT (2018) Predicting earthquake-induced soil liquefaction based on a hybridization of kernel Fisher discriminant analysis and a least squares support vector machine: A multi-dataset study. *Bulletin of Engineering Geology and the Environment* 77(1): 191–204.
- Hosmer DW and Lemeshow S (2000) *Applied Logistic Regression*. 2nd ed. New York: John Wiley.
- Housner GW, Bergman LA, Caughey TK, Chassiakos AG, Claus RO, Masri SF, Skelton RE, Soong TT, Spencer BF and Yao JTP (1997) Structural control: Past, present, and future. *Journal of Engineering Mechanics* 123(9): 897–971.
- Huang CS, Hung SL, Wen CM and Tu TT (2003) A neural network approach for structural identification and diagnosis of a building from seismic response data. *Earthquake Engineering and Structural Dynamics* 32(2): 187–206.
- Huang G, Qiu W and Zhang J (2017) Modelling seismic fragility of a rock mountain tunnel based on support vector machine. *Soil Dynamics and Earthquake Engineering* 102: 160–171.

- Huang H and Burton HV (2019) Classification of in-plane failure modes for reinforced concrete frames with infills using machine learning. *Journal of Building Engineering* 25: 100767.
- Hung SL, Huang CS, Wen CM and Hsu YC (2003) Nonparametric identification of a building structure from experimental data using wavelet neural network. *Computer-aided Civil and Infrastructure Engineering* 18(5): 356–368.
- Hutchinson T, Franke K, Arduino P, Asimaki D, Barrett B, Dafni J, Dashti S, El Mohtar CS, Hastings N, Lemnitzer A, Lo E, Martinez A, Montalva G, Montgomery J and Wood C (2019) *GEER Puebla-Mexico City Earthquake*. DesignSafe-CI. DOI: 10.17603/ds2-aetm-8891.
- Idriss IM and Boulanger RW (2006) Semi-empirical procedures for evaluating liquefaction potential during earthquakes. *Soil Dynamics and Earthquake Engineering* 26: 115–130.
- Javadi AA, Rezaia M and Nezhad MM (2006) Evaluation of liquefaction induced lateral displacements using genetic programming. *Computers and Geotechnics* 33(4–5): 222–233.
- Javan-Emrooz H, Eskandari-Ghadi M and Mirzaei N (2018) Prediction equations for horizontal and vertical PGA, PGV, and PGD in northern Iran using prefix gene expression programming. *Bulletin of the Seismological Society of America* 108(4): 2305–2332.
- Jeng C-H and Mo YL (2004) Quick seismic response estimation of prestressed concrete bridges using artificial neural networks. *Journal of Computing in Civil Engineering* 18(4): 360–372.
- Jeon JS, Mangalathu S, Song J and Desroches R (2019) Parameterized seismic fragility curves for curved multi-frame concrete box-girder bridges using Bayesian parameter estimation. *Journal of Earthquake Engineering* 23: 954–979.
- Jeon J-S, Shafieezadeh A and DesRoches R (2014) Statistical models for shear strength of RC beam-column joints using machine-learning techniques. *Earthquake Engineering & Structural Dynamics* 43: 2075–2095.
- Jiang X and Adeli H (2005) Dynamic wavelet neural network for nonlinear identification of highrise buildings. *Computer-Aided Civil and Infrastructure Engineering* 20(5): 316–330.
- Jiang X and Adeli H (2007) Pseudospectra, MUSIC, and dynamic wavelet neural network for damage detection of highrise buildings. *International Journal for Numerical Methods in Engineering* 71: 606–629.
- Jiang X and Adeli H (2008a) Dynamic fuzzy wavelet neuroemulator for non-linear control of irregular building structures. *International Journal for Numerical Methods in Engineering* 74: 1045–1066.
- Jiang X and Adeli H (2008b) Neuro-genetic algorithm for non-linear active control of structures. *International Journal for Numerical Methods in Engineering* 75: 770–786.
- Jiang X and Mahadevan S (2008) Bayesian probabilistic inference for nonparametric damage detection of structures. *Journal of Engineering Mechanics* 134(10): 820–831.
- Jiang X, Mahadevan S and Yuan Y (2017) Fuzzy stochastic neural network model for structural system identification. *Mechanical Systems and Signal Processing* 82: 394–411.
- Joghataie A and Farrokh M (2008) Dynamic analysis of nonlinear frames by Prandtl neural networks. *Journal of Engineering Mechanics* 134(11): 961–969.
- Jolliffe IT (2002) *Principal Component Analysis (Springer Series in Statistics)*. 2nd ed. New York: Springer.
- Juang CH and Chen CJ (1999) CPT-based liquefaction evaluation using artificial neural networks. *Computer-Aided Civil and Infrastructure Engineering* 14(3): 221–229.
- Kameshwar S and Padgett JE (2014) Multi-hazard risk assessment of highway bridges subjected to earthquake and hurricane hazards. *Engineering Structures* 78: 154–166.
- Kameshwar S and Padgett JE (2018) Effect of vehicle bridge interaction on seismic response and fragility of bridges. *Earthquake Engineering and Structural Dynamics* 47(3): 697–713.
- Kameshwar S, Vishnu N and Padgett J (2019) *Response and Fragility Modeling of Aging Bridges Subjected to Earthquakes and Truck Loads*. DesignSafe-CI. DOI: 10.17603/ds2-5tzv-qz91
- Karimi I, Khaji N, Ahmadi MT and Mirzayee M (2010) System identification of concrete gravity dams using artificial neural networks based on a hybrid finite element-boundary element approach. *Engineering Structures* 32(11): 3583–3591.

- Karpatne A, Atluri G, Faghmous JH, Steinbach M, Banerjee A, Ganguly A, Shekhar S, Samatova NF and Kumar V (2017) Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Transactions on Knowledge and Data Engineering* 29(10): 2318–2331.
- Kaveh A, Bakhshpoori T and Hamze-Ziabari SM (2016) Derivation of new equations for prediction of principal ground-motion parameters using M5' algorithm. *Journal of Earthquake Engineering* 20(6): 910–930.
- Kaya Z (2016) Predicting liquefaction-induced lateral spreading by using neural network and neuro-fuzzy techniques. *International Journal of Geomechanics* 16(4): 04015095.
- Kerh T and Ting SB (2005) Neural network estimation of ground peak acceleration at stations along Taiwan high-speed rail system. *Engineering Applications of Artificial Intelligence* 18(7): 857–866.
- Khalatbarisoltani A, Soleymani M and Khodadadi M (2019) Online control of an active seismic system via reinforcement learning. *Structural Control and Health Monitoring* 26(3): 1–18.
- Khodabandolehlou H, Pekcan G, Fadali MS and Salem MMA (2018) Active neural predictive control of seismically isolated structures. *Structural Control and Health Monitoring* 25(1): 1–15.
- Khosravikia F, Clayton P and Nagy Z (2019) Artificial neural network-based framework for developing ground-motion models for natural and induced earthquakes in Oklahoma, Kansas, and Texas. *Seismological Research Letters* 90(2A): 604–613.
- Khosravikia F, Potter A, Prakhov V, Zalachoris G, Cheng T, Tiwari A, Clayton P, Cox B, Rathje E, Williamson E, Paine J and Frohlich C (2018) *Seismic Vulnerability and Post-event Actions for Texas Bridge Infrastructure*. Austin, TX: The University of Texas at Austin.
- Kia A and Sensoy S (2014) Classification of earthquake-induced damage for R/C slab column frames using multiclass SVM and its combination with MLP neural network. *Mathematical Problems in Engineering* 2014: 2014: 734072.
- Kicinger R, Arciszewski T and De Jong K (2005) Evolutionary computation and structural design: A survey of the state-of-the-art. *Computers and Structures* 83(23–24): 1943–1978.
- Kim DH and Lee IW (2001) Neuro-control of seismically excited steel structure through sensitivity evaluation scheme. *Earthquake Engineering and Structural Dynamics* 30(9): 1361–1377.
- Kim DH, Kim D, Chang S and Jung HY (2008) Active control strategy of structures based on lattice type probabilistic neural network. *Probabilistic Engineering Mechanics* 23(1): 45–50.
- Kim J-T, Jung H-J and Lee I-W (2000) Optimal structural control using neural networks. *Journal of Engineering Mechanics* 126(2): 201–205.
- Kohestani VR and Ardakani MH (2015) Evaluation of liquefaction potential based on CPT data using random forest. *Natural Hazards* 79(2): 1079–1089.
- Kong Q, Trugman DT, Ross ZE, Bianco MJ, Meade BJ and Gerstoft P (2018) Machine learning in seismology: Turning data into insights. *Seismological Research Letters* 90(1): 3–14.
- Kong Z, Tang B, Deng L, Liu W and Han Y (2020) Condition monitoring of wind turbines based on spatio-temporal fusion of SCADA data by convolutional neural networks and gated recurrent units. *Renewable Energy* 146: 760–768.
- Koutsourelakis PS (2010) Assessing structural vulnerability against earthquakes using multi-dimensional fragility surfaces: A Bayesian framework. *Probabilistic Engineering Mechanics* 25(1): 49–60.
- Krishnamoorthy A, Bhat S and Bhasari D (2017) Radial basis function neural network algorithm for semi-active control of base-isolated structures. *Structural Control and Health Monitoring* 24(10): 1–11.
- Lagaros ND and Fragiadakis M (2007) Fragility assessment of steel frames using neural networks. *Earthquake Spectra* 23(4): 735–752.
- Lagaros ND, Tsompanakis Y, Psarropoulos PN and Georgopoulos EC (2009) Computationally efficient seismic fragility analysis of geostructures. *Computers and Structures* 87(19–20): 1195–1203.
- Lattanzi D, Miller GR, Eberhard MO and Haraldsson OS (2015) Bridge column maximum drift estimation via computer vision. *Journal of Computing in Civil Engineering* 30(4): 04015051.
- LeCun Y, Bengio Y and Hinton G (2015) Deep learning. *Nature* 521: 436–444.

- Lee HJ, Jung HJ, Cho SW and Lee IW (2008) An experimental study of semiactive modal neuro-control scheme using MR damper for building structure. *Journal of Intelligent Material Systems and Structures* 19(9): 1005–1015.
- Lee HJ, Yang G, Jung HJ, Spencer BF and Lee IW (2006) Semi-active neurocontrol of a base-isolated benchmark structure. *Structural Control and Health Monitoring* 13(2–3): 682–692.
- Li HN and Yang H (2007) System identification of dynamic structure by the multi-branch BPNN. *Neurocomputing* 70(4–6): 835–841.
- Liao TW, Egbelu PJ, Sarker BR and Leu SS (2011) Metaheuristics for project and construction management—A state-of-the-art review. *Automation in Construction* 20(5): 491–505.
- Liel AB, Haselton CB, Deierlein GG and Baker JW (2009) Incorporating modeling uncertainties in the assessment of seismic collapse risk of buildings. *Structural Safety* 31(2): 197–211.
- Liu Z and Tesfamariam S (2012) Prediction of lateral spread displacement: Data-driven approaches. *Bulletin of Earthquake Engineering* 10(5): 1431–1454.
- Liu Z and Zhang Z (2018) Artificial neural network based method for seismic fragility analysis of steel frames. *KSCE Journal of Civil Engineering* 22(2): 708–717.
- Lu P, Chen S and Zheng Y (2012) Artificial intelligence in civil engineering. *Mathematical Problems in Engineering* 2012: 145974.
- Luco N, Manuel L, Baldava S and Bazzurro P (2005) Correlation of damage of steel moment-resisting frames to a vector-valued set of ground motion parameters. In: *Proceedings of the 9th international conference on structural safety and reliability (ICOSSAR05)*, Rome, 19–23 June, pp. 2719–2726. Rotterdam: Millpress.
- Luo H and Paal SG (2018) Machine learning-based backbone curve model of reinforced concrete columns subjected to cyclic loading reversals. *Journal of Computing in Civil Engineering* 32(5): 04018042.
- Luo H and Paal SG (2019) A locally weighted machine learning model for generalized prediction of drift capacity in seismic vulnerability assessments. *Computer-Aided Civil and Infrastructure Engineering* 34: 935–950.
- Madan A (2005) Vibration control of building structures using self-organizing and self-learning neural networks. *Journal of Sound and Vibration* 287(4–5): 759–784.
- Mahmoudi SN and Chouinard L (2016) Seismic fragility assessment of highway bridges using support vector machines. *Bulletin of Earthquake Engineering* 14(6): 1571–1587.
- Mangalathu S and Burton HV (2019) Deep learning-based classification of earthquake-impacted buildings using textual damage descriptions. *International Journal of Disaster Risk Reduction* 36: 101111.
- Mangalathu S and Jeon JS (2018) Classification of failure mode and prediction of shear strength for reinforced concrete beam-column joints using machine learning techniques. *Engineering Structures* 160: 85–94.
- Mangalathu S and Jeon J-S (2019a) Machine learning-based failure mode recognition of circular reinforced concrete bridge columns: Comparative study. *Journal of Structural Engineering* 145(10): 04019104.
- Mangalathu S and Jeon J-S (2019b) Stripe-based fragility analysis of concrete bridge classes using machine learning techniques. *Earthquake Engineering & Structural Dynamics* 48: 1238–1255.
- Mangalathu S, Heo G and Jeon JS (2018a) Artificial neural network based multi-dimensional fragility development of skewed concrete bridge classes. *Engineering Structures* 162: 166–176.
- Mangalathu S, Jeon JS and DesRoches R (2018b) Critical uncertainty parameters influencing seismic performance of bridges using Lasso regression. *Earthquake Engineering and Structural Dynamics* 47(3): 784–801.
- Markič Š and Stankovski V (2013) An equation-discovery approach to earthquake-ground-motion prediction. *Engineering Applications of Artificial Intelligence* 26(4): 1339–1347.
- Memarzadeh M and Pozzi M (2019) Model-free reinforcement learning with model-based safe exploration: Optimizing adaptive recovery process of infrastructure systems. *Structural Safety* 80: 46–55.

- Mitra N, Mitra S and Lowes LN (2011) Probabilistic model for failure initiation of reinforced concrete interior beam-column connections subjected to seismic loading. *Engineering Structures* 33(1): 154–162.
- Mitropoulou CC and Papadrakakis M (2011) Developing fragility curves based on neural network IDA predictions. *Engineering Structures* 33(12): 3409–3421.
- Mohammadnejad AK, Mousavi SM, Torabi M, Mousavi M and Alavi AH (2012) Robust attenuation relations for peak time-domain parameters of strong ground motions. *Environmental Earth Sciences* 67(1): 53–70.
- Molas GL and Yamazaki F (1995) Neural networks for quick earthquake damage estimation. *Earthquake Engineering & Structural Dynamics* 24(4): 505–516.
- Morfidis K and Kostinakis K (2017) Seismic parameters' combinations for the optimum prediction of the damage state of R/C buildings using neural networks. *Advances in Engineering Software* 106: 1–16.
- Morfidis K and Kostinakis K (2018) Approaches to the rapid seismic damage prediction of r/c buildings using artificial neural networks. *Engineering Structures* 165: 120–141.
- Mosalam K, Abuchar V, Archbold J, Arteta C, Fischer E, Gunay S, Hakhmaneshi M, Hassan W, Micheli L, Muin S, Pajaro Miranda C, Pretell Ductram AR, Peng H, Robertson I, Roueche D and Ziotopoulou K (2019) *Steer—M6.4 and M7.1 Ridgecrest, CA Earthquakes on July 4-5, 2019: Preliminary Virtual Reconnaissance Report (PVRR)*. DesignSafe-CI. DOI: 10.17603/ds2-xqfh-1631.
- Moss RE, Seed RB, Kayen RE, Stewart JP, Der Kiureghian A and Cetin KO (2006) CPT-based probabilistic and deterministic assessment of in situ seismic soil liquefaction potential. *Journal of Geotechnical and Geoenvironmental Engineering* 132(8): 1032–1051.
- Murphy KP (2012) *Machine Learning: A Probabilistic Perspective*. *Chance Encounters: Probability in Education*. Cambridge, MA; London: The MIT Press.
- Naderpour H and Mirrashid M (2019) Shear failure capacity prediction of concrete beam-column joints in terms of ANFIS and GMDH. *Practice Periodical on Structural Design and Construction* 24(2): 04019006.
- Nakamura M, Masri SF, Chassiakos AG and CauGhey TK (1998) A method for non-parametric damage detection through the use of neural networks. *Earthquake Engineering and Structural Dynamics* 27: 997–1010.
- Nguyen KTP and Medjaher K (2019) A new dynamic predictive maintenance framework using deep learning for failure prognostics. *Reliability Engineering and System Safety* 188: 251–262.
- Oommen T and Baise LG (2010) Model development and validation for intelligent data collection for lateral spread displacements. *Journal of Computing in Civil Engineering* 24(6): 467–477.
- Paal SG, Jeon J-S, Brilakis I and DesRoches R (2014) Automated damage index estimation of reinforced concrete columns for post-earthquake evaluations. *Journal of Structural Engineering* 141(9): 04014228.
- Pal M (2006) Support vector machines-based modelling of seismic liquefaction potential. *International Journal for Numerical and Analytical Methods in Geomechanics* 30(10): 983–996.
- Pan Y, Agrawal AK and Ghosn M (2007) Seismic fragility of continuous steel highway bridges in New York State. *Journal of Bridge Engineering* 12(6): 689–699.
- Pang Y, Dang X and Yuan W (2014) An artificial neural network based method for seismic fragility analysis of highway bridges. *Advances in Structural Engineering* 17(3): 413–428.
- Park J and Towashiraporn P (2014) Rapid seismic damage assessment of railway bridges using the response-surface statistical model. *Structural Safety* 47: 1–12.
- Perlovsky L (2000) *Neural Networks and Intellect: Using Model-Based Concepts*. 1st ed. Oxford: Oxford University Press.
- Perotti F, Domaneschi M and De Grandis S (2013) The numerical computation of seismic fragility of base-isolated nuclear power plants buildings. *Nuclear Engineering and Design* 262: 189–200.
- Peyk-Herfeh M and Shahbahrani A (2014) Evaluation of adaptive boosting and neural network in earthquake damage levels detection. *International Journal of Computer Applications* 100(3): 23–29.
- Pirhadi N, Tang X, Yang Q and Kang F (2018) A new equation to evaluate liquefaction triggering using the response surface method and parametric sensitivity analysis. *Sustainability* 11(1): 1–24.

- Porter KA (2003) An overview of PEER's performance-based earthquake engineering methodology. In: *Ninth international conference on applications of statistics and probability in civil engineering*, San Francisco, CA, 6–9 July, pp. 1–8.
- Raissi M and Karniadakis GE (2018) Hidden physics models: Machine learning of nonlinear partial differential equations. *Journal of Computational Physics* 357: 125–141.
- Rajeev P and Tesfamariam S (2012) Seismic fragilities for reinforced concrete buildings with consideration of irregularities. *Structural Safety* 39: 1–13.
- Ramkrishnan D, Singh TN, Purwar N, Barde KS, Gulati A and Gupta S (2008) Artificial neural network and liquefaction susceptibility assessment: A case study using the 2001 Bhuj earthquake data, Gujarat, India. *Computational Geosciences* 12(4): 491–501.
- Rao MM and Datta TK (2006) Modal seismic control of building frames by artificial neural network. *Journal of Computing in Civil Engineering* 20(1): 69–73.
- Rathje EM, Dawson C, Padgett JE, Pinelli J-P, Stanzione D, Adair A, Arduino P, Brandenberg SJ, Cockerill T, Dey C, Esteva M, Haan FL Jr, Hanlon M, Kareem A, Lowes L, Mock S and Mosqueda G (2017) DesignSafe: New cyberinfrastructure for natural hazards engineering. *Natural Hazards Review* 18(3): 06017001.
- Ravi Kiran A, Reddy GR and Agrawal MK (2019) Seismic fragility analysis of pressurized piping systems considering ratcheting: A case study. *International Journal of Pressure Vessels and Piping* 169(412): 26–36.
- Rawlings JO, Pantula SG and Dickey DA (1998) *Applied Regression Analysis: A Research Tool* (Springer Texts in Statistics). 2nd ed. New York: Springer.
- Rezania M, Faramarzi A and Javadi AA (2011) An evolutionary based approach for assessment of earthquake-induced soil liquefaction and lateral displacement. *Engineering Applications of Artificial Intelligence* 24(1): 142–153.
- Rezania M, Javadi AA and Giustolisi O (2010) Evaluation of liquefaction potential based on CPT results using evolutionary polynomial regression. *Computers and Geotechnics* 37(1–2): 82–92.
- Ricci P, De Risi MT, Verderame GM and Manfredi G (2013) Influence of infill distribution and design typology on seismic performance of low- and mid-rise RC buildings. *Bulletin of Earthquake Engineering* 11(5): 1585–1616.
- Saha SK, Matsagar V and Chakraborty S (2016) Uncertainty quantification and seismic fragility of base-isolated liquid storage tanks using response surface models. *Probabilistic Engineering Mechanics* 43: 20–35.
- Sahoo DM and Chakraverty S (2018) Functional link neural network learning for response prediction of tall shear buildings with respect to earthquake data. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 48(1): 1–10.
- Sahu D, Nishanth M, Dhir PK, Sarkar P, Davis R and Mangalathu S (2019) Stochastic response of reinforced concrete buildings using high dimensional model representation. *Engineering Structures* 179: 412–422.
- Saka MP and Geem ZW (2013) Mathematical and metaheuristic applications in design optimization of steel frame structures: An extensive review. *Mathematical Problems in Engineering* 2013: 271031.
- Salehi H and Burgueño R (2018) Emerging artificial intelligence methods in structural engineering. *Engineering Structures* 171: 170–189.
- Samuel AL (1959) Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development* 3(3): 535–554.
- Schöbi R, Sudret B and Marelli S (2017) Rare event estimation using polynomial-chaos kriging. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* 3(2): 1–12.
- Sen D, Aghazadeh A, Mousavi A, Nagarajaiah S, Baraniuk R and Dabak A (2019) Data-driven semi-supervised and supervised learning algorithms for health monitoring of pipes. *Mechanical Systems and Signal Processing* 131: 524–537.
- Seo J and Linzell DG (2010) Probabilistic vulnerability scenarios for horizontally curved steel I-girder bridges under earthquake loads. *Transportation Research Record: Journal of the Transportation Research Board* 2202: 206–211.

- Seo J and Linzell DG (2012) Horizontally curved steel bridge seismic vulnerability assessment. *Engineering Structures* 34: 21–32.
- Seo J and Linzell DG (2013) Use of response surface metamodels to generate system level fragilities for existing curved steel bridges. *Engineering Structures* 52: 642–653.
- Seo J and Park H (2017) Probabilistic seismic restoration cost estimation for transportation infrastructure portfolios with an emphasis on curved steel I-girder bridges. *Structural Safety* 65: 27–34.
- Seo J, Dueñas-Orsorio L, Craig JI and Goodno BJ (2012) Metamodel-based regional vulnerability estimate of irregular steel moment-frame structures subjected to earthquake events. *Engineering Structures* 45: 585–597.
- Shahin MA (2014) A review of artificial intelligence applications in shallow foundations. *International Journal of Geotechnical Engineering* 9(1): 49–60.
- Shahin MA (2016) State-of-the-art review of some artificial intelligence applications in pile foundations. *Geoscience Frontiers* 7(1): 33–44.
- Sichani ME (2018) *Seismic risk assessment of vertical concrete dry casks*. PhD Dissertation, Rice University, Houston, TX.
- Silva MS and Garcia L (2001) Earthquake damage assessment based on fuzzy logic and neural networks. *Earthquake Spectra* 17(1): 89–112.
- Stewart JP, Kramer SL, Kwak DY, Greenfield MW, Kayen RE, Tokimatsu K, Bray JD, Beyzaei CZ, Cubrinovski M, Sekiguchi T, Nakai S and Bozorgnia Y (2016) PEER-NGL project: Open source global database and model development for the next-generation of liquefaction assessment procedures. *Soil Dynamics and Earthquake Engineering* 91: 317–328.
- Subasri R, Suresh S and Natarajan AM (2014) Discrete direct adaptive ELM controller for active vibration control of nonlinear base isolation buildings. *Neurocomputing* 129: 246–256.
- Suresh S, Narasimhan S, Nagarajaiah S and Sundararajan N (2010) Fault-tolerant adaptive control of nonlinear base-isolated buildings using EMRAN. *Engineering Structures* 32(8): 2477–2487.
- Tezcan J and Cheng Q (2012) Support vector regression for estimating earthquake response spectra. *Bulletin of Earthquake Engineering* 10(4): 1205–1219.
- Thomas S, Pillai GN and Pal K (2017) Prediction of peak ground acceleration using ϵ -SVR, ν -SVR and Ls-SVR algorithm. *Geomatics, Natural Hazards and Risk* 8(2): 177–193.
- Thomas S, Pillai GN, Pal K and Jagtap P (2016) Prediction of ground motion parameters using randomized ANFIS (RANFIS). *Applied Soft Computing Journal* 40: 624–634.
- Tibshirani R (1996) Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* 58(1): 267–288.
- Trugman DT and Shearer PM (2018) Strong correlation between stress drop and peak ground acceleration for recent M 1–4 earthquakes in the San Francisco bay area. *Bulletin of the Seismological Society of America* 108(2): 929–945.
- Tsompanakis Y, Lagaros ND, Psarropoulos PN and Georgopoulos EC (2009) Simulating the seismic response of embankments via artificial neural networks. *Advances in Engineering Software* 40(8): 640–651.
- Ülgen D and Engin HK (2007) A study of CPT based liquefaction assessment using artificial neural networks. In: *4th international conference on earthquake geotechnical engineering*, Thessaloniki, 25–28 June, paper no. 1219, pp. 1–12. New York: Springer.
- Vafaei M, Adnan AB and Rahman ABA (2013) Real-time seismic damage detection of concrete shear walls using artificial neural networks. *Journal of Earthquake Engineering* 17(1): 137–154.
- Vafaei M, Adnan AB and Rahman ABA (2014) A neuro-wavelet technique for seismic damage identification of cantilever structures. *Structure and Infrastructure Engineering* 10(12): 1666–1684.
- Vapnik V (1998) *Statistical Learning Theory*. 1st ed. New York: John Wiley & Sons.
- Verderame GM, Ricci P, De Luca F, Del Gaudio C and De Risi MT (2014) Damage scenarios for RC buildings during the 2012 Emilia (Italy) earthquake. *Soil Dynamics and Earthquake Engineering* 66: 385–400.
- Wagner N and Rondinelli JM (2016) Theory-guided machine learning in materials science. *Frontiers in Materials* 3: 1–9.

- Wallen RD (2004) *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance. Biomedical Instrumentation and Technology*, vol. 38. 2nd ed. Boca Raton, FL; London; New York: CRC Press.
- Wang J and Rahman MS (1999) A neural network model for liquefaction-induced horizontal ground displacement. *Soil Dynamics and Earthquake Engineering* 18(8): 555–568.
- Wang Z, Pedroni N, Zentner I and Zio E (2018) Seismic fragility analysis with artificial neural networks: Application to nuclear power plant equipment. *Engineering Structures* 162: 213–225.
- Wen YK, Ghaboussi J, Venini P and Nikzad K (1995) Control of structures using neural networks. *Smart Materials and Structures* 4(1A): 149–157.
- Whitley D, Starkweather T and Bogart C (1990) Genetic algorithms and neural networks: Optimizing connections and connectivity. *Parallel Computing* 14(3): 347–361.
- Wu Z, Xu B and Yokoyama K (2002) Decentralized parametric damage detection based on neural networks. *Computer-Aided Civil and Infrastructure Engineering* 17(3): 175–184.
- Xia PQ (2003) An inverse model of MR damper using optimal neural network and system identification. *Journal of Sound and Vibration* 266(5): 1009–1023.
- Xie Y and DesRoches R (2019) Sensitivity of seismic demands and fragility estimates of a typical California highway bridge to uncertainties in its soil-structure interaction modeling. *Engineering Structures* 189: 605–617.
- Xie Y and Zhang J (2016) Optimal design of seismic protective devices for highway bridges using performance-based methodology and multiobjective genetic optimization. *Journal of Bridge Engineering* 22(3): 04016129.
- Xie Y and Zhang J (2018) Design and optimization of seismic isolation and damping devices for highway bridges based on probabilistic repair cost ratio. *Journal of Structural Engineering* 144(8): 04018125.
- Xie Y, Zhang J, DesRoches R and Padgett JE (2019a) Seismic fragilities of single-column highway bridges with rocking column-footing. *Earthquake Engineering & Structural Dynamics* 48: 843–864.
- Xie Y, Zheng Q, Yang CSW, Zhang W, DesRoches R, Padgett JE and Taciroglu E (2019b) Probabilistic models of abutment backfills for regional seismic assessment of highway bridges in California. *Engineering Structures* 180: 452–467.
- Xu B, Wu Z, Chen G and Yokoyama K (2004) Direct identification of structural parameters from dynamic responses with neural networks. *Engineering Applications of Artificial Intelligence* 17(8): 931–943.
- Xu B, Wu Z, Yokoyama K, Harada T and Chen G (2005) A soft post-earthquake damage identification methodology using vibration time series. *Smart Materials and Structures* 14(3): 116–124.
- Xu ZD and Guo YQ (2008) Neuro-fuzzy control strategy for earthquake-excited nonlinear magnetorheological structures. *Soil Dynamics and Earthquake Engineering* 28(9): 717–727.
- Xu ZD, Shen YP and Guo YQ (2003) Semi-active control of structures incorporated with magnetorheological dampers using neural networks. *Smart Materials and Structures* 12(1): 80–87.
- Yakut O and Alli H (2011) Neural based sliding-mode control with moving sliding surface for the seismic isolation of structures. *Journal of Vibration and Control* 17(14): 2103–2116.
- Yang Y, Dorn C, Mancini T, Talken Z, Nagarajaiah S, Kenyon G, Farrar C and Mascareñas D (2017) Blind identification of full-field vibration modes of output-only structures from uniformly-sampled, possibly temporally-aliased (sub-Nyquist), video measurements. *Journal of Sound and Vibration* 390: 232–256.
- Yaseen ZM, Afan HA and Tran MT (2018) Beam-column joint shear prediction using hybridized deep learning neural network with genetic algorithm. *IOP Conference Series: Earth and Environmental Science* 143(1): 1–7.
- Yazdi AJ, Haukaas T, Yang T and Gardoni P (2016) Multivariate fragility models for earthquake engineering. *Earthquake Spectra* 32(1): 441–461.
- Yerlikaya-Ozkurt F, Askan A and Weber G-W (2014) An alternative approach to the ground motion prediction problem by a non-parametric adaptive regression method. *Engineering Optimization* 46(12): 1651–1668.

- Youd TL (2018) Application of MLR procedure for prediction of liquefaction-induced lateral spread displacement. *Journal of Geotechnical and Geoenvironmental Engineering* 144(6): 04018033.
- Youd TL, Hansen CM and Bartlett SF (2002) Revised multilinear regression equations for prediction of lateral spread displacement. *Journal of Geotechnical and Geoenvironmental Engineering* 128(12): 1007–1017.
- Yun HB, Tasbighoo F, Masri SF, Caffrey JP, Wolfe RW, Makris N and Black C (2008) Comparison of modeling approaches for full-scale nonlinear viscous dampers. *JVC/Journal of Vibration and Control* 14(1–2): 51–76.
- Zhang J, Sato T, Iai S and Hutchinson T (2008) A pattern recognition technique for structural identification using observed vibration signals: Linear case studies. *Engineering Structures* 30(5): 1439–1446.
- Zhang J, Xie Y and Wu G (2019a) Seismic responses of bridges with rocking column-foundation: A dimensionless regression analysis. *Earthquake Engineering and Structural Dynamics* 48(1): 152–170.
- Zhang J, Zhang LM and Huang HW (2013) Evaluation of generalized linear models for soil liquefaction probability prediction. *Environmental Earth Sciences* 68(7): 1925–1933.
- Zhang R, Chen Z, Chen S, Zheng J, Büyüköztürk O and Sun H (2019b) Deep long short-term memory networks for nonlinear structural seismic response prediction. *Computers and Structures* 220: 55–68.
- Zhang Y and Burton HV (2019) Pattern recognition approach to assess the residual structural capacity of damaged tall buildings. *Structural Safety* 78: 12–22.
- Zhang Y, Burton HV, Sun H and Shokrabadi M (2018) A machine learning framework for assessing post-earthquake structural safety. *Structural Safety* 72: 1–16.
- Zhou J, Li E, Wang M, Chen X, Shi X and Jiang L (2019) Feasibility of stochastic gradient boosting approach for evaluating seismic liquefaction potential based on SPT and CPT case histories. *Journal of Performance of Constructed Facilities* 33(3): 04019024.
- Zou H and Hastie T (2005) Regression and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67(2): 301–320.