HUMAN-MACHINE COLLABORATION FRAMEWORK FOR BRIDGE HEALTH MONITORING

Sifat Muin[*], Chrystal Chern[†], and Khalid M. Mosalam[‡]

University of California, Berkeley

Abstract

The importance and relevance of structural health monitoring (SHM) for highway bridges in the United States is highlighted by the bridges' poor condition and the growing amount of resources for data-driven condition assessment in the field of artificial intelligence (AI), particularly, machine learning (ML). To tackle this issue, a human-machine collaboration (H-MC) framework for highway bridge SHM is developed in this study to take advantage of the strengths of both AI and engineering domain expertise. The H-MC framework uses a *physics-based* model (human) to conduct probability of exceedance (POE) analysis coupled with *novelty detection* ML model (machine) to establish a damage detection and assessment algorithm. To produce the training data for the model, nonlinear time history analyses (NTHA) are performed on analytical bridge model for vibration responses to many selected ground motions. Feature extraction and selection are performed using an *Ordinal Fisher Score* analysis with *k-fold cross-validation* parameter tuning to produce the input for training the ML model. A component capacity-based damage state model is used to produce the output for training the ML model. During model validation, the benefit of the H-MC is demonstrated through significant increase in the classification accuracy when the POE analysis is coupled with the ML model. Once the ML model is trained, it is tested on the seismic responses of two instrumented bridges, El Centro - Hwy8/Meloland Overpass and Parkfield - Hwy46/Cholame Creek Bridge. The model accurately classified all 14 undamaged events and one damaged event, including damage assessment consistent with reported visual inspection of the bridge after the damaging event.

Introduction

Structural health monitoring (SHM) of highway bridges is becoming essential as the civil infrastructures in the United States are aging. According to ASCE's 2019 report card for California's infrastructures, 7% of California's bridges are structurally deficient and the largest percentage of bridges in "Poor" condition in the nation belongs to California. Moreover, approximately 50% of bridges in the state have exceeded their design life. The backlog of recommended maintenance, repair, and replacement continues to grow. Therefore, continued health monitoring of these bridges is essential for safe operation and for prioritizing replacement. The monitoring of these vulnerable structures becomes more critical after natural disasters such as earthquakes. The loss of accessibility due to damages and closures of the transportation network can greatly affect the rescue and recovery of a city after major seismic

[*]Post-Doctoral Researcher, sifat.muin@berkeley.edu.

[†]PhD Student, cchern@berkeley.edu.

[‡]Corresponding author, Taisei Professor of Civil Engineering & PEER Director, mosalam@berkeley.edu.

events, turning them into disasters. Utilizing the advances in remote sensing, computing technologies, and data science is an effective and feasible way of structural monitoring of bridges.

Data-driven SHM brings in the capability of machine learning (ML) and deep learning (DL) techniques in developing automated and online damage detection and/or assessment tools for civil infrastructures. With the recent advances in sensor technologies, wireless sensor networks, remote sensing, and computer vision, data science can now be used to develop SHM techniques to assess and quantify the conditions of structures in near-real-time utilizing ML and DL algorithms. ML gives computers the ability to learn about the trends and features of a process without being explicitly programmed (Samuel, 1959) for this process. Currently, ML is the core of artificial intelligence (AI) which has excelled in classification and prediction capabilities. However, ML still lacks contextual intelligence and minimal reasoning capability. The main problem with this approach to AI is that it requires thousands, if not millions, of data examples for training the ML algorithm to be able to function properly. Therefore, ML cannot be adopted in the current form for applications where data is limited. One such application is SHM of bridges to detect damage conditions and categories after seismic events.

In order to enable widespread adoption, AI systems need to shift away from this data-heavy approach and focus on methods that use a combination of different approaches, including systems modelling based on domain expertise. Through such collaborative approaches, humans and AI will actively enhance each other's complementary strengths. However, to fully extract the benefit of this collaboration, humans need to understand and trust the AI systems as they become more involved in making critical decisions alongside the humans. Transparency of the explainable AI (XAI) systems can foster this trust. One way to gain explainability in AI systems is to use ML algorithms that are inherently explainable. In this paper, such a ML tool and structural engineering expertise are utilized to develop the human-machine collaboration (H-MC) framework for highway bridge health monitoring.

Proposed Framework

Human-machine collaboration (H-MC) is a model in which humans co-work with artificial intelligence to complete specific tasks (National Research Council, 2012). The goal of an H-MC model is to utilize the forte of both types of intelligence and minimize the impact of their weaknesses. In this study, an algorithm is developed by following the basic principal of the H-MC model which is employed towards developing an SHM framework. The proposed algorithm leverages the advancements in the field of ML and utilizes response data from undamaged bridge structures (more abundantly available than from damaged bridges) to develop a ML model for a specific structure. Concurrently, an analytical model is developed by a domain expert, i.e., the structural engineer in this case, in order to simulate data of the damaged structure. This approach reduces the uncertainty associated with the computational model by using real data recorded from the structure and at the same time it provides a possible range of data (possibly unseen) from the damaged condition which is not usually available from in-situ measurements of structures. This is because engineered structures are typically damaged due to rare (low occurrence probability) events of natural hazards with high consequences, e.g., earthquakes or hurricanes.

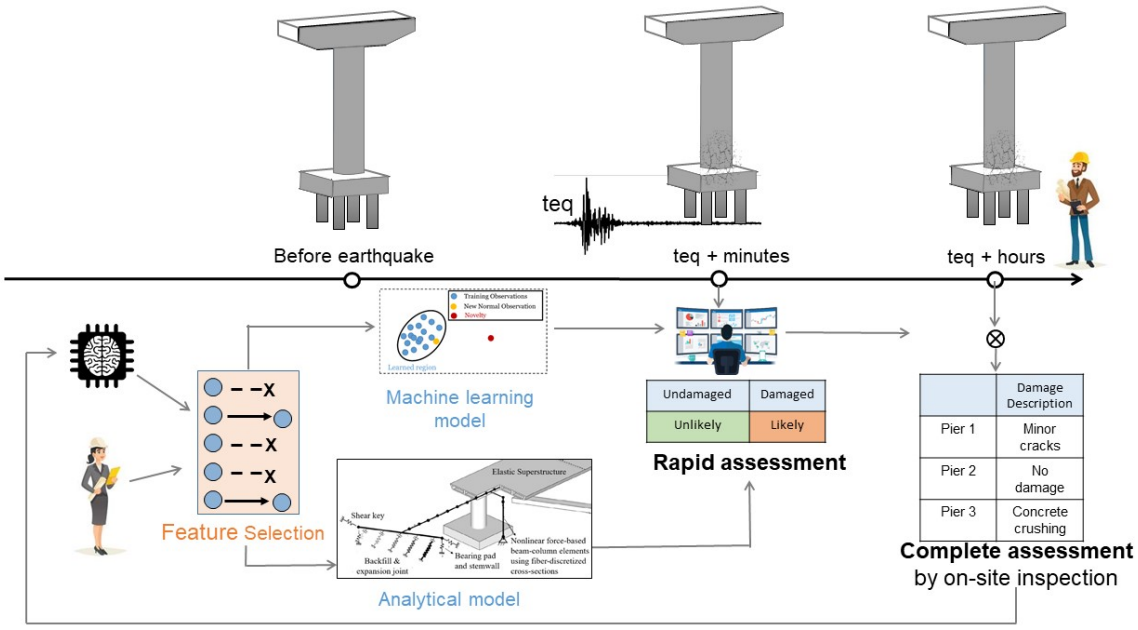


Figure 1: H-MC framework for SHM making use of ML and knowledge of human experts.

Figure 1 illustrates the H-MC framework for bridge SHM. There are two distinct phases. The first phase of the framework is completed before an earthquake occurs. In this phase, the task of feature selection is conducted, and both the analytical model and the ML model are developed[1]. The second phase of the framework takes place during and shortly following an earthquake, where the response of an instrumented bridge is measured and relayed to a remote location where damage sensitive features can be rapidly generated. Consequently, using the previously developed H-MC model, damage detection can be performed in *real time* within minutes after an earthquake. Detailed analytical models along with ML tools can provide further information on damage levels and locations within minutes. However, if a detailed model is missing, the measured response data can be reviewed by a skilled human counterpart to provide a *near real-time* assessment within hours after an event. This can be followed up by a complete on-site investigation by a professional structural engineer as part of a field inspection team if it is deemed necessary following the assessment phase. This complete assessment including field inspection may take days after the event. Following the complete assessment process, the ML model can be updated using the annotated response data including the damage information. In the subsequent sections, this proposed H-MC framework is described in detail including several applications.

ML Model

The proposed framework uses the *novelty detection* method as the ML model. Novelty detection has gained much research attention in application domains of critical systems particularly with those involving large datasets. Conventional ML typically focuses on the classifica-

[1]This dual H-MC model is sometimes referred to in modern literature as a *digital twin*.

tion of two or more classes. The problem of novelty detection, however, is approached within the framework of one-class classification. It can be defined as the task of recognizing unseen data that differs in some respect from the training data.

In the novelty detection approach to classification, "normal" patterns are available for training, while "abnormal" ones are missing. A model of normality M(θ), where θ represents the free parameters of the model, is inferred and used to assign novelty scores $z(x)$ to previously unseen test data x. Larger $z(x)$ corresponds to increased "abnormality" with respect to the model of normality. A threshold $z(x) = k$ is defined to identify novelty as follows,

$$\text{If } z(x) \geq k \rightarrow x \text{ is novelty.} \tag{1}$$

Due to the practical importance and challenging nature of novelty detection, many approaches have been proposed to determine the model of normality M(θ) and the threshold k. In this study, a distance-based approach has been conceived. The responses from undamaged structures are used to develop the M(θ). Multiple features, selected following a feature engineering process (described in a later section of this paper), are utilized. The distance of a point to the centroid of M(θ) is computed using the Mahalanobis distance (Mahalanobis, 1936). Since M(θ) is a multivariate model, not only distance but also the shape and size of the data need to be considered, quantified by the covariance matrix. The Mahalanobis distance[2] takes this covariance matrix into account. For a p-dimensional multivariate sample $\mathbf{x_i} = \{x_{i1}, \ldots, x_{ip}\}$, the Mahalanobis distance is defined as follows,

$$MD_i = \sqrt{(\mathbf{x_i} - \boldsymbol{\mu})^T \mathbf{C}^{-1} (\mathbf{x_i} - \boldsymbol{\mu})} \tag{2}$$

where $\boldsymbol{\mu}$ is the p-dimensional arithmetic mean vector consisting of the mean values of all the variables, and $\mathbf{C}$ is the covariance matrix. For multivariate normally distributed data, the values of the square of MD_i are approximately chi-square distributed with p degrees of freedom, χ^2_p. Conventionally, a certain quantile q (e.g., 0.95) of the χ^2_p distribution is set as the threshold k. In this study, q is considered as a parameter of the novelty model and estimated using the k-fold cross-validation technique.

POE Envelope

The human aspect of the framework is manifested in the probability of exceedance (POE) envelope where structural engineering domain expertise is applied. For this part of the analysis, a structure-specific analytical model based on a single degree of freedom (SDOF) idealization, is developed. This idealization is selected herein for simplicity but more sophisticated models can be equally utilized. The SDOF model properties are based on the structure under consideration and discussed in the "Numerical Example" Section. For this numerical example, the SDOF properties are listed in Table 1. The SDOF model has a nonlinear (inelastic) material that follows Özdemir's rate-independent force-displacement model (Özdemir, 1976).

[2]It is a measure of the distance between a point $\mathbf{x}$ and a distribution $\mathscr{D}$ as a multi-dimensional generalization of measuring how many standard deviations away $\mathbf{x}$ is from the mean of $\mathscr{D}$.

Table 1: Material and dynamic properties of the SDOF model representing Bridge-A.

Yield stress	68 ksi
Yield displacement	0.67 in
Hardening ratio	0.01
Period	0.6 sec
Damping	5%

Nonlinear time history analysis (NTHA) is performed on the model using a set of ground motions from the PEER[3] NGA[4]-West2 database (PEER, 2020; Ancheta et al., 2014). At the time of writing this paper, this database has 21,539 records of shallow crustal earthquakes in active tectonic regions. Due to the possibility of anomalies from older data collection systems, only records from the past 30 years have been selected. Moreover, records with peak ground acceleration (PGA) less than 1% of the acceleration of gravity (g) may not produce enough excitation useful for this study. Therefore, only records with PGA more than 0.01g are considered. To avoid homogeneity in response, not more than 20 records from a single earthquake event are selected. Thus, $n = 1,710$ records matched these criteria. From the SDOF analysis, a damage feature vector $\mathbf{X} = \{X_1, X_2, \ldots, X_m\}$ is computed from the responses. Subsequently, a probability of exceedance (POE) envelope is developed as follows,

$$\begin{aligned}\hat{G}_i = P(DS \geq 1|\mathbf{X_i}) &= \frac{P(DS \geq 1, \mathbf{X} = \mathbf{X_i})}{P(\mathbf{X} = \mathbf{X_i})} \\ &= \frac{\frac{1}{n}\sum_{j=1}^{n}[I(DS)K_H(\mathbf{X_j} - \mathbf{X_i})]}{\frac{1}{n}\sum_{j=1}^{n}[K_H(\mathbf{X_j} - \mathbf{X_i})]} = \frac{\sum_{j=1}^{n}[I(DS)K_H(\mathbf{X_j} - \mathbf{X_i})]}{\sum_{j=1}^{n}[K_H(\mathbf{X_j} - \mathbf{X_i})]}\end{aligned} \tag{3}$$

where $\hat{G}_i$ is the POE of damage at the i^{th} sample (i.e., earthquake event) $(i = 1, 2, \ldots, n)$, DS is the damage state or the damage class with discrete values. Here, $DS = 0$ corresponds to undamaged state and $DS = 1, 2, \ldots$ correspond to different levels of damage. Moreover, $\mathbf{X_i}$ corresponds to the m-dimensional feature vector, $I(DS)$ is the indicator function ($I = 0$ for undamaged state, i.e., $DS = 0$, and $I = 1$ for the damaged state, i.e., $DS \geq 1$), $j = 1, 2, \ldots, n$ refers to the observed data represented by the considered events, and K_H is a multivariate Gaussian kernel.[5] For any vector $\mathbf{V}$, $K_H(\mathbf{V})$ is defined as follows,

$$K_H(\mathbf{V}) = \frac{1}{\sqrt{(2\pi)^m |\mathbf{H}|}} \frac{1}{\sqrt{e^{\mathbf{V}^T \mathbf{H}^{-1} \mathbf{V}}}} \tag{4}$$

where the $m \times m$ bandwidth matrix $\mathbf{H}$ is analogous to the covariance matrix controlling the amount and orientation of the induced smoothing, $|\mathbf{H}|$ is the 2-norm of the matrix $\mathbf{H}$, and superscript T indicates transpose. The kernel assigns a different weight to each data point of $\mathbf{X_i}$

[3]Pacific Earthquake Engineering Research Center.

[4]Next Generation Attenuation Ground Motion Database.

[5]Kernel density estimation is a nonparametric technique for the estimation of the probability density function. In this case the estimate of the density function with bandwidth (smoothing) characterized by H is expressed as $\hat{f}_H(\mathbf{X_i}) = \frac{1}{n}\sum_{j=1}^{n} K_H(\mathbf{X_j} - \mathbf{X_i})$.

that is inversely related to the distance between the data point j and the point of interest i. The Gaussian kernel is centered at the event of interest, i.e., at $\mathbf{X_i}$ and the developed envelope is analogous to the fragility curves. Several studies have applied Gaussian kernel densities to develop univariate fragility curves (Noh et al., 2012, 2015; Lallemant et al., 2015). However, in this study, multiple variables are used.

Figure 2 shows an example POE envelope on a parallel coordinates plot. In this plot, each feature is represented by a vertical axis and individual samples (i.e., earthquake events) are displayed as a line connected across each axis. The colored areas show the POE envelope. The color represents the probability values of the distribution. Green and red represent undamaged and damaged zones, respectively. Moreover, probability of damage increases with shade of the zone, with red zone having the high probability of damage. In this figure two example samples are shown. One falls in the undamaged zone and is hence labelled as undamaged. The other one falls in the red zone and is labelled as damaged.

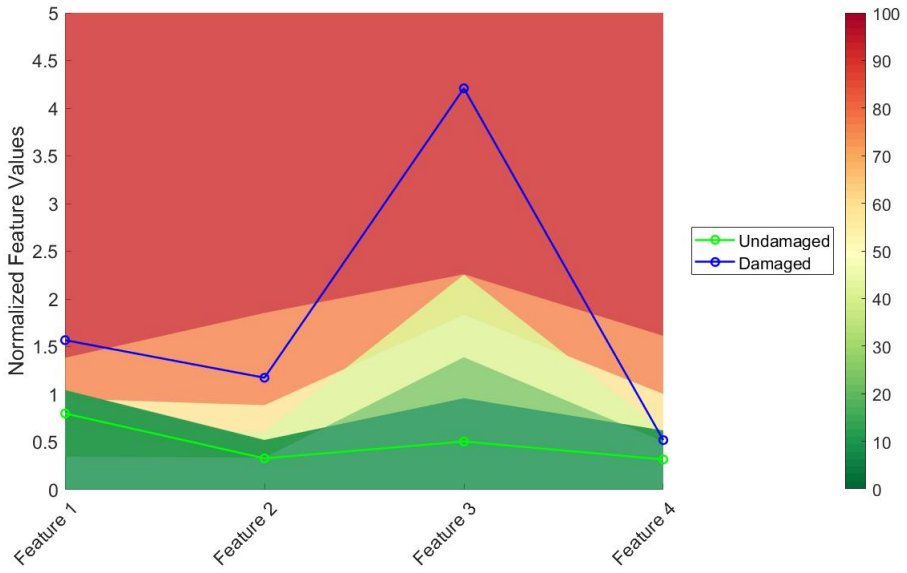


Figure 2: Parallel coordinate plot showing POE envelope with undamaged and damaged zones.

Damage Detection

In this section, the damage detection routine of the H-MC framework is described. For an instrumented bridge, data from previous earthquakes is used to develop the novelty model $\mathrm{M}(\theta)$. The POE envelope for the bridge is developed simultaneously using a SDOF model prior to the earthquakes. In this POE model, the threshold POE value (POE_{th}) for undamaged events and the entries of the bandwidth matrix $\mathbf{H}$ are used as model parameters. The optimal values of these parameters are determined using a 5-fold cross-validation technique.

When data becomes available from a new earthquake, the H-MC framework is activated. It first calculates the features and runs the novelty model. If novelty is not detected, the bridge is identified as undamaged. On the other hand, if novelty is detected, the second phase

of the framework is initiated where the POE value is computed as follows,

$$\hat{G}_i = P(DS \geq 1|\mathbf{X_i^{rec}}) = \frac{\sum_{j=1}^{n}[I(DS)K_H(\mathbf{X_j^{com}} - \mathbf{X_i^{rec}})]}{\sum_{j=1}^{n}[K_H(\mathbf{X_j^{com}} - \mathbf{X_i^{rec}})]} \tag{5}$$

where $\mathbf{X_i^{rec}}$ and $\mathbf{X_j^{com}}$ are respectively the m-dimensional feature vectors of the recorded data and the computed data using the SDOF analysis where $j = 1, 2, \ldots, n$. Subsequently, $\hat{G}_i$ is compared against POE_{th}. When $\hat{G}_i \leq \text{POE}_{th}$, the bridge response to the i^{th} event is identified as undamaged. On the other hand, when $\hat{G}_i > \text{POE}_{th}$, i.e., the data falls in the high probability region, the bridge response to the i^{th} event is identified as damaged.

Damage Assessment

The probability of different damage states for the i^{th} sample is calculated as follows,

$$\hat{G}_{id} = P(DS = d|\mathbf{X_i^{rec}}) = \frac{\sum_{j=1}^{n}[I(DS_d)K_H(\mathbf{X_j^{com}} - \mathbf{X_i^{rec}})]}{\sum_{j=1}^{n}[K_H(\mathbf{X_j^{com}} - \mathbf{X_i^{rec}})]} \tag{6}$$

where $\hat{G}_{id}$ is the probability that the i^{th} sample (i.e., earthquake event) belongs to the damage state d for different levels of damage, i.e., $d = 1, 2, \ldots, D$ where D is the number of considered damage (limit) states, and $I(DS_d)$ is the indicator function ($I = 1$ when $DS = d$, and $I = 0$ otherwise).

When data becomes available from a new earthquake, novelty analysis is conducted. When the data is detected as a novelty, first the damage detection with the POE envelope is checked. The damage assessment follows the damage detection phase. This assessment step is conducted only on the samples for which damage has been detected in the previous step. Once the values of $\hat{G}_{id} \; \forall \; d = 1, 2, \ldots, D$ are computed for the i^{th} sample, the damage state with the highest probability value is assigned as the damage state of this sample.

Numerical Example

A finite element (FE) model of a selected bridge is developed for the study to determine important damage features. A damage feature is some quantity extracted from the measured system response data that is used to indicate the presence of damage in a structure. The simulated vibration response histories of the FE bridge model to a set of 400 selected ground motions is used to extract initial damage features, such as peak accelerations, and corresponding damage states, which range from no damage to extreme damage. The dataset of corresponding damage features and damage states is later used in a feature engineering process called *feature selection* to identify the best damage features for classification of the damage states.

Analytical Bridge Model

The selected bridge for the analytical model is the Jack Tone Road Overcrossing Bridge (referred to for simplicity as Bridge-A), which is a two-span reinforced concrete (RC) highway bridge located in San Joaquin County, California, spanning approximately 220 feet with a

total height of approximately 25 feet. As shown in the photograph of the structure and its original design elevation drawings in Figure 3, the bridge is supported on a single column bent and on seat-type abutments. The FE model of the bridge includes 20 nodes along the deck length and a fiber-section model of the column bent. NTHA of this bridge has been studied extensively to determine the effect of abutment skew-angle on the probability of collapse (Kaviani et al., 2012), improve direct integration algorithms for nonlinear seismic response (Liang et al., 2016), and select and modify bidirectional ground motions for bridge seismic behavior evaluation (Liang and Mosalam, 2020).

Figure 3: Photograph and design elevation of Jack Tone Road Overcrossing Bridge (Bridge-A).

The superstructure and column bent of the Jack Tone Road Overcrossing Bridge both have concrete strength of 5 ksi. The column bent is reinforced with ASTM A706 steel, with 44 #11 bundles of 2 longitudinal bars at a volumetric ratio of 2.00%, and #6 spiral transverse reinforcement at a spacing of 3.34 inches. The structural properties are summarized in Table 2.

Table 2: Jack Tone Road Overcrossing Bridge (Bridge-A) properties (Liang et al., 2016).

# of spans	2
Column bent	Single-column
Column radius	33.1 in.
Column height	22.0 ft.
Abutment	Seat type
Seat length	33.85 in.
Superstructure concrete	$f'_c = 5$ ksi, $E_c = 4030.5$ ksi
Column bent concrete & reinforcing materials	Concrete: 5 ksi, Steel: ASTM A706
Reinf. details of column bent cross-section	Long.: 44 #11 (bundles of 2), $\rho_l = 2.00\%$ Trans.: Spiral, #6 @ 3.34 in.

The bridge FE model simplifies the deck into two symmetric spans of $110'$-$3''$ each, with 20 nodes along the length of the deck. The column bent is modeled with three nodes at

the base of the column, the interface of the column-to-deck joint, and the center of the column-to-deck joint. For the abutments, nonlinear springs connected in series to gap elements are used to model the passive backfill response and the expansion joint, and the shear key response is modeled using an elastic-perfectly-plastic backbone curve (Liang et al., 2016). The super-structure FE modeling details are shown in Figure 4. In order to capture progressive column yielding and damage, the column bent is modeled as a nonlinear beam-column with nonlinear fiber-based beam FE's at 10 quadrature points (Kaviani et al., 2012); refer to Figure 5.

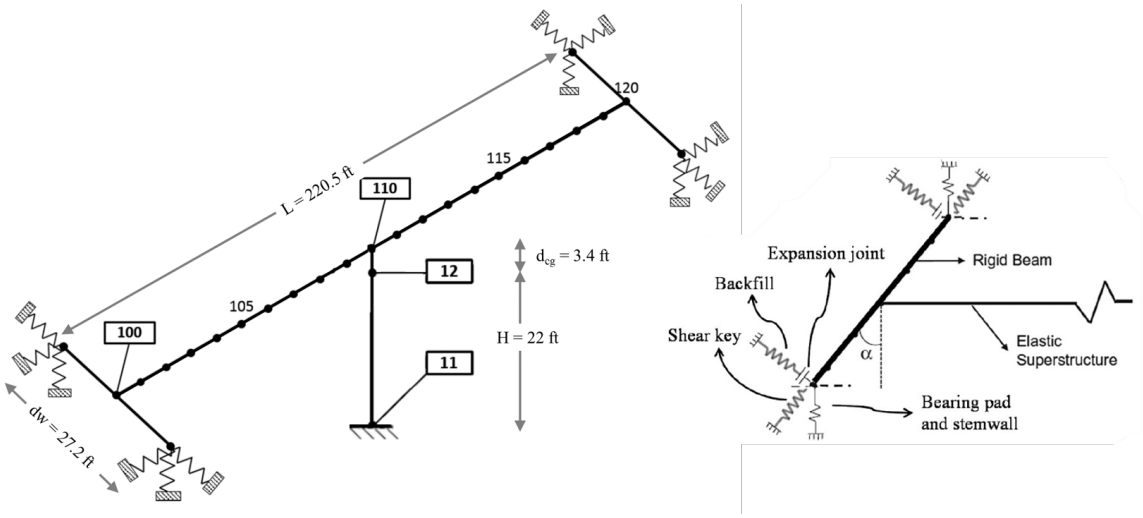


Figure 4: Bridge-A superstructure FE modeling details (Liang et al., 2016).

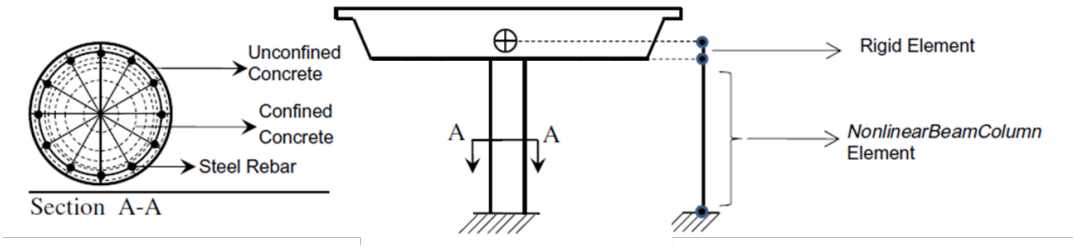


Figure 5: Bridge-A column FE modeling details (Kaviani et al., 2012).

Ground Motion Selection

To produce the vibration response histories from which damage features are extracted, NTHA is performed on the bridge model using 400 ground motions, selected and scaled from the PEER NGA-West2 (Bozorgnia et al., 2014) database, which has 21,539 records of shallow crustal earthquakes in active tectonic regions. Before selecting and modifying the records, they are restricted to the 1,710 ground motions mentioned in "POE Envelope" Section.

Unconditional Selection (US), a spectrum shape matching ground motion selection and modification method, is used to select and scale the ground motions. In the US method, the

median+1.5 standard deviation, σ, spectrum is used as the target spectrum. First, a Monte Carlo simulation is used to probabilistically generate multiple response spectra from a multivariate normal distribution with the logarithmic spectral acceleration means and variances. Then, ground motion records, whose response spectra match the simulated spectra, are selected. Finally, a *greedy optimization* algorithm is applied to improve the match between the target and the sample means and variances. In each iteration of the optimization, one previously selected record is replaced with a record from the database that generates the best improvement in the match (Jayaram et al., 2011).

Out of four common ground motion selection and modification methods, the US method was shown in (Liang and Mosalam, 2020) to produce the best prediction for the covariances of the probability distributions of the seismic demands (PDSD), and the most conservative estimates for the probability of collapse for RC highway bridges with nonlinear responses due to large bidirectional ground motions. Thus, the US method was used to select ground motions for three ground motion scenarios, each with site soil shear wave velocity V_{s30} of 222.95 m/s, strike-slip fault type, and target spectrum of 1.5σ above the median spectrum (for all periods) using the attenuation model in (Campbell and Bozorgnia, 2008). Each ground motion scenario is summarized in Table 3 with 100, 50, and 50 ground motions for the M7d23, M5d50, and M8d23 ground motion scenarios, respectively. These ground motions were each rotated by 90° for a total of 400 ground motions. The results of the US method are summarized in Table 4, which lists the PEER NGA-West2 earthquake ID numbers, earthquake names, and the number of ground motions per earthquake event for the selected set of ground motions with the corresponding response spectra shown in Figure 6.

Damage State Definition

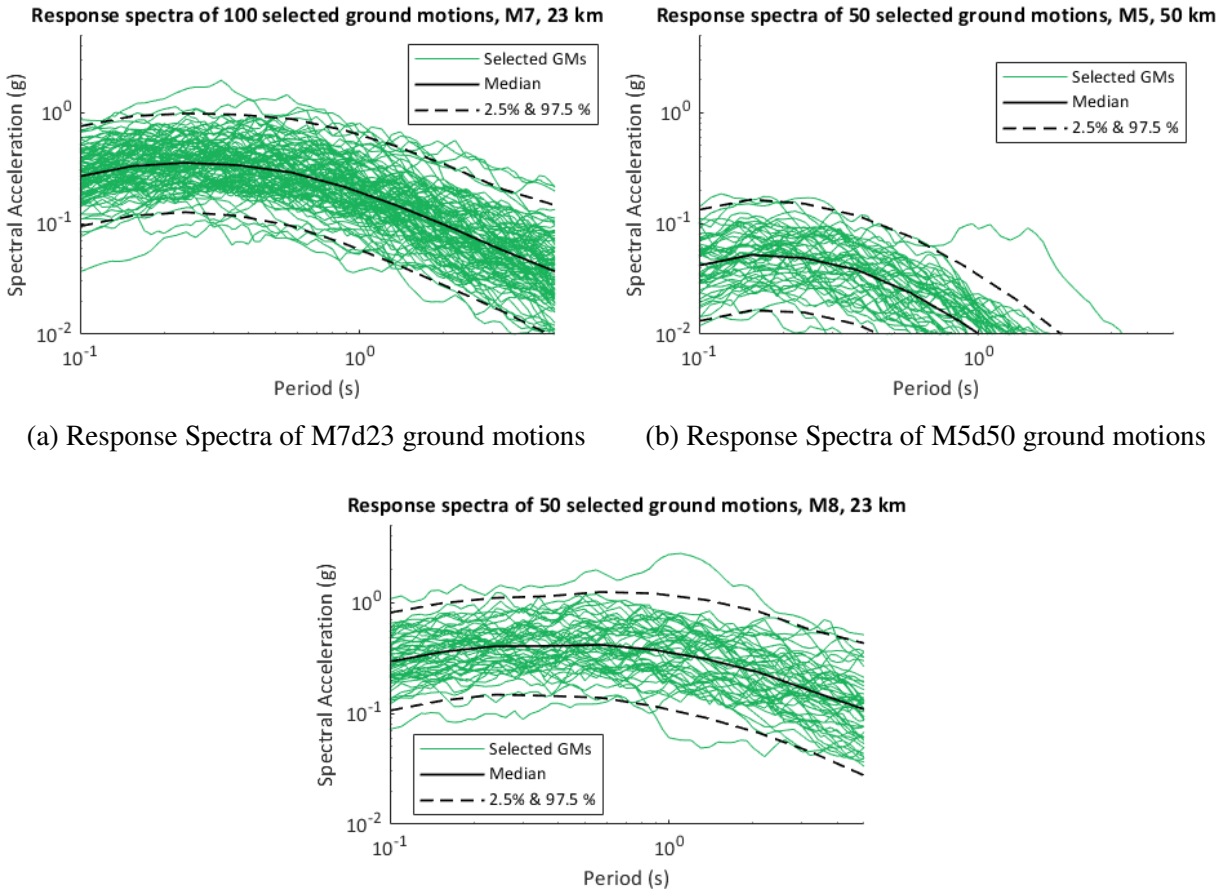
During the NTHA of the bridge model response to the selected ground motions, the stress and strain response histories are recorded at the base and top of the bridge column in order to determine the resulting damage state of the bridge. This study employs a component capacity-based damage state model. In this model, the damage states are defined in seven levels. The damage state criteria are based on Caltrans and PEER visual inspection criteria for levels of column flexural capacity for RC highway bridges. The capacity levels are translated into strain-based limit states at different locations of the cover concrete, core concrete, and longitudinal reinforcing bars, as listed in Table 5. The pushover curve for the bridge model is labeled with each component capacity-based damage state in Figure 7. For the analytical bridge model used in this study, limit states 2, 3, and 4 lie in close proximity to each other on the pushover curve because of the large size of the column, which results in the cover thickness

Table 3: Ground motion scenario summary.

Ground Motion Scenario	Magnitude (M)	Distance (R) [km]	# Records Selected	# Rotated (90°) Records	Total # Records
M7d23	7.0	23	100	100	200
M5d50	5.0	50	50	50	100
M8d23	8.0	23	50	50	100
		Total			**400**

Table 4: Description of selected ground motions.

PEER NGA-West2 earthquake ID number	Earthquake name	Number of ground motions per earthquake event
163	Anza-02	2
170	Big Bear City	1
123	Cape Mendocino	1
173	Chi-Chi (aftershock 3), Taiwan	1
174	Chi-Chi (aftershock 4), Taiwan	1
137	Chi-Chi, Taiwan	34
171	Chi-Chi, Taiwan-02	4
172	Chi-Chi, Taiwan-03	5
173	Chi-Chi, Taiwan-04	6
174	Chi-Chi, Taiwan-05	5
175	Chi-Chi, Taiwan-06	2
346	Christchurch, New Zealand	5
278	Chuetsu-oki	7
281	Darfield, New Zealand	11
169	Denali, Alaska	5
138	Duzce, Turkey	1
280	El Mayor-Cucapah	15
166	Gilroy	1
164	Gulf of California	1
158	Hector Mine	8
279	Iwate	12
129	Kobe, Japan	2
136	Kocaeli, Turkey	3
125	Landers	14
275	L'Aquila (aftershock 1), Italy	2
274	L'Aquila, Italy	1
152	Little Skull Mtn,NV	1
162	Mohawk Val, Portola	1
180	Niigata, Japan	3
127	Northridge-01	5
150	Northridge-05	1
151	Northridge-06	1
179	Parkfield-02, CA	2
176	Tottori, Japan	2
250	Umbria Marche (aftershock 16), Italy	1
277	Wenchuan, China	8
283	Wenchuan, China-02	1
284	Wenchuan, China-03	1
285	Wenchuan, China-04	1
292	Wenchuan, China-11	1
295	Wenchuan, China-14	1
297	Wenchuan, China-16	1
320	Wenchuan, China-39	1
167	Yorba Linda	1
1015	9753485	2
1018	10370141	1
1011	10410337	5
1014	14138080	1
1003	14151344	1
1002	14383980	3
1049	21266207	2
1001	40204628	2
Total (before rotation)		**200**



(a) Response Spectra of M7d23 ground motions

(b) Response Spectra of M5d50 ground motions

(c) Response Spectra of M8d23 ground motions

Figure 6: Response spectra plots for each of the three ground motion scenarios: (a) M7d23, (b) M5d50, and (c) M8d23, including median and 2.5 & 97.5 percentiles.

constituting a small percentage (i.e., ∼6%) of the column depth. Using the assumption of a linear strain profile throughout the column section (i.e., the assumption made by the Bernoulli-Euler beam theory that "plane sections remain plane"), small changes in the curvature result in the inner cover fibers reach the spalling strain quickly after the outer cover fibers reach the spalling strain. The definitions of limit states 2, 3, and 4 are highly dependent on the model discretization of the cover due to fiber mesh sensitivity. Such sensitivity to model discretization in the nonlinear response of RC beam-columns is discussed in (Kenawy et al., 2020).

Feature Engineering

Feature Extraction

Feature extraction refers to the process of transforming the measured data into some alternative form where the correlation with the damage is more readily observed. After NTHA is performed on the bridge model with the details shown in Figure 4 and using the selected ground motions from the "Ground Motion Selection" Section, damage features are extracted from the resulting vibration response histories. The vibration response at the top of the col-

Table 5: Damage state criteria for component capacity-based model.

Limit State	Caltrans-PEER Description	Criteria Used to Define Strain-Based Limit State	Strain-Based Limit State		
			Fiber Location	Compression or Tension	Strain
0	No damage	–	–	–	–
1	EQ-related tight cracking of cover	Cover cracking: the cover surface reaches tensile strength	Any outermost cover fiber	Tension	$\varepsilon_t = 1.32 \times 10^{-4}$
2	Moderate cracking (mixed orientations) & minor spalling/flaking	Minor spalling: the cover surface reaches compressive strength	Any outermost cover fiber	Compression	$\varepsilon_{sp} = 0.005$
3	Open cracking or major spalling (exterior to confinement)	Major spalling: a significant depth of the cover reaches compressive strength	Any cover fiber at 1/2-3/4 of the cover depth	Compression	$\varepsilon_{sp} = 0.005$
4	Exposed core (interior of confinement)	Exposed core: the entire depth of the cover reaches compressive strength	Any innermost cover fiber	Compression	$\varepsilon_{sp} = 0.005$
5	Visible bar buckling; confinement loss or core shedding	Core shedding: the outer surface of the core begins to fail	Any outermost core fiber	Compression	$\varepsilon_{cu} = 0.025$
	Multi-bar rupture or buckling; large drift; or core crushing	Bar rupture: longitudinal bars reach ultimate tensile strength	Any longitudinal bar	Tension	$\varepsilon_{su} = 0.10$
6	Column collapse (Near-total loss of axial capacity)	Loss of axial capacity: approximately 1/2 of the core fails	All core fibers at 1/4 of the core depth	Compression	$\varepsilon_{cu} = 0.025$

umn bent (Node 110 in Figure 4) is used because the top of the bridge piers are commonly instrumented locations, allowing direct future comparisons to instrumented bridges. The damage features extracted from the vibration response histories include one-dimensional vibration characteristics that have been studied as structural damage indicators by researchers in the past (Farrar and Worden, 2012). In addition to indicating damage on a global level, these features can also be measured at specific locations on the structure, as in the case of this study, to identify the local damage. Other examples of features which have been studied as structural damage indicators include cumulative absolute velocity (*CAV*) (Muin and Mosalam, 2018), higher exponentiations of the acceleration intensity (Sajedi and Liang, 2019), standardized *CAV* (CAV_{STD}) (Campbell and Bozorgnia, 2012), and instantaneous power (*IP*) (Zengin and Abrahamson, 2020). Sixteen damage features, Table 6, are extracted herein in two directions, longitudinal and transverse, for a total of 32 damage features. After the features are extracted, feature selection is performed to identify the best features for the intended damage classification.

Fisher Score Analysis

Feature selection is the process of identifying a subset of the original feature set which increases the learning efficiency while minimizing the reduction in the classification performance. It is performed to determine the best feature vector for this study. Moreover, a filter-based feature selection method, or one in which a *score* is assigned to each feature in order to compute its expected contribution to solving the classification task, is chosen as a computationally inexpensive approach that allows feature selection before the learning phase (Baccianella et al., 2014). One such method is the *Fisher score*, which determines the relative ability of fea-

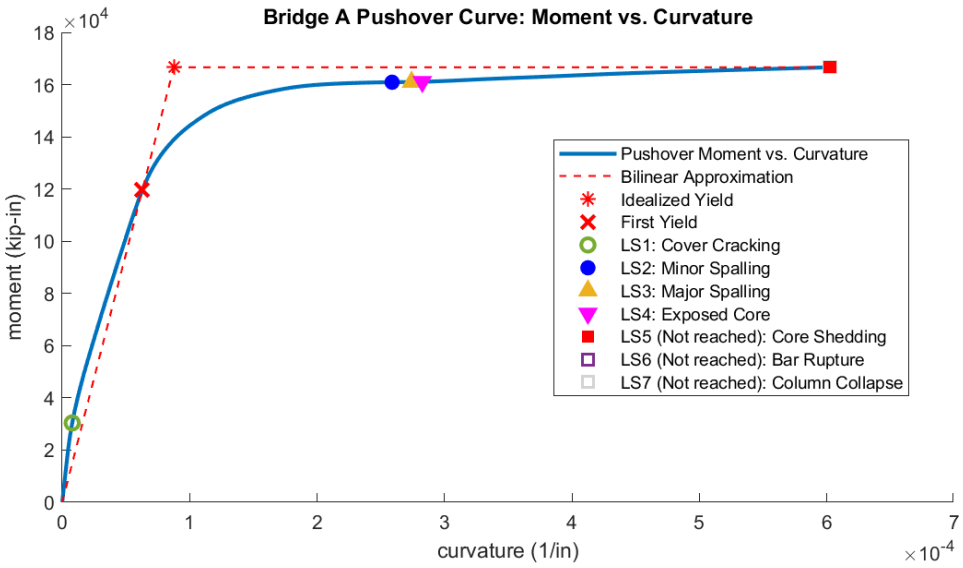


Figure 7: Bridge-A model pushover curve labeled with component capacity-based damage states.

tures to discriminate between categorical classes in a classification model. The Fisher Score is also known in statistical modeling as Fisher ratio, Fisher method, Fisher combined probability score, or information score. The Fisher score for a given feature is defined as the average distance between classes, normalized by the average spread of each class, and a larger Fisher score indicates that the feature is more discriminative. A visualization of the discriminatory power concept used to develop the Fisher score is shown in Figure 8, where 8a shows two features showing a clear distinction among classes, and thus a high Fisher score, and 8b shows two features not ideal for detecting damage, and thus a low Fisher score. Although the Fisher score is known to have limitations in that it considers each feature separately and therefore cannot reveal *mutual information* between features, it is widely used as a heuristic algorithm for feature selection (Gu et al., 2012).

In this study, the goal of the classification model is to determine which class of damage state is indicated by the feature vector. Due to the ordering of the classes in this model (e.g., the 5^{th} damage (limit) state is more severe, and thus ordered higher than the 3^{rd} damage (limit) state, refer to Table 5), feature selection which considers ordinality is required. Thus, an ordinal variant of the Fisher score, termed here as the *Ordinal Fisher Score*, is used for feature selection in this study. The Ordinal Fisher Score is shown to perform better for feature selection in an ordinal classification model than the original Fisher score (Pérez-Ortiz et al., 2016).

The Ordinal Fisher Score for the i^{th} feature, $F_{OR}(x^i)$, is computed following a procedure discussed below. First, the class discrimination term, $F_O(x^i)$ is defined as follows,

$$F_O(x^i) = \frac{\sum_{k=1}^{K}\sum_{j=1}^{K}|k-j|\cdot d^i(\mathscr{C}_k,\mathscr{C}_j)}{(K-1)\sum_{k=1}^{K}(\sigma_k^i)^2} \tag{7}$$

where the class $\mathscr{C}_k$ indicates data from one of the K classes, $k \in 1,\ 2,\ldots,K$, σ_k^i indicates the variance of the i^{th} feature in class k, and $d^i(\mathscr{C}_k,\mathscr{C}_j)$ indicates the distance between classes

Table 6: Description of initial damage features extracted from NTHA for later feature selection.

Feature	Definition or Equation
Peak acceleration	$\max_t(\lvert\ddot{u}(t)\rvert)$
Peak displacement	$\max_t(\lvert u(t)\rvert)$
Peak velocity	$\max_t(\lvert\dot{u}(t)\rvert)$
Arias Intensity	$I_a = \int_0^T [\ddot{u}(t)]^2 dt$
Significant Duration	time between 5% & 95% of the total Arias energy $= t_{0.95} - t_{0.05}$
Cubed Absolute Acceleration Intensity (Sajedi and Liang, 2019)	$\int_0^T [\ddot{u}(t)]^3 dt$
Spectral Acceleration at 0.6 s	$S_a(0.6)$
Spectral Acceleration at 1.0 s	$S_a(1.0)$
Spectral Acceleration at 2.0 s	$S_a(2.0)$
Spectral Acceleration at 5.0 s	$S_a(5.0)$
CAV	$CAV = \int_0^T \lvert\ddot{u}(t)\rvert dt$
R_{CAV}	$R_{CAV} = \frac{CAV_s}{CAV_l}$
Sum of absolute displacements	$\int_0^T \lvert u(t)\rvert dt$
Sum of absolute velocities	$\int_0^T \lvert\dot{u}(t)\rvert dt$
CAV_{STD} (Campbell and Bozorgnia, 2012)	$CAV_{STD} = \sum_{i=1}^{N} \left(H(PGA_i - 0.025) \int_{i-1}^{i} \lvert\ddot{u}(t)\rvert dt \right)$ $H(x)$: Heaviside function ($= 0$ if $x < 0$ & $= 1$ otherwise) N: # non-overlapping 1 sec. time intervals
Instantaneous Power (Zengin and Abrahamson, 2020)	$IP(T_1) = \max_t(\frac{1}{\Delta t} \int_t^{t+\Delta t} V_{filtered}^2(t) dt)$ $V_{filtered} = bandpass\,(\ddot{u}(t), [0.2T_1, 3T_1])$

$\mathscr{C}_k$ and $\mathscr{C}_j$. Next, an ordinality term, $O_R(x^i)$, which incorporates the relative distance between classes, is defined as follows,

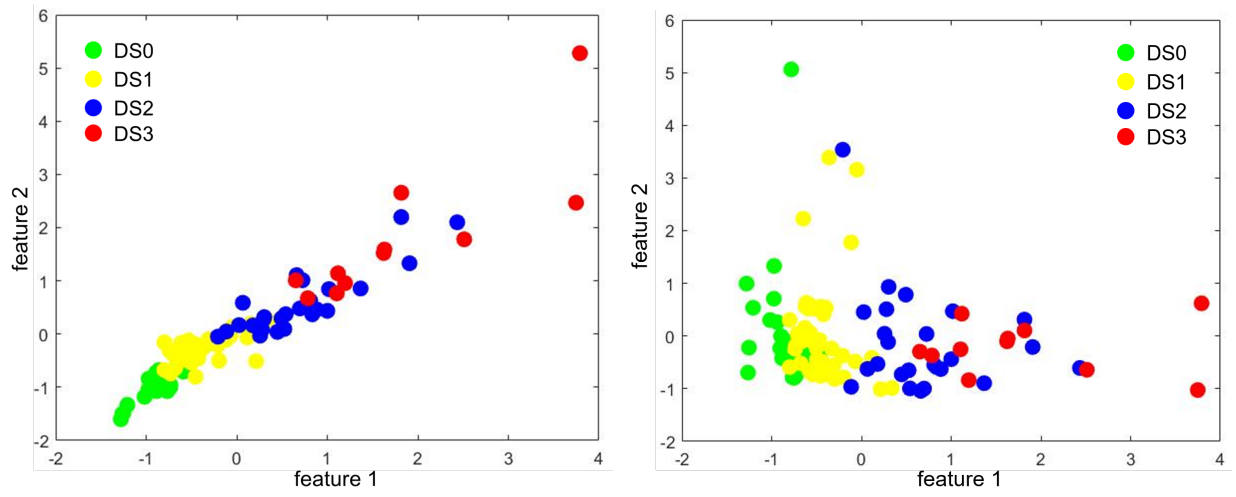
$$O_R(x^i) = \frac{\sum_{k=1}^{K-2} \sum_{j=k+1}^{K-1} \sum_{h=j+1}^{K} [\![(d^i(\mathscr{C}_k, \mathscr{C}_h) - d^i(\mathscr{C}_k, \mathscr{C}_j)) > 0]\!]}{\sum_{j=2}^{K-1} (K-j)} \tag{8}$$

where $[\![\cdot]\!]$ is a Boolean test which is 1 if the inner condition is true, and 0 otherwise. This $O_R(x^i)$ score measures the number of ordinal requirements fulfilled for a specific feature i. Finally, the two terms $F_O(x^i)$ and $O_R(x^i)$ are combined in a weighted sum to determine the Ordinal Fisher Score for the i^{th} feature, $F_{OR}(x^i)$ as follows,

$$F_{OR}(x^i) = \alpha \cdot F_O(x^i) + (1-\alpha) \cdot O_R(x^i), \quad \alpha \in (0,1). \tag{9}$$

The distance between classes, $d^i(\mathscr{C}_z, \mathscr{C}_j)$, can be computed in several ways, and this study uses the *Hausdorff distance*, $d_h^i(\mathscr{C}_z, \mathscr{C}_j)$, formulated to allow for use of nonlinear, multi-modal, or non-normal data, and shown in (Pérez-Ortiz et al., 2016) to perform best for ordinal classification. The Hausdorff distance is defined as follows,

$$d_h^i(\mathscr{C}_z, \mathscr{C}_j) = \max \left\{ \max_{x_h^i \in \mathscr{C}_z} \Delta_m(x_h^i, \mathscr{C}_j), \max_{x_v^i \in \mathscr{C}_j} \Delta_m(x_v^i, \mathscr{C}_z) \right\} \tag{10}$$



(a) Features with higher discriminatory power (b) Features with lower discriminatory power

Figure 8: Visualization of discriminatory power using a 2D feature class comparison. (a) Features with a clear distinction among different classes have higher discriminatory power and are more useful for classification. (b) Features without clear distinction are not useful in classification.

where $\Delta_m(x_h^i, \mathscr{C}_j)$ is the minimum Euclidean distance between a point x_h^i and the points in class $\mathscr{C}_j$, expressed as follows,

$$\Delta_m(x_h^i, \mathscr{C}_j) = \min_{x_v^i \in C_j} d(x_h^i, x_v^i) = \min_{x_v^i \in C_j} \sqrt{(x_h^i - x_v^i)^2}. \tag{11}$$

Starting with the 32 initial damage features extracted from the NTHA (Table 6), the Ordinal Fisher Score with weighting factor set to $\alpha = 0.5$ is used to identify which features best discriminate between the different classes of the damage states. The scores of each feature are compared in Figure 9. The results of the Ordinal Fisher Score analysis show that the best features for classification are: 1) peak acceleration in X (longitudinal) direction, 2) spectral acceleration at 0.6 seconds, $S_a(0.6)$, in X direction, 3) sum of absolute velocities in X direction, 4) CAV_{STD} in X direction, 5) CAV_{STD} in Y (transverse) direction, 6) CAV in X direction, and 7) and CAV in Y direction. The results indicate that these features discriminate well between classes of the damage states. The best features for classification based on the Ordinal Fisher Score analysis, Table 7, form a reduced feature vector which allows for greater computational efficiency and are next evaluated in the H-MC analysis for classification accuracy.

k-fold Cross-Validation Analysis

k-fold cross-validation is a resampling procedure used to evaluate ML models where k is the number of groups a given data sample is split into. The general procedure is as follows:

1. Shuffle the dataset randomly.
2. Split the dataset into k groups.

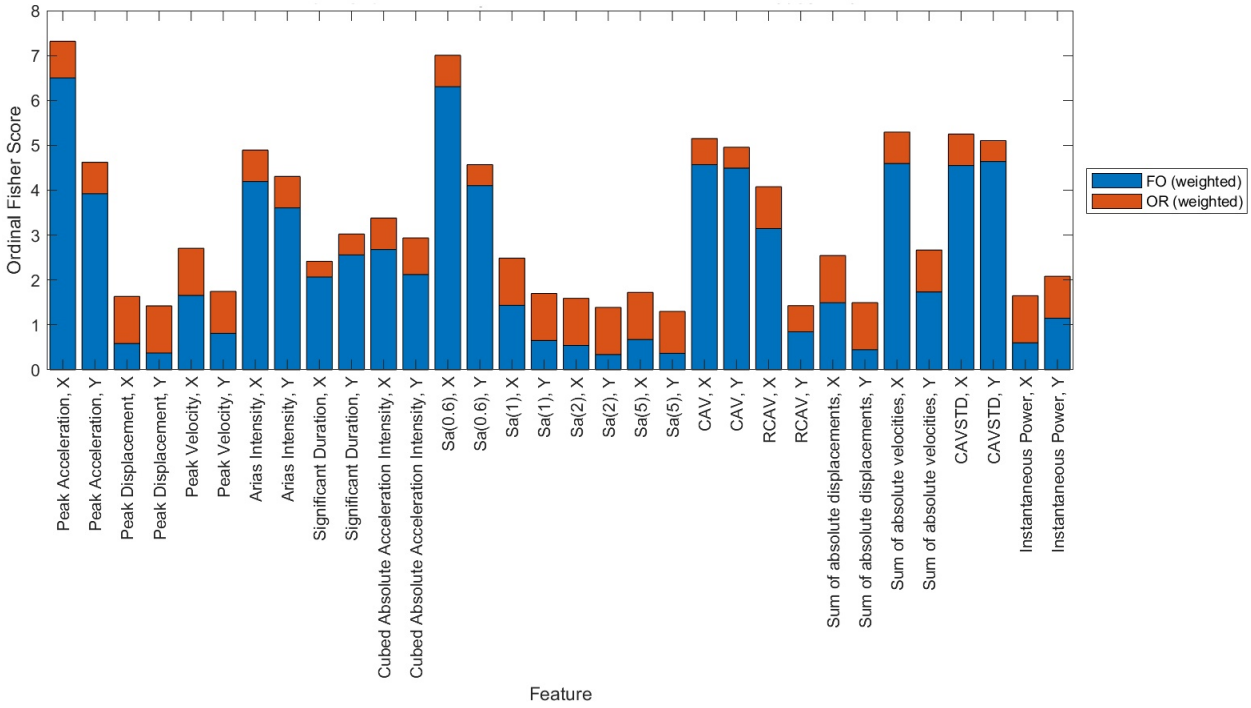


Figure 9: Comparison of Ordinal Fisher Scores for the extracted 32 damage features for the component capacity-based damage state model of Bridge A. *X* & *Y* denote the bridge longitudinal and transverse directions, respectively.

3. For each unique group:
    - Take the group as the validation data set;
    - Take the remaining groups as a training data set; and
    - Fit a model on the training set and evaluate the accuracy on the validation set.
4. Summarize the skill[6] of the model using the average accuracy.

This validation phase gives an opportunity to tune the complexity of the ML model to perform better. In this study, this technique is utilized to finalize the number of features. The first seven selected features from the Ordinal Fisher Score analysis are used to develop a POE model with the SDOF response data. At first, the 5-fold cross-validation technique is applied to the model with one feature (feature with the highest F_{OR}). For each training-validation group, the accuracy is computed as follows,

$$\text{Accuracy} = \frac{\text{\# of correct prediction in the validation set}}{\text{Total \# of samples in the validation set}} \times 100\%. \tag{12}$$

[6]Prediction *skill* is a measure of the accuracy and/or degree of association of prediction to an observation or estimate of the actual value of what is being predicted. An example of a skill calculation is based on the error metric "Mean Squared Error (MSE)".

Table 7: Ordinal Fisher Score analysis results for component capacity-based damage state model. X & Y denote the bridge longitudinal and transverse directions, respectively.

Features with highest Ordinal Fisher Scores
Peak acceleration in X (Feature 1)
$S_a(0.6)$ in X (Feature 2)
Sum of absolute velocities in X (Feature 3)
CAV_{STD} in X (Campbell and Bozorgnia, 2012) (Feature 4)
CAV_{STD} in Y (Campbell and Bozorgnia, 2012) (Feature 5)
CAV in X (Feature 6)
CAV in Y (Feature 7)

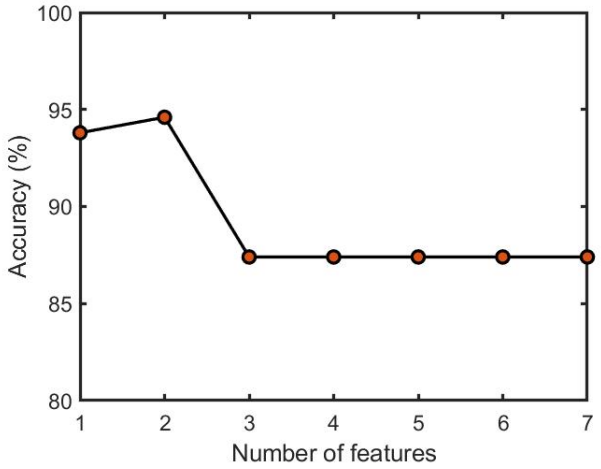


Figure 10: Model complexity (number of features) vs. validation accuracy (skill).

At the end of the 5^{th} groups' analysis, the average accuracy is calculated for the model with one feature. Subsequently, average accuracy is calculated for increasing model complexity (i.e., number of features). Figure 10 shows that the accuracy improves when feature number 2 is introduced in the model but it drops once feature number 3 is added. For the rest of the features, the accuracy remains stable. Therefore, feature number 3 (the sum of absolute velocities in the X direction) is removed as a feature. The final features utilized in the model development are listed in Table 8.

H-MC Analysis

H-MC analysis is performed on the analytical bridge model of Bridge-A using the six selected features (Table 8) and the 400 (Table 4) event responses mentioned earlier. The novelty model is trained on the undamaged responses only. For this purpose, the undamaged responses are separated out of the 400 responses. Three different training and test sets (Set-1, Set-2, and Set-3) are used in the study to evaluate the performance of the model. For each set, 80% of the undamaged sample is randomly selected to train the novelty model. The rest of the undamaged sample is combined with randomly selected damaged events to make up the test set (150 samples).

Table 8: Final feature set after k-fold cross-validation tuning for optimal model complexity; X & Y denote the bridge longitudinal and transverse directions, respectively.

Final Feature Set
Peak acceleration, X (Feature 1)
$S_a(0.6)$, X (Feature 2)
CAV_{STD} (Campbell and Bozorgnia, 2012), X & Y (Features 3 & 4)
CAV, X & Y (Features 5 & 6)

Table 9 presents the accuracy values of the analysis for the three test sets indicating that an average accuracy of 90.2% is achieved when novelty detection alone is used. In contrast, when coupled with the POE analysis for the complete H-MC framework, the average accuracy improves significantly to 98.7%. Therefore, it is evident that the H-MC model is very successful in detecting damage. However, when it comes to assessing damage, the model does a good job with an average accuracy of 82.0%. Figure 11 shows visually the performance of the novelty detection alone and the H-MC method on a tSNE[7] plot. It shows that the novelty detection successfully detects the majority of the damaged events but also misclassifies some of the undamaged events as damaged ones, hence, the accuracy decreases. However, when combined with POE analysis in the H-MC framework, these misclassifed events are correctly classified as undamaged ones.

Although accuracy is an important evaluation measure, it does not present the whole picture. Figure 12 presents the confusion matrices for damage detection using novelty only and H-MC. It shows that novelty detection alone misclassifies a significant portion of the undamaged cases as damaged (58% for Set-1 and Set-2 and 38% for Set-3). In contrast, the H-MC analysis has misclassification rate in all three sets of less than 4% for both the undamaged and damaged states, which further shows the effectiveness of the H-MC framework in damage detection. H-MC classifies a very small fraction of the damaged samples as undamaged which could be a problem for SHM. However, this can be improved by a better model for POE beyond the SDOF.Figure 13 shows the confusion matrices for damage assessment which shows DS1 (Damage (Limit) States (DS) are shown in Table 5) as the dominant class, i.e., most of the test samples are classified as DS1. This can be attributed to the fact that both datasets (60% of the Bridge-A data and 57% of the SDOF data) are governed by the DS1 limit state. Furthermore, none of the 400 samples (corresponding to the simulations using the 400 ground motions of the selected 200 records augmented by their 90° rotations, Table 4) belonged to DS2 and only one sample belonged to DS3. In contrast, DS1 and DS4 had 240 and 57 samples, respectively. This disproportionate distribution of samples occurred due to the close proximity of DS2, DS3, & DS4, which was also reflected on the SDOF training dataset. Therefore, it is expected that the damage assessment performance will be improved when limit states are well-separated and data from all the damage states are equally represented in the SDOF dataset.

[7] t-Distributed Stochastic Neighbor Embedding (tSNE) is a nonlinear technique for dimensionality reduction that is particularly well suited for the visualization of high-dimensional datasets. It is extensively utilized in ML applications. The axes of the tSNE plots, herein labeled $tSNE_1$ and $tSNE_2$, have non-physical meaning due to mappings through iterative nonlinear transformations.

Table 9: Accuracy of the novelty and H-MC model for the three test sets.

Test	Accuracy		
	Novelty Detection	H-MC: Damage Detection	H-MC: Damage Assessment-POE
Set-1	90.7%	100%	82.7%
Set-2	90.7%	97.3%	81.3%
Set-3	89.3%	98.7%	82.0%
Average	**90.23%**	**98.7%**	**82.0%**

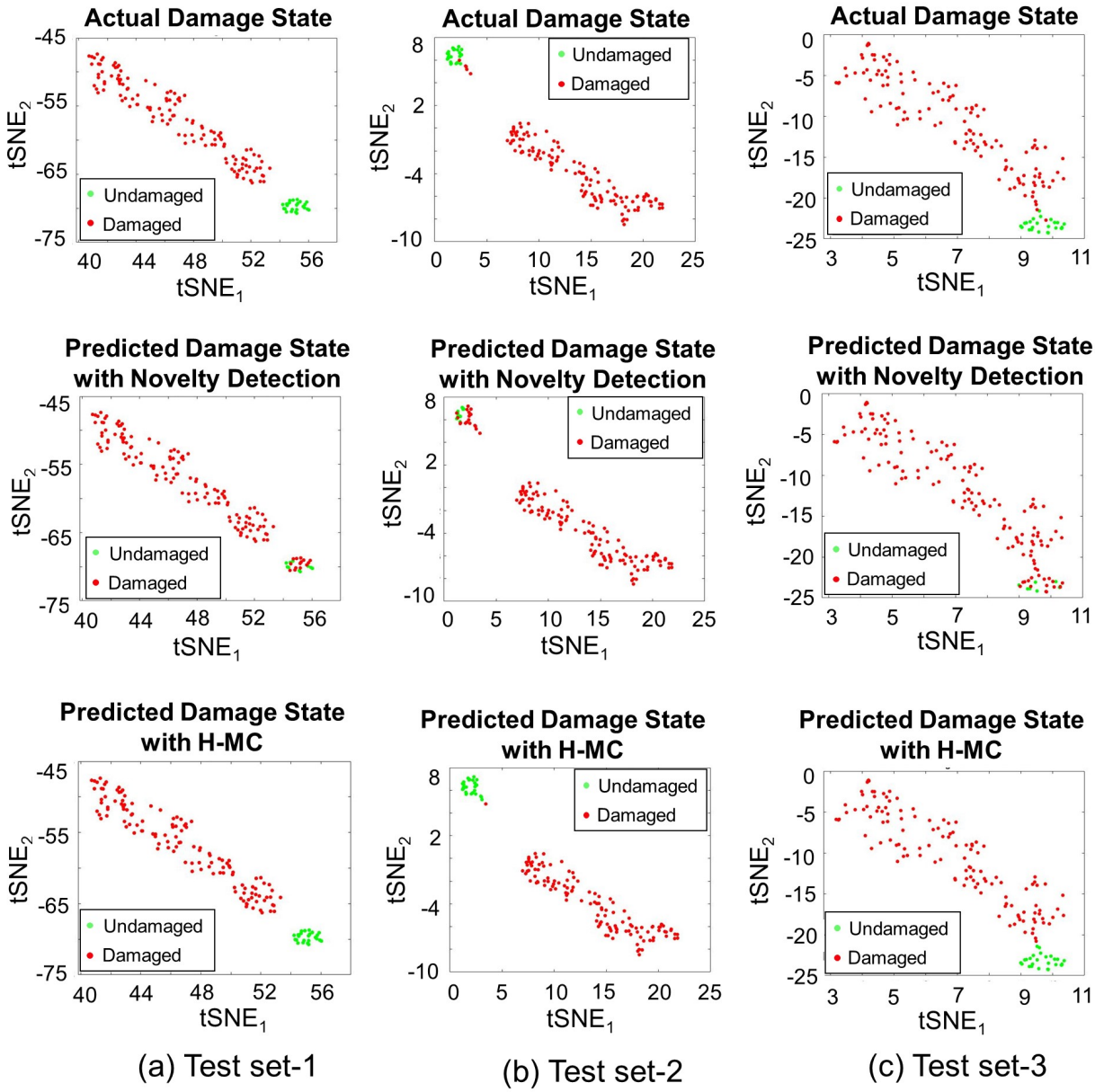


Figure 11: tSNE visualization of the actual and predicted damage states for test sets: (a) Set-1, (b) Set-2, and (c) Set-3.

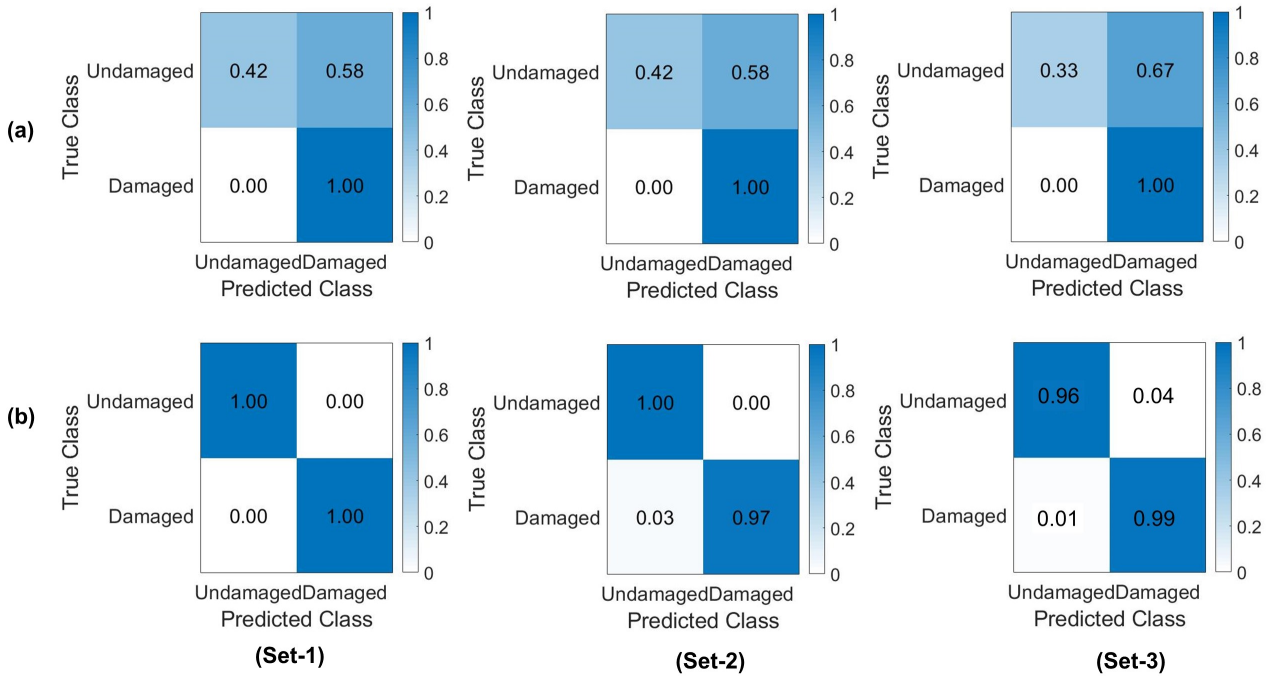


Figure 12: Confusion matrices for damage detection of the three test sets for: (a) Novelty detection model, and (b) H-MC framework (novelty + POE).

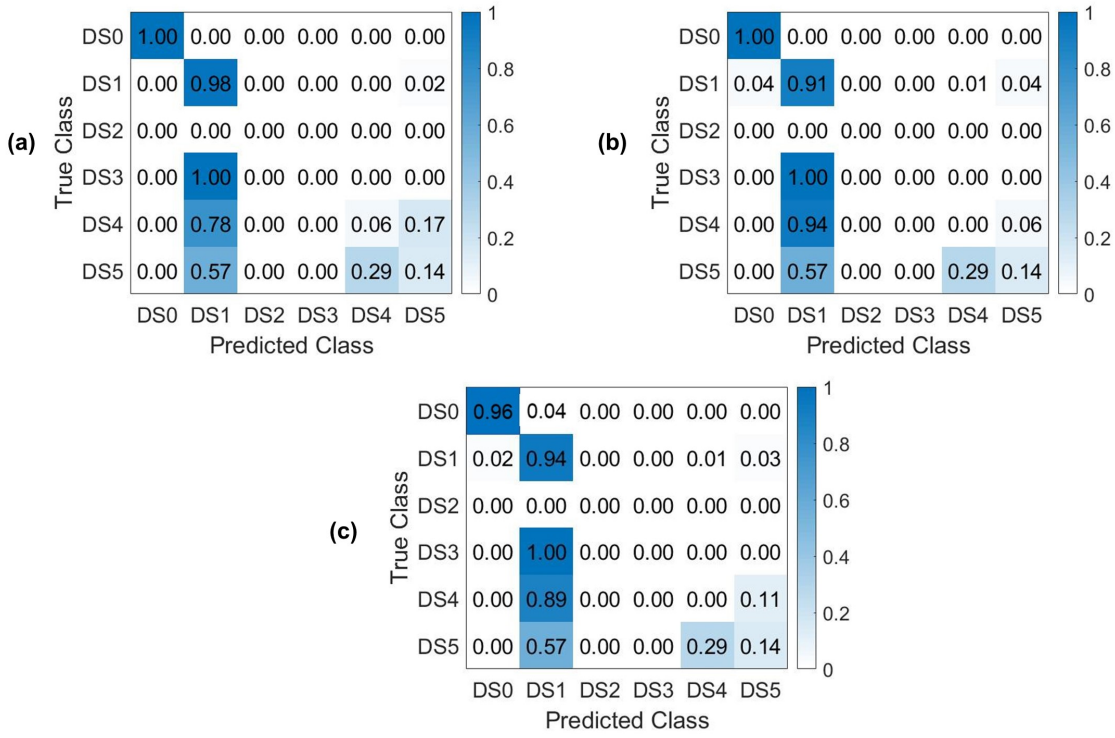


Figure 13: Confusion matrices for damage assessment using H-MC for test sets: (a) Set-1; (b) Set-2; (c) Set-3.

Applications

El Centro - Hwy8/Meloland Overpass

The Meloland Road Overpass which is located near El Centro, California is a non-skewed bridge. It is approximately 208 ft long and 34 ft wide with each span measuring 104 ft. The depth of the deck is 5.5 ft. The height of its 5 ft-diameter column is approximately 21 ft, which is supported on 25 timber piles with a square concrete cap. The monolithic abutment backwalls have a height of approximately 13 ft. The bridge is instrumented with 29 sensors. Figure 14 shows the sensor locations of the bridge. The average feature values of the sensors from 13 recorded earthquake events are utilized in this study.

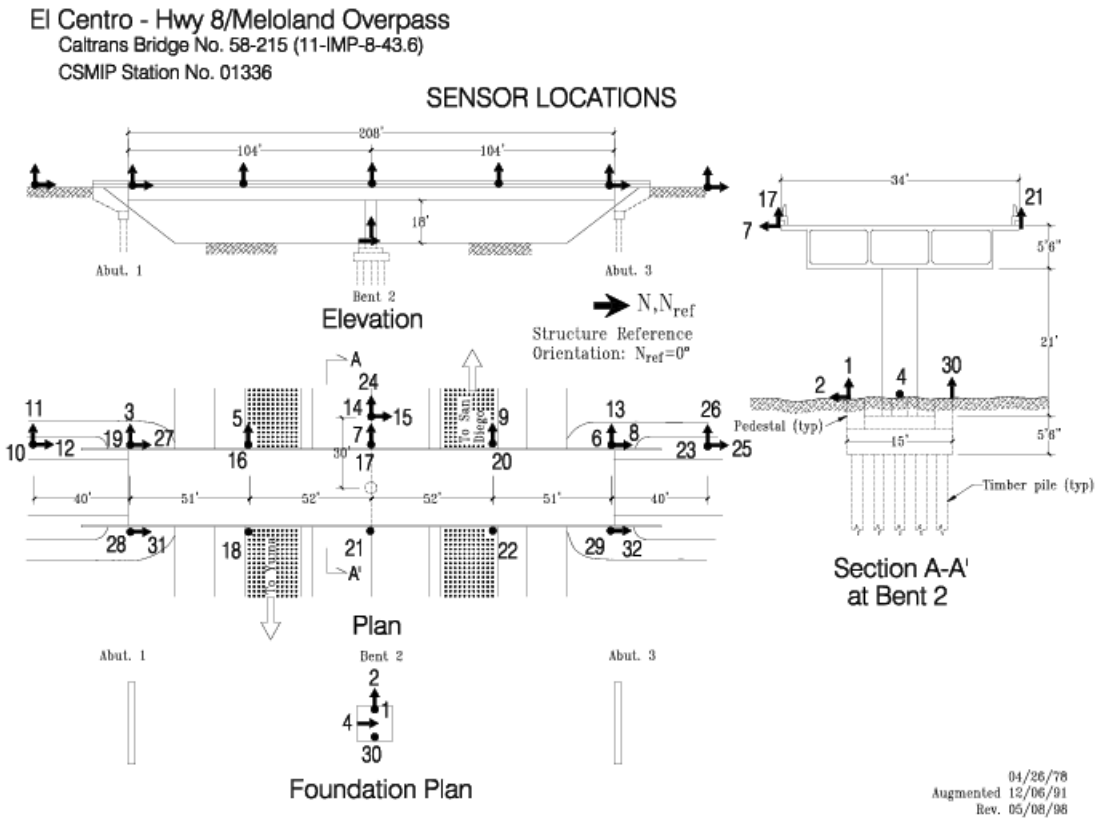
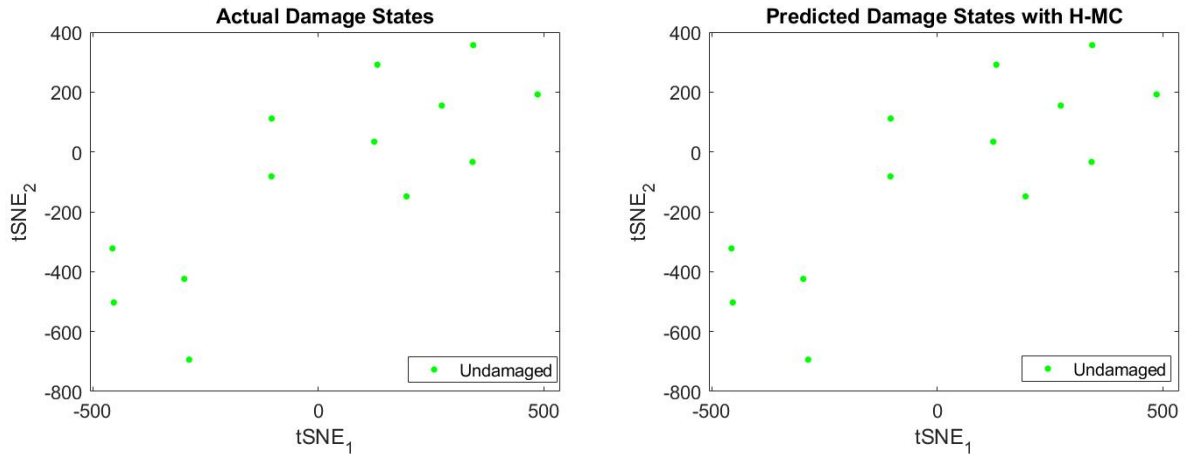


Figure 14: Sensor locations of the Meloland Road Overpass.

For the novelty model, 7 randomly selected events were used to train the model. All 13 events are utilized as the test set. In order to develop the POE envelope, the period reported in (Shamsabadi et al., 2011) is used to develop the SDOF model. Figure 15 shows the actual and predicted tSNE plots for the 13 events. It shows that all the undamaged events have been correctly detected as undamaged. Figure 16 compares the events with the POE envelope. Here, the POE envelope for Feature 1 (peak acceleration, X) spans a smaller distribution than the



(a) Actual damage states (b) Predicted damage states

Figure 15: tSNE plots for the 13 events of Meloland Overpass Bridge.

other features, which results in white space at the top left corner of the plot. In addition, overlapping regions between POE values result in variations in darkness of the shaded regions. The order of features from left to right corresponds to Ordinal Fisher Scores of these features from highest to lowest. Figure 16 shows that the bridge experienced a higher level of shaking for one event (Calexico Earthquake on April 4, 2010). However, this event in 2010 was not strong enough to cause any damage.

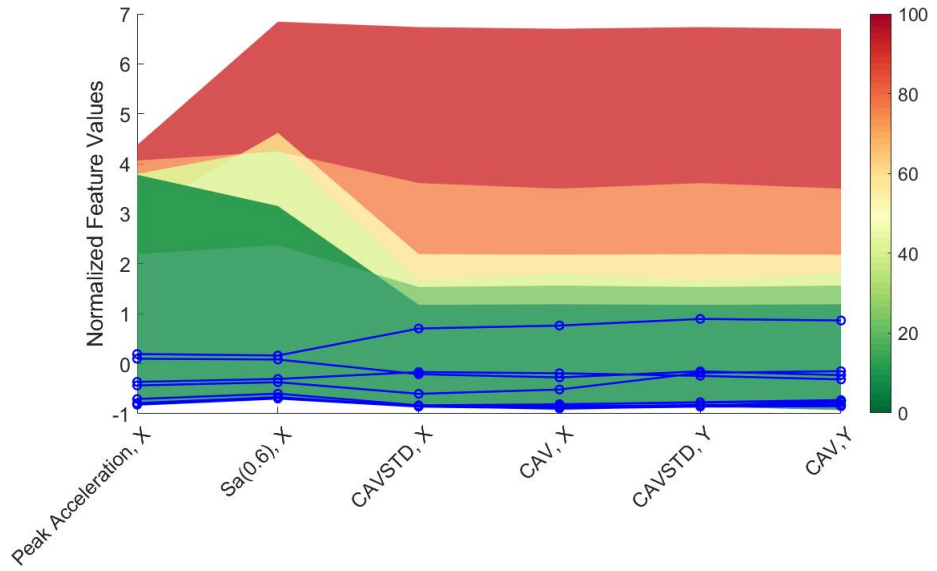


Figure 16: Normalized features and POE envelope of the Meloland Overpass Bridge.

Parkfield - Hwy46/Cholame Creek Bridge

The Cholame Creek Bridge located near Parkfield, California is a slightly-skewed RC slab bridge with five spans and supported by piles. The bridge spans Cholame Creek and is

130 ft long and 44 ft wide, widened in 1979 from its original width of 33 ft. The west end of the bridge is supported by a monolithic diaphragm abutment and the east end is supported on a seat abutment at the original portion and a diaphragm abutment at the widened portion. The bridge is instrumented with six sensors, and Figure 17 shows the sensor locations. The average feature values of the sensors from 2 recorded earthquakes are utilized in this study. In particular, one of these recorded earthquakes, the 2004 M_w[8] 6.0 Parkfield Earthquake, induced a peak structural horizontal acceleration of 1.047g measured by the bridge sensors, and signs of minor damage were observed during inspection after the event. With only 2 recorded events, it was not possible to develop the novelty model using recorded data. In this case, the novelty model for Bridge-A was utilized. The two recorded events are used as the test set. In order to develop the POE envelope, the period reported in (Boardman et al., 2006) is used to develop the SDOF model.

Figure 18 shows the actual and predicted tSNE plots for the 2 events. It is noted that the novelty model labelled both events as novelty. However, the H-MC method correctly detected the undamaged and damaged cases. Figure 19 compares the events with the POE envelope. It shows that San-Simeon earthquake (December 22, 2003) produced low levels of shaking for which the bridge falls in the green (undamaged) zone. For the Parkfield earthquake (September 28, 2004), however, the first two features fall in the red (damaged) zone. The damage assessment algorithm indicated the damage level to be DS1 in this case, consistent with the observed minor damage upon inspection following the 2004 earthquake.

Discussion and Conclusion

This study establishes an H-MC framework for bridge structural health monitoring. The framework uses a *novelty detection* ML model and a feature engineering procedure including feature selection using the *Ordinal Fisher Score* and model parameter tuning using *k-fold cross-validation*. First, the response of an analytical bridge model from nonlinear time history analysis with 400 ground motions is used to extract 32 one-dimensional damage features and damage states. The damage features are based on vibration characteristics with potential to indicate structural damage, and the damage states are based on levels of strain response according to a component capacity-based model. Next, an Ordinal Fisher Score feature selection analysis is used to determine the features with highest discriminatory power between classes of damage state. Using the seven highest-scoring features, k-fold cross-validation with $k = 5$ is performed with a POE analysis to determine the set of features which achieves optimal model complexity.

Following the feature engineering process, a ML model is trained using the selected feature set and the 400 ground motion responses to form the H-MC analysis, which is used to perform binary damage detection and multiclass damage assessment. Coupling the novelty detection model with a POE analysis is shown to increase average damage detection accuracy to 98.7%, an 8% increase from the novelty detection model alone. Furthermore, misclassification within the undamaged test set decreased by 34-58% when the POE analysis is applied together with the novelty detection. For multiclass damage assessment, the H-MC analysis achieved an

[8]The moment magnitude scale is a measure of an earthquake's magnitude ("size" or strength) based on its seismic moment (a measure of the work done by the earthquake).

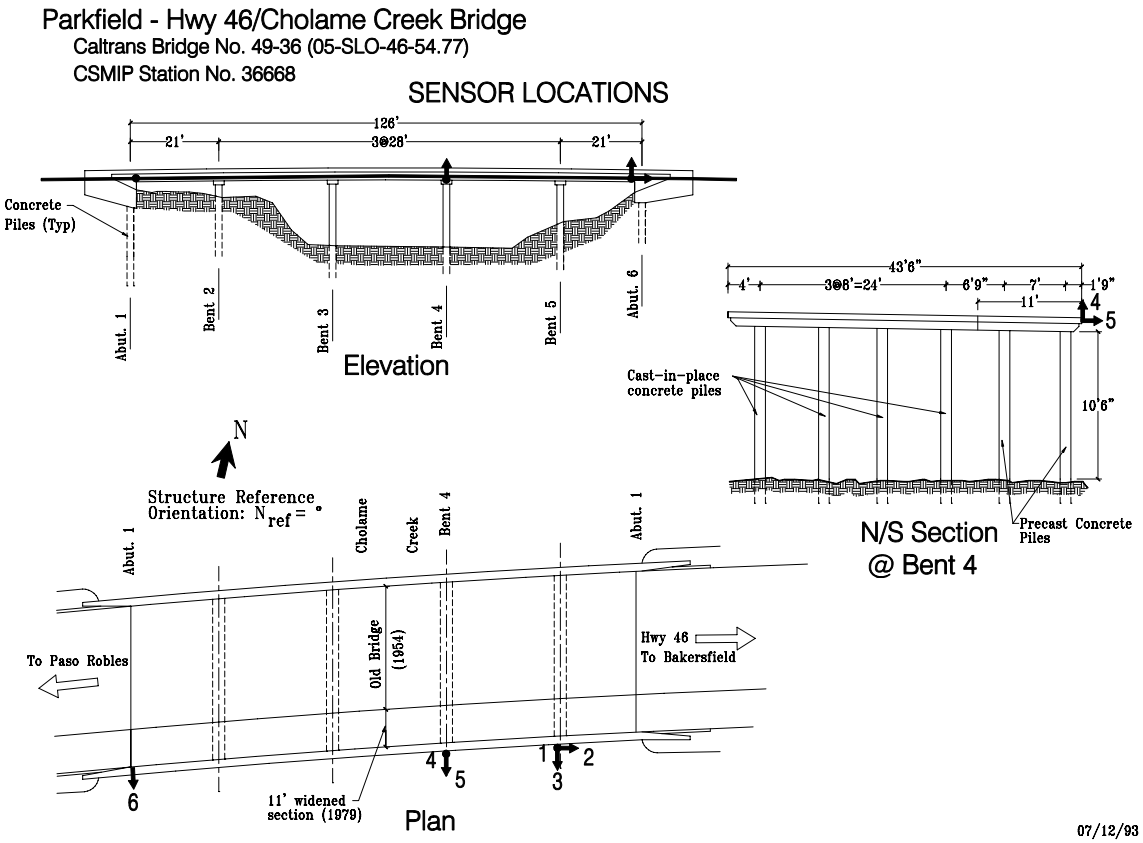
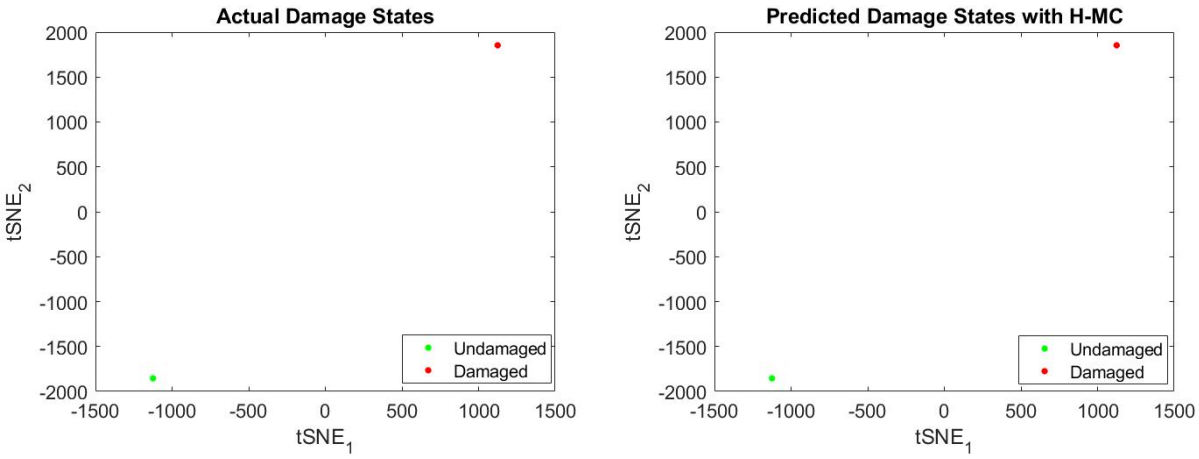


Figure 17: Sensor locations of the Cholame Creek Bridge.



(a) Actual damage states

(b) Predicted damage states

Figure 18: tSNE plots for the 2 events of the Parkfield Cholame Creek Bridge.

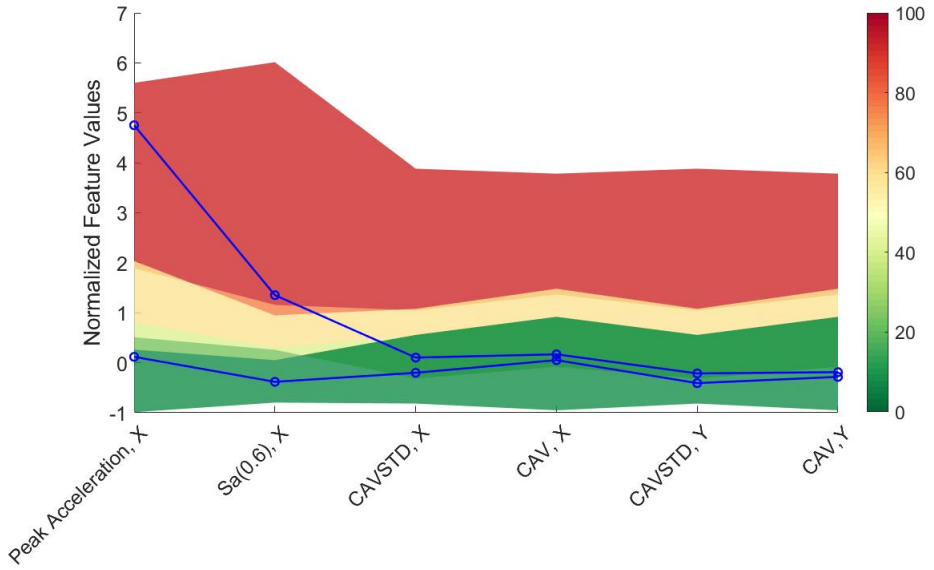


Figure 19: Normalized features and POE envelope of the Parkfield Cholame Creek Bridge.

average accuracy of 82%, which is judged to be reasonable and can be improved further in the future by using a more sophisticated analytical model than the SDOF for the POE analysis.

The application of the H-MC framework on two CSMIP[9]-instrumented bridges, namely El Centro-Hwy8/Meloland Overpass and Parkfield-Hwy46/Cholame Creek bridge, resulted in accurate classification of all fourteen undamaged events as undamaged and the one damaged event as damaged. In addition, the one damaged event was classified as DS1, which is consistent with the reported minor cracking observed during visual inspection after the event. Overall, the proposed framework will facilitate a rapid and efficient decision-making process regarding emergency response and immediate use/closure of bridges. These lifelines are essential transportation infrastructure that should remain functioning for community resiliency, before, during and after major earthquake events.

Acknowledgements

This research is supported by the California Department of Conservation, California Geological Survey, Strong Motion Instrumentation Program agreement 1019-016. The second author thanks the US NSF for the financial support provided by the NSF's Graduate Research Fellowship Program. The authors appreciate Dr. S. Günay's valuable input during this study.

References

ANCHETA, T. D., R. B. DARRAGH, J. P. STEWART, E. SEYHAN, W. J. SILVA, B. S.-J. CHIOU, K. E. WOODDELL, R. W. GRAVES, A. R. KOTTKE, D. M. BOORE ET AL., "NGA-West2 database," *Earthquake Spectra* 30 (2014), 989–1005.

[9]California Strong Motion Instrumentation Program.

BACCIANELLA, S., A. ESULI AND F. SEBASTIANI, "Feature selection for ordinal text classification," *Neural computation* 26 (2014), 557–591.

BOARDMAN, B. T., A. V. SANCHEZ, G. MARTIN, Z. ZAFIR, E. RINNE AND J. TOGNOLI, "Seismic Response of the Hwy 46/Cholame Creek Bridge During the 2004 Parkfield Earthquake," in *Fifth National Seismic Conference on Bridges & HighwaysMultidisciplinary Center for Earthquake Engineering ResearchCalifornia Department of TransportationFederal Highway AdministrationTransportation Research Board*A23 (2006).

BOZORGNIA, Y., N. A. ABRAHAMSON, L. A. ATIK, T. D. ANCHETA, G. M. ATKINSON, J. W. BAKER, A. BALTAY, D. M. BOORE, K. W. CAMPBELL, B. S.-J. CHIOU ET AL., "NGA-West2 research project," *Earthquake Spectra* 30 (2014), 973–987.

CAMPBELL, K. AND Y. BOZORGNIA, "Use of cumulative absolute velocity (CAV) in damage assessment," in *Proc. of 15th World Conf. in Earthquake Eng* (2012), 1–10.

CAMPBELL, K. W. AND Y. BOZORGNIA, "NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10 s," *Earthquake Spectra* 24 (2008), 139–171.

FARRAR, C. R. AND K. WORDEN, *Structural health monitoring: a machine learning perspective* (John Wiley & Sons, 2012).

GU, Q., Z. LI AND J. HAN, "Generalized fisher score for feature selection," *arXiv preprint arXiv:1202.3725* (2012).

JAYARAM, N., T. LIN AND J. W. BAKER, "A computationally efficient ground-motion selection algorithm for matching a target response spectrum mean and variance," *Earthquake Spectra* 27 (2011), 797–815.

KAVIANI, P., F. ZAREIAN AND E. TACIROGLU, "Seismic behavior of reinforced concrete bridges with skew-angled seat-type abutments," *Engineering Structures* 45 (2012), 137–150.

KENAWY, M., S. KUNNATH, S. KOLWANKAR AND A. KANVINDE, "Concrete Uniaxial Nonlocal Damage-Plasticity Model for Simulating Post-Peak Response of Reinforced Concrete Beam-Columns under Cyclic Loading," *Journal of Structural Engineering* 146 (2020), 04020052.

LALLEMANT, D., A. KIREMIDJIAN AND H. BURTON, "Statistical procedures for developing earthquake damage fragility curves," *Earthquake Engineering & Structural Dynamics* 44 (2015), 1373–1389.

LIANG, X. AND K. M. MOSALAM, "Ground motion selection and modification evaluation for highway bridges subjected to Bi-directional horizontal excitation," *Soil Dynamics and Earthquake Engineering* 130 (2020), 105994.

LIANG, X., K. M. MOSALAM AND S. GÜNAY, "Direct integration algorithms for efficient nonlinear seismic response of reinforced concrete highway bridges," *Journal of Bridge Engineering* 21 (2016), 04016041.

MAHALANOBIS, P. C., "On the generalized distance in statistics," (National Institute of Science of India, 1936).

MUIN, S. AND K. M. MOSALAM, "Localized Damage Detection of CSMIP Instrumented Buildings using Cumulative Absolute Velocity: A Machine Learning Approach," in *SMIP18 Seminar on Utilization of Strong-Motion Data, Sacramento, California* (2018).

NATIONAL RESEARCH COUNCIL, *Intelligent Human-machine Collaboration: Summary of a Workshop* (National Academies Press, 2012).

NOH, H. Y., D. LALLEMANT AND A. S. KIREMIDJIAN, "Development of empirical and analytical fragility functions using kernel smoothing methods," *Earthquake Engineering & Structural Dynamics* 44 (2015), 1163–1180.

NOH, H. Y., D. G. LIGNOS, K. K. NAIR AND A. S. KIREMIDJIAN, "Development of fragility functions as a damage classification/prediction method for steel moment-resisting frames using a wavelet-based damage sensitive feature," *Earthquake Engineering & Structural Dynamics* 41 (2012), 681–696.

ÖZDEMIR, H., "Nonlinear transient dynamic analysis of yielding structures [Ph.D. Thesis]," (1976).

PEER, *PEER NGA West2 Database* (2020), available at `https://ngawest2.berkeley.edu/`, accessed February 3, 2020.

PÉREZ-ORTIZ, M., M. TORRES-JIMÉNEZ, P. A. GUTIÉRREZ, J. SÁNCHEZ-MONEDERO AND C. HERVÁS-MARTÍNEZ, "Fisher Score-Based Feature Selection for Ordinal Classification: A Social Survey on Subjective Well-Being," in *International Conference on Hybrid Artificial Intelligence Systems* (Springer, 2016), 597–608.

SAJEDI, S. O. AND X. LIANG, "Vibration-based semantic damage segmentation for large-scale structural health monitoring," *Computer-Aided Civil and Infrastructure Engineering* (2019).

SAMUEL, A. L., "Some Studies in Machine Learning Using the Game of Checkers," *IBM Journal of Research and Development* 3 (1959), 211–229.

SHAMSABADI, A., T. OSTROM AND E. TACIROGLU, "Three dimensional global nonlinear time history analyses of instrumented bridges to validate current bridge seismic design procedures," in *Seminar on Utilization of Strong-Motion Data, Los Angeles, CA* (2011).

ZENGIN, E. AND N. A. ABRAHAMSON, "A vector-valued intensity measure for near-fault ground motions," *Earthquake Engineering & Structural Dynamics* 49 (2020), 716–734.