Vibration based bridge scour evaluation: A data-driven method using support vector machines

Zhiming Zhang[1a], Chao Sun[*1], Changbin Li[2b] and Mingxuan Sun[3c]

[1]*Department of Civil and Environmental Engineering, Louisiana State University,*
3255 Patrick F. Taylor Hall, Baton Rouge, LA 70803, United States
[2]*Department of Computer Science, University of Texas at Dallas,*
800 W Campbell Rd, Richardson, TX 75080, United States
[3]*Division of Computer Science and Engineering, Louisiana State University,*
3325 Patrick F. Taylor Hall, Baton Rouge, LA 70803, United States

Abstract. Bridge scour is one of the predominant causes of bridge failure. Current climate deterioration leads to increase of flooding frequency and severity and thus poses a higher risk of bridge scour failure than before. Recent studies have explored extensively the vibration-based scour monitoring technique by analyzing the structural modal properties before and after damage. However, the state-of-art of this area lacks a systematic approach with sufficient robustness and credibility for practical decision making. This paper attempts to develop a data-driven methodology for bridge scour monitoring using support vector machines. This study extracts features from the bridge dynamic responses based on a generic sensitivity study on the bridge's modal properties and selects the features that are significantly contributive to bridge scour detection. Results indicate that the proposed data-driven method can quantify the bridge scour damage with satisfactory accuracy for most cases. This paper provides an alternative methodology for bridge scour evaluation using the machine learning method. It has the potential to be practically applied for bridge safety assessment in case that scour happens.

Keywords: bridge scour; modal properties; machine learning; feature extraction; feature selection; support vector machines

1. Introduction

Bridge scour is among the predominant causes of bridge collapse in the United States (Deng and Cai 2009, Melville and Coleman 2000, Azhari1a *et al.* 2015). Scour around the bridge foundation damnifies the bridge's structural stability and has the potential to fail the bridge pier abruptly without warning. In the past 30 years, scour has failed 600 bridges, including the Sava Bridge in Zagreb and the Malahide Viaduct, resulting in significant transportation operating interruption and economic losses (Shirole and Holt 1991).

To improve the structural capacity for resisting scour damage, hydraulic and structural countermeasures can be adopted during the design stage (Shirole and Holt 2009, May *et al.* 2002,

---

∗Corresponding author, Dr., E-mail: csun@lsu.edu

Bolduc *et al.* 2008). In addition, monitoring its evolution in real time and implementing remedy measures when necessary is an alternative method with more effectiveness and economic benefits (Briaud *et al.* (2011)). Visual inspection and instrumentalized scour depth recording is limited because of potential inaccessibility and the high cost of installation and maintenance. In comparison, a vibration-based method that captures the variation of bridge structural dynamic responses has received increasing attention in recent years (Bao *et al.* 2017). Scour removes the soil around the bridge foundation and thus weakens the boundary constraints of the structure, which leads to the variation of bridge's dynamic responses. Many studies investigated the feasibility of detecting the bridge scour presence via measuring bridge's dynamic responses.

Foti and Sabia (2010) explored bridge scour detection through comparing the bridge dynamic responses before and after retrofitting. The features being compared include the bridge girder modal frequencies and the covariance matrix of foundation system acceleration signals from the structural responses under live traffic loading. It was found that the modal frequencies vary slightly after the retrofitting (no larger than 7%), and thus cannot provide sufficient reference for scour damage decision. The covariance matrix of certain piers shows significant asymmetry after retrofitting, which is an indicator of scour existence, though it cannot be used for scour severity evaluation. Briaud *et al.* (2011) studied the effectiveness of various instruments for bridge scour monitoring through lab experiments, numerical simulations, and field testing on real bridges. The evaluation parameters include modal frequencies and root mean square ratio of accelerations from two different directions. Results show that the dynamic responses in the flow direction shows the highest sensitivity to the scour development. This study provides an alternative for the specification of bridge scour warning criteria. Elsaid (2012) and Elsaid and Seracino (2014) investigated scour detection using three vibration based features including mode shape curvature, flexibility-based deflection, and flexibility based curvature through numerical simulation and experimental tests. The results indicate that the structural vibration in the horizontal direction demonstrates sufficient variation when scour happens. Prendergast *et al.* (2013) studied the feasibility of using the first modal frequency to detect bridge scour and evaluate its severity. The experiments show significant frequency reduction of the first mode when scour happened. Ju (2013) studied the variation of bridge natural frequencies with the increase of scour depth using a finite element method considering the soil-fluid-structure interaction. The results show significant decrease of natural frequencies as the scour develops, especially when the scour level reaches below the pile cap. Kong and Cai (2016) analyzed the effect of bridge scour on bridge natural frequencies and vehicle responses considering the bridge-vehicle-wave-interaction. The authors provided an alternative to monitor the bridge scour progression with vehicle responses. Prendergast *et al.* (2016) numerically developed a method of bridge scour detection using the first modal frequency of the pier vibration excited by the passing vehicles on bridges. Sensitivity study shows that the proposed method applies for a wide range of soil conditions, road roughness levels, and signal qualities.

In summary, the literature on vibration-based bridge scour monitoring uses features extracted from collected responses on superstructure or substructure with sensors, mostly accelerometer, to detect and evaluate the scour progression. The extracted dynamic features include the natural frequency (Foti and Sabia 2010, Briaud *et al.* 2011, Prendergast *et al.* 2013, 2016, Bao *et al.* 2017, signal asymmetry Foti and Sabia (2010)), signal power ratio (Briaud *et al.* 2011), mode shape curvature (Elsaid 2012, Elsaid and Seracino 2014, Chen *et al.* 2018, 2019), flexibility-based deflection (Elsaid 2012, Elsaid and Seracino 2014), flexibility based curvature (Elsaid 2012, Elsaid and Seracino 2014), and vehicle responses (Kong and Cai 2016). Though using the same

features, not all the studies show significant variation of the features with scour progression. For example, the first modal frequency of bridge substructure's vibration shows little variation with the scour depth increase in (Bao *et al.* 2017) while varies significantly in (Prendergast *et al.* 2013). Moreover, most of the existing studies make decisions based on a parametric study of primary modal properties, which lacks robustness for real applications that incorporates considerable random factors influencing the results.

Data-based methods, based on statistical rather than physical properties of structures, generalize damage identification to a supervised learning problem in the statistical pattern recognition paradigm (Farra and Worden 2012, Huang and Nagarajaiah 2014, Li *et al.* 2018, Sen *et al.* 2019, Yao and Pakzad 2012, Yu *et al.* 2018). Damage localization is pursued by a classification algorithm that evaluates the difference in damage-sensitive features between the current structural state and the initial baseline state.

To the best of the authors' knowledge, there is little research published on bridge scour evaluation using a data-driven method. Therefore, this paper attempts to develop a data driven method of bridge scour monitoring using support vector machines (SVM). Structural response of the bridge substructure is modeled using the finite element method. In the simulated dynamic test, the bridge model is excited by an impact load on the bridge pier and the structural responses at specified locations are collected. Then an output-only method is used to extract the structural modal information. Damage-sensitive features are extracted via comparing these modal properties, which are then selected and used for SVM classifier training and testing. The testing results indicate that the trained classification model can evaluate the bridge scour level with satisfactory accuracy regardless of the constraint type at the pile end. The data-based method possesses improved robustness and reliability through considering a large number of scour cases and significant noise disturbance that simulates the uncertain factors including structural variability, signal collection quality, and environmental disturbance. The rest of this paper first describes the numerical model used in this study. A generic sensitivity study is conducted prior to feature extraction and selection. Using the selected scour-sensitive features, the data-based scour detection is conducted. Finally, this paper presents conclusions and recommendations.

2. Numerical model

The present study uses the continuous bridge in (Lv *et al.* 2015) as an example for the bridge scour analysis as shown in Fig. 1. The bridge has a span layout of 65 m+65 m+65 m+40 m. The bridge girder has a height of 13.69 m and a width of 32.67 m. The pier in the middle and the corresponding piles are selected for analysis in this study. The middle pier is 20 m in height and 2.5 m in diameter. Each pile has a height of 34 m and a diameter of 1.8m. The bridge piles span across six soil layers from its top to bottom. This study applies a linear model of the soil properties. Table 1 lists the depth and mechanical properties of each soil layer at the bridge site.

The present study simulates the bridge structure and the effects of soil-structure interaction using the commercial finite element program, ANSYS. Fig. 2 shows the schematic of the numerical model simulating the middle part of the bridge in (Lv *et al.* 2015). A unit length of the girder across the middle pier is abstracted to resemble the constraint from the bridge girder onto the substructure. The four piles below the pier are represented using a single column with equivalent stiffness in the numerical model. Due to the requirement of large data quantity in data-driven method, this study uses beam element for the numerical model to reduce

computational cost without loss of generality regarding the structural behavior and the boundary conditions. The soil effect is simulated with spring and dashpot (COMBIN 14 element in ANSYS), as shown in Fig. 2. Each element simulates the mechanical effects of the soil of 1m in depth. This numerical model ignores the effect of soil mass and fluid-structure interaction which has been proved to have negligible influence on the structural dynamic behavior (Prendergast *et al.* 2013), Ju 2013). As to the boundary condition, this study simplifies the deck support with a vertical displacement constraint and fixes the end of the soil elements. The boundary condition at the pile end is assumed as fully fixed in this study.

Table 1 Soil components and their depth and mechanical properties

soil component	soil layer depth, m	specific weight , kN·m^{-3}	shear velocity , m·s^{-1}	undrained shear strength, kN/m	damping coefficient , kN/m s^{-1}	soil stiffness, kN/m
muck	4	16.5	100	4	0.90	305250
mucky soil	8	17.0	135	6	0.82	943500
silty clay 1	4	18.1	220	16	0.77	1387500
silty clay 2	4	19.2	270	35	0.61	1831500
coarse sand	8	18.0	410	-	0.10	3163500
silty clay 3	4	19.2	300	27	0.60	2275500

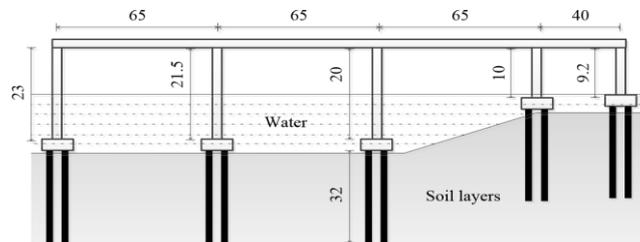

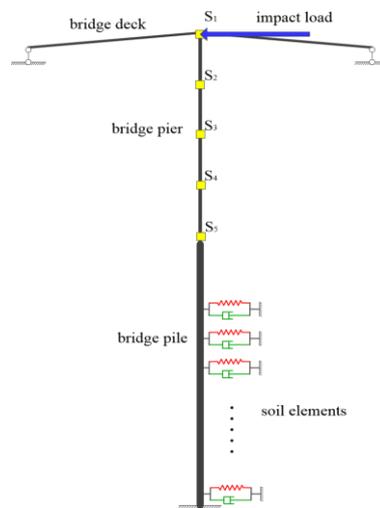Fig. 1 The layout of bridge [Unit: m, reproduced from (Lv *et al.* 2015)]

Fig. 2 Schematic model of bridge deck, pier, and pile

3. Feature extraction and selection

3.1 Sensitivity study

This section analyzes the sensitivity of bridge modal properties that are widely used for vibration-based scour monitoring. The dynamic features used in this sensitivity analysis include the natural frequency, the mode shape, and the mode shape curvature. These modal information are extracted with a blind source separation method (Yang and Nagarajaiah 2012), which calculates the modal information from only the structural displacement responses collected during an impact loading test. Fig. 2 shows the locations of the impact load and the sensors installed on the bridge pier (S_1 to S_5) for collecting the bridge dynamic responses.

This paper adds white noise to the calculated structural responses to simulate the effect of environmental and equipment noise disturbance in real applications. The white noise has a standard deviation equal to 5% of the response maximum. Considering the slight participation of higher modes in the structural response in real situations, this study uses the modal information of the first two modes for sensitivity analysis, feature extraction, and bridge scour evaluation. The scour progression is simulated by decreasing the real constant corresponding to stiffness (i.e., K) of the COMBIN elements in the numerical model. In this sensitivity study, the spring stiffness is decreased gradually by 50% from the soil top to its bottom. Once a spring's stiffness is decreased to 0, it indicates a 100% failure of the corresponding soil layer.

Fig. 3 compares the displacement responses collected from the five sensors with gradual development of the scour. For the sake of clear demonstration, the responses are not disturbed by white noise for this comparison. Fig. 3 shows that as scour develops from the top soil layer to the bottom layer, the bridge pier vibration is gradually impeded, signaling that that the scour development reduces the overall stiffness of the bridge structure and elongates its vibration period. In addition, the response magnitudes are increased as the scour develops.

Fig. 4 shows the modal frequency change with the scour development. It can be seen that as the scour progresses, both the first and second modes experience significant frequency reduction, especially when scour happens in the first four soil layers. However, the relative change of modal frequencies between different scour levels are not remarkable. For example, complete erosion of the muck layer causes a frequency reduction of no more than 7.2% for the first mode and 6.5% for the second mode. Moreover, when the scour develops further to the layer of coarse sand, both the first two modes experience little frequency variation. This indicates that only the frequency variation of the first two modes might not be sufficient for accurate scour identification considering uncertainty in real situation. Fig. 5 shows the variation of the mode shapes at different scour levels. For fair comparison, the first element of the mode shape vector is normalized to be unity. As scour progresses, the first mode shape exhibits clear difference; the second mode shape, though does not show a consistent change with different scour levels, still variates significantly. Fig. 6 shows the mode shape curvature of the first two modes presented in Fig. 5. In comparison with frequencies and mode shapes, the mode shape curvatures experience more significant variations.

The sensitivity analysis above shows that the modal properties of bridge pier experience significant variation as scour develops. In addition to modal frequency, the mode shape and mode shape curvature provide valuable information for determining the scour level. Furthermore, the mode frequencies, especially the 1st mode frequency, denoted as predominant natural frequency (PNF) in (Briaud *et al.* 2010, Bao *et al.* 2017), don't suffice the bridge scour detection and

quantification, as inconsiderable relative variation of frequencies are observed at a lower scour level. However, the authors expect satisfactory scour monitoring performance when using a data-driven method with substantial features consisting of modal frequencies, mode shapes, and mode shape curvatures reflecting the structural variation with scour progression and sufficiently large datasets. This will be explored in the rest of this paper.
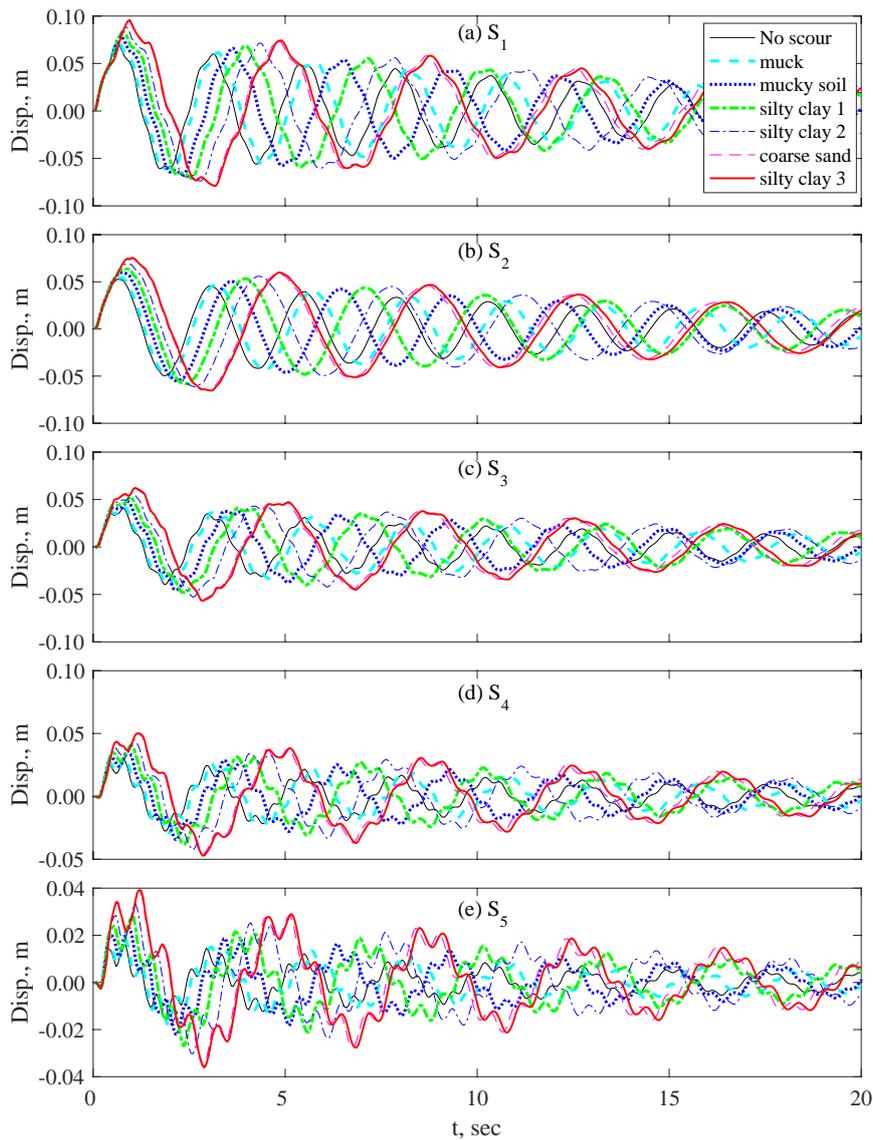

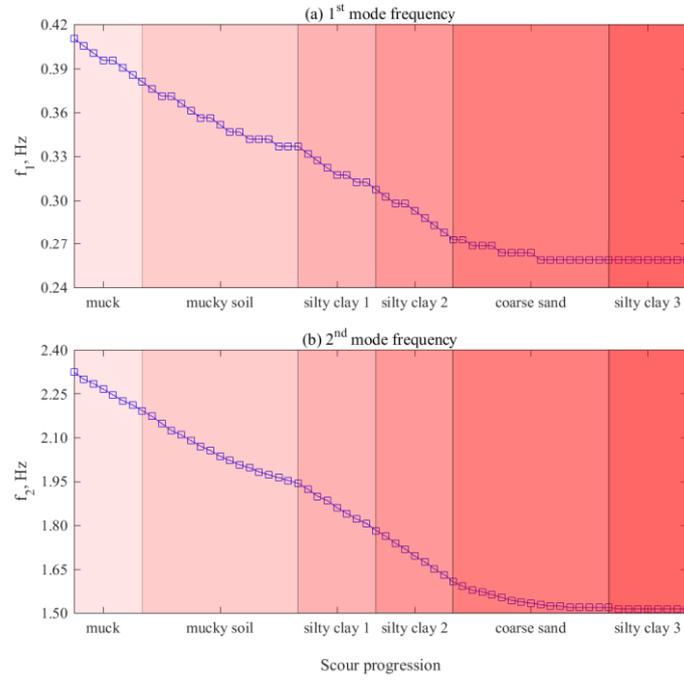Fig. 3 Comparison of bridge responses under different scour conditions

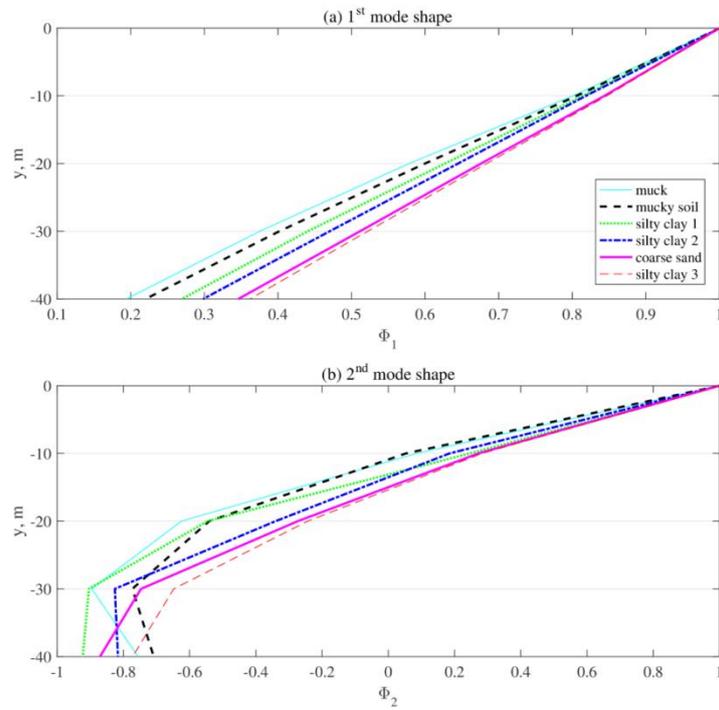Fig. 4 Frequency change with bridge scour development

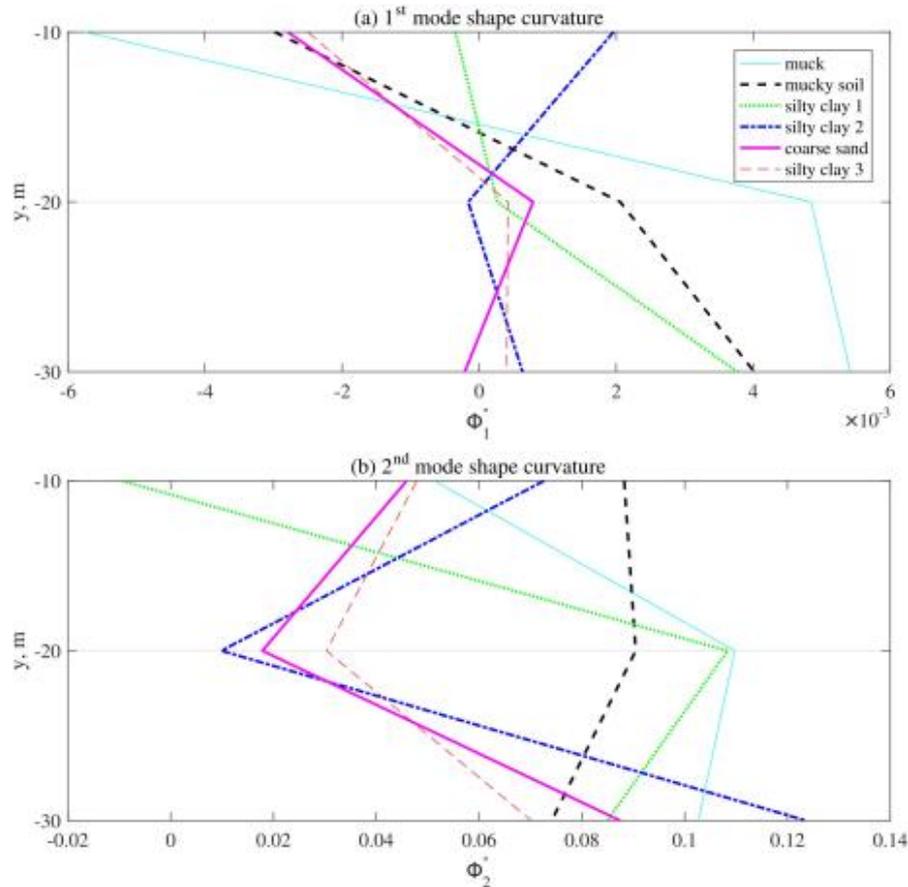Fig. 5 Mode shape change with bridge scour development

Fig. 6 Mode curvature change with bridge scour development

3.2 Feature extraction

For efficient structural damage detection and evaluation through pattern recognition using machine learning methods, the extraction of damage-sensitive features that are capable of reflecting the damage location and severity is of critical importance. Extracting and selecting features that can distinguish damaged structure and that of varied damage severities is one primary component of structural health monitoring (Doebling *et al.* 1996, Nagarajaiah and Erazo 2016, Nagarajaiah and Chen 2016, Sen *et al.* 2017). Indicative features of structural damage lead to accurate and efficient damage diagnosis, while irrelevant features results in misleading judgment on damage locations and severities. The following subsections present a review on some features extracted from structural dynamic properties that are widely used for structural health monitoring.

3.2.1 Natural frequency features
Adams *et al.* initiated damage detection using frequency variations in 1970s (Adams *et al.* 1978), and it was extended by Crawly and Adams (1979). The global nature of natural frequency makes it not an superior candidate for damage localization (Lee and Chung 2000, Kim *et al.* 2003,

Casas and Aparicio 1994). Nevertheless, it has been verified that the frequency change ratio of two modes caused by damage is not a function of the damage extent but its occurrence location (Cawley and Adams 1979). In addition, the normalized frequency change ratio (NFCR) is also proved to be related to the damage location and independent from the damage severity (Kaminski 1995). The fractional frequency change (FFC) for the ith mode is defined as

$$\text{FFC}_i = \frac{f_{\text{u}i} - f_{\text{d}i}}{f_{\text{u}i}} \tag{1}$$

where $f_{\text{u}i}$ and $f_{\text{d}i}$ are the ith mode frequencies of the structure in intact and damaged states, respectively.

The NFCR for the ith mode is defined a

$$\text{NFCR}_i = \frac{\text{FFC}_i}{\sum\limits_{j=1}^{N} \text{FFC}_j} \tag{2}$$

3.2.2 Mode shape features

It is worth noting that all the frequency based features cannot discriminate between two symmetric damage locations. In comparison, mode shapes contain structural spatial information that is valuable for damage localization (Rytter 1993) and are less susceptible to environmental influence (Farrar and James 1997).

3.2.2.1 DSI (Damage Signature Index)

A combined damage signature index (**DSI**) consisting of both mode shapes and frequencies were proposed in (Lam *et al.* 1998). It was found that the index **DSI** is dependent on damage location only. The **DSI** is defined as the deformation of mode shape vector normalized by the modal frequency square variation, that is,

$$\mathbf{DSI}_i = \frac{\mathbf{\Phi}_{\text{u}i} - \mathbf{\Phi}_{\text{d}i}}{|f_{\text{u}i}^2 - f_{\text{d}i}^2|} \tag{3}$$

where $\mathbf{\Phi}_{\text{u}i}$ and $\mathbf{\Phi}_{\text{d}i}$ are the modal shape vectors of the ith mode in undamaged and damaged conditions, respectively.

3.2.2.2 SDSI (Simplified Damage Signature Index)

Instead of a scalar parameter, **DSI** is series of vectors, with its quantity equal to the number of selected modes and its dimension equal to the number of degree of freedoms (DOFs) for mode shape extraction. For simplicity, the authors of this paper simplify this mode shape feature to an index that averages the contribution of each mode for a certain DOF, following the practice of mode shape curvature feature that will be introduced later. This simplification generates the simplified damage signature index (**SDSI**) as

$$\mathbf{SDSI} = \frac{1}{N} \sum_{i=1}^{N} |\mathbf{DSI}_i| = \frac{1}{N} \sum_{i=1}^{N} \frac{|\mathbf{\Phi}_{\text{u}i} - \mathbf{\Phi}_{\text{d}i}|}{|f_{\text{u}i}^2 - f_{\text{d}i}^2|} \tag{4}$$

where N is the number of selected modes for damage detection.

3.2.2.3 MAC (Modal Assurance Criterion)

Modal Assurance Criterion (**MAC**), proposed by Allemang and Brown (1982) detects mode shape shifts taking advantage of the orthogonality of mode shapes. **MAC** for the ith mode shape is defined as

$$MAC_i = \frac{|\mathbf{\Phi}_{ui}^T \mathbf{\Phi}_{di}|^2}{(\mathbf{\Phi}_{ui}^T \mathbf{\Phi}_{ui})(\mathbf{\Phi}_{di}^T \mathbf{\Phi}_{di})}$$

(5)

where $\mathbf{\Phi}_{ui}$ is the ith mode shape of the undamaged structure and $\mathbf{\Phi}_{di}$ is that of the damaged structure. **MAC** is a metric that measures the correlation level of mode shapes of a structure before and after damage. The value of **MAC** ranges between 0 and 1. A value of 1 indicates 100% vector correlation and no obvious mode shape variation after damage occurrence; a value of 0 denotes no correlation, and it means significant mode shifts due to the damage occurrence. Zhao and Zhang proved that **MAC**s of certain modes are sensitive to structural damage using an example of truss bridge (Zhao and Zhang 2012).

3.2.2.4 COMAC (Coordinate Modal Assurance Criterion)

Kim *et al.* (1992) extended **MAC** and developed the coordinate modal assurance criterion (**COMAC**) that identifies the coordinates where the mode shapes do not agree before and after structural damage occurrence.

$$COMAC_i = \frac{[\sum_{j=1}^{N} \mathbf{\Phi}_{u,j}^i \mathbf{\Phi}_{d,j}^i]^2}{\sum_{j=1}^{N}(\mathbf{\Phi}_{u,j}^i)^2 \sum_{j=1}^{N}(\mathbf{\Phi}_{d,j}^i)^2}$$

(6)

where the subscript i stands for the ith DOF of mode shape, j the mode order, and N the total number of modes taken into consideration.

3.2.3 Mode shape curvature features

Pandey *et al.* (1991) for the first time proposed to use the modal curvature for damage detection. The modal curvature of the jth mode at DOF i can be calculated using Eq. (7).

$$\mathbf{\Phi}_j^{"i} = \frac{\mathbf{\Phi}_j^{i+1} - 2\mathbf{\Phi}_j^i + \mathbf{\Phi}_j^{i-1}}{h^2}$$

(7)

where $\mathbf{\Phi}_j^i$ is the jth mode shape amplitude at the ith DOF; h is the distance between two successive measured locations.

Wahab and Guido (1999) used the modal curvature to detect bridge damage and obtained promising results. However, modal curvature works well for lower mode shape while higher modal curvature might produce false damage indication. Therefore, features extracted from mode shape curvatures should be used with caution.

3.2.3.1 CDF (Curvature Damage Factor)

Wahab and Roeck proposed the curvature damage factor that summarizes the contribution of curvature from all the modes (Wahab and Guido 1999), as

Table 2 Results of feature extraction

Feature ID	Feature Content
1	NFCR of 1st mode
2	NFCR of 2nd mode
3-7	SDSI at S_1 to S_5
8	MAC of 1st mode
9	MAC of 2nd mode
10-14	COMAC at S_1 to S_5
15-17	CDF at S_2 to S_4

$$\mathbf{CDF} = \frac{1}{N}\left\{\sum_{j=1}^{N}\left|\mathbf{\Phi}_{uj}^{"} - \mathbf{\Phi}_{dj}^{"}\right|\right\}$$

(8)

where N is the total number of modes considered in the analysis, $\mathbf{\Phi}_{uj}^{"}$ is the jth mode shape curvature of the intact structure, and $\mathbf{\Phi}_{dj}^{"}$ is that of the damaged structure.

Table 2 lists the result of feature extraction using the modal properties. The present study uses the first two modal information for feature extraction considering the difficulty of obtaining higher-frequency modes in practice. For the features from mode shape and mode shape curvatures, the five locations with sensors on the bridge pier from the top of pier to its bottom are used for modal information extraction. As a result, COMAC has a dimension of 5 and CDF 3. It can be expected that not all the feature listed in Table 2 are significantly contributive to bridge scour detection. Feature selection in Section 3.3 will eliminate the features with little contribution while preserve those with considerable importance.

3.3 Feature selection

Feature selection is the process of ranking and selecting contributive features from the feature candidates that are prepared during feature extraction (Qin *et al.* 2015). Feature selection produces a vector of features that have more discriminative and less redundant information. Feature selection also refines the feature extraction step by crediting effective features while disparaging features without significant positive contribution. Feature selection is necessary to improve the computational efficiency and the pattern recognition performance using machine learning methods. Additionally, the outcome of feature selection provides in-depth insights into the evaluation system by marking the features as contributive or not.

3.3.1 L1-based feature selection

Features can be selected through evaluating the classification quality using a certain model with selected features and eliminating those features considered unimportant (Ng (2004)). L1-based feature selection method adds *Lasso* penalty to the loss function during classification forcing features with weak contributions to obtain zero coefficients. Therefore, L1 regularization fulfills feature selection through producing sparse solutions that excludes features with little significance. This paper uses L1 regularized support vector machines for L1-based feature selection.

3.3.2 Recursive feature elimination

The recursive feature elimination method evaluates the performance of a smaller group of features recursively and selects the smallest feature group giving the best learning results. The size of feature groups is reduced through pruning the features with smallest weights after each round of estimator training. It also performs cross-validation to finalize the optimal features.

3.3.3 Tree-based feature selection

This selection method computes the feature importance using tree based estimators and discards insignificant features. For example, random forests implements feature selection using two straightforward methods: (1) mean decrease impurity and (2) mean decrease accuracy. The first method ranks the features by how much each of them decreases the weighted impurity in a decision tree and excludes the ones with lowest contribution; the second method measures the influence of permuting each feature on the classification or regression accuracy and removes those with little effect from permutation.

3.3.4 Univariate feature selection

Univariate feature selection selects the best features using the results of univariate statistical tests. This method individually evaluates each feature's informativeness and ranks the features accordingly. It finally applies cross-validation to determine the number of features to include from the top of the ranked feature list. Univariate feature selection method is often used for data preprocessing before applying machine learning methods (Lai *et al.* 2006).

3.3.5 Variance-based feature selection

This method removes those features with a variance below a certain threshold, as the variance of data along a dimension is a reflection of its representative power (He *et al.* 2006). As variance is a range-dependent statistic, normalization is a prerequisite before using this method.

Fig. 7 shows the results of feature selection using the five methods introduced above. The labels of *y*-axis, "L1", "Recursive", "Tree", "Univariate", and "Variance", denote L1-based feature selection, recursive feature elimination, tree-based feature selection, univariate feature selection, and variance-based feature selection, respectively. "Negative" and "Positive" on *z*-axis indicate the decision of feature selection using a certain method on a specific feature. For example, "L1" method determines that NFCR of the first mode has limited contribution to the scour level prediction, while "Tree" method yields positive decision. For a certain feature, the difference in the *x* direction indicates that the decision on its contribution using the five methods can be somewhat divergent. The present study screens out those features labeled as negative by two or more methods and finally collects those commonly selected features across all three constraint conditions, which guarantees the generality and robust effectiveness of the features used for bridge scour detection. The finally selected features are listed in Table 3.

To have a brief view of their distribution and representativeness of different bridge scour levels, this study reduces the selected features by projecting the features to two dimensions with the largest and second largest variance using principal component analysis. Fig. 8 illustrates the distribution of the first two principal components. It shows that the instances belonging to the same scour level gather together forming one to several clusters. The distribution does not show significant overlapping between clusters of different scour levels, which evidences the distinguishing capability of the selected features for scour level identification.
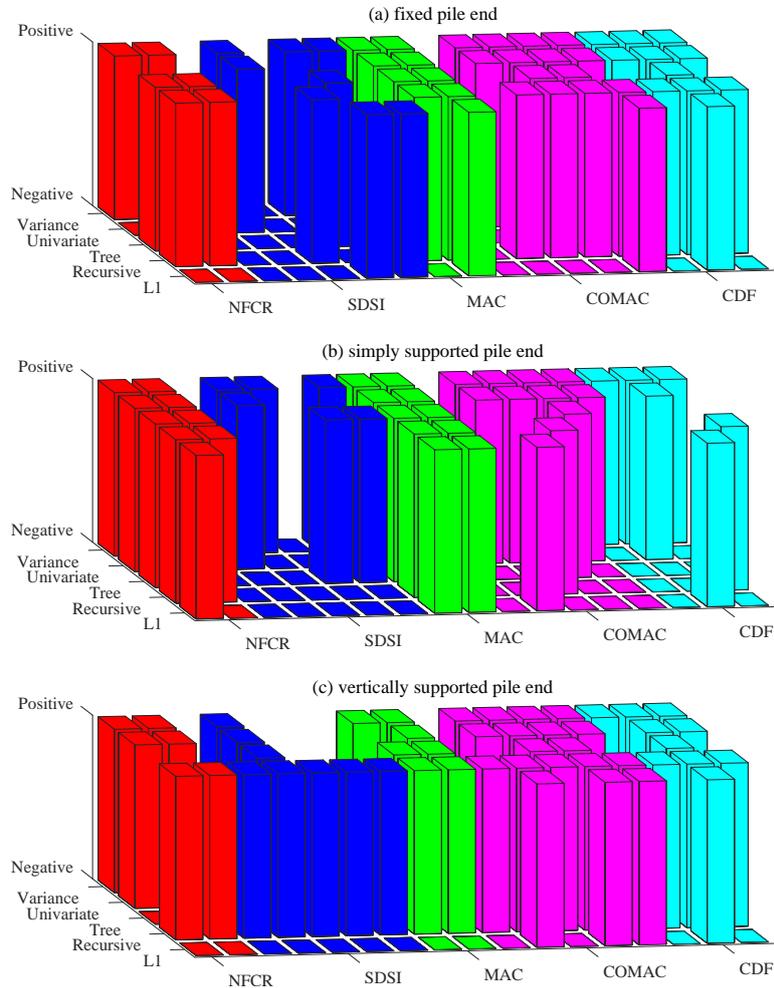
Fig. 7 Feature selection results

Table 3 Results of feature selection

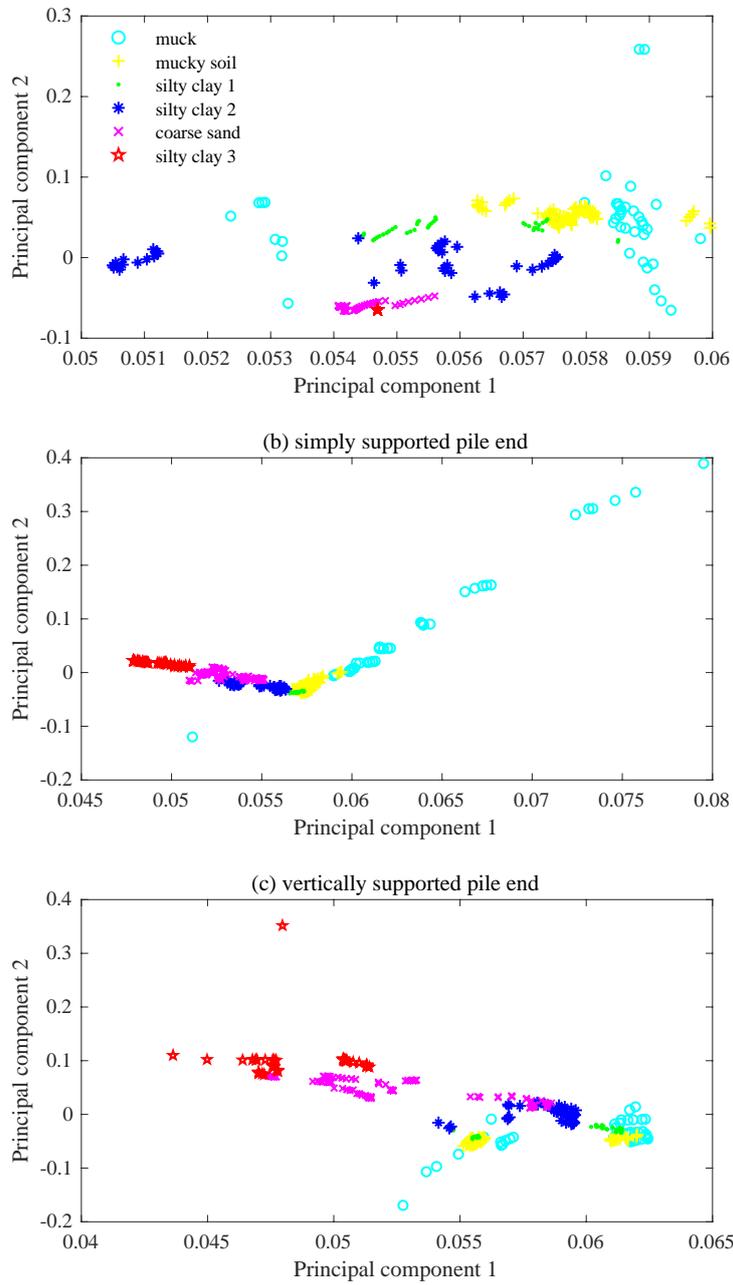| Selected Feature ID	Feature Content
1	NFCR of 1st mode
2	NFCR of 2nd mode
8	MAC of 1st mode
9	MAC of 2nd mode
11	COMAC at S_2
12	COMAC at S_3
13	COMAC at S_4
16	CDF at S_3

Fig. 8 Scatter plot of reduced features

4. Bridge scour detection

This paper uses a widely used machine learning method, support vector machines, to evaluate the bridge scour severity through supervised classification using the finally selected features listed in Table 3. As described above, the present study gradually fails the soil layers represented by the COMBIN element in the numerical model to simulate the continuous scour progress. Differing from the sensitivity study in which the step size is set as 50% stiffness reduction for each soil spring, this section adopts a step size of 1% as the soil failure increment to generate sufficient data for the machine learning classification, ending up with a total number of 3200 samples for six soil layers. It is worth noting that all the features are extracted from the dynamic responses with certain noise disturbance that simulates the random influence in practical implementation. The six scour stages from the muck layer to the layer of silty clay 3 are labeled as Class I-VI, respectively.

This study applies a ten-fold cross validation with 80% data for training, 10% for validation, and 10% for testing within each fold to evaluate the classification performance of the models, so that the hyper-parameters of the classifier can be well tuned. The classification quality is evaluated on the blind test data that is not used for training or hyper-parameter tuning using three metrics, including precision, recall, and f-score that will be introduced as follows. In a classification problem, the precision for a certain category is defined as the ratio between the number of true positives and the sum of quantity of true positives and false positives. True positives are the instances that are correctly classified into that category, while false positives are the instances that are wrongly labeled as belonging to that category. Recall is also a ratio where the numerator is the number of true positives and the denominator is the total of instances that indeed belong to the corresponding category.

Eqs. (9) and (10) express the formula of precision and recall, respectively (Powers 2011), in which N_{tp}, N_{fp}, and N_{fn} denote the number of true positives, the number of false positives, and the number of false negatives, respectively. A precision of 1.0 for a category indicates that all the instances that are classified into that category are indeed instances of that class, while a recall of 1.0 means that all the instances in that category have been classified correctly.

F-score is a measure of classification quality that combines the concepts of precision and recall (Powers 2011). As can be seen from Eq. (11), the value of f-score equals 1.0 only when both precision and recall have values of 1.0.

$$\text{Precision} = \frac{N_{tp}}{N_{tp} + N_{fp}} \tag{9}$$

$$\text{Recall} = \frac{N_{tp}}{N_{tp} + N_{fn}} \tag{10}$$

$$\text{F-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{11}$$

4.1 Introduction of support vector machines

Support vector machines (SVM) are popular methods for classification, regression, and novelty detection (Christopher 2016). In binary classification, using the selected kernel method that maps the input data to a n-dimensional space, SVM aims to construct a *n*-1 hyperplane that separates the two classes of data with the maximum margin. New samples are projected to the same space and

classified to one of the classes based on the partition of the hyperplane learned from training samples (Saha *et al.* 2009). The radial basis function kernel, or RBF kernel, is the most widely used kernel function used in SVM, because it has localized and finite responses along the entire axis. The RBF kernel on two feature vectors $\mathbf{x}$ and $\mathbf{x}'$ is defined as

$$K(\mathbf{x},\mathbf{x}') = \exp\left(-\frac{\|\mathbf{x}-\mathbf{x}'\|^2}{2\sigma^2}\right)$$
(12)

in which $\|\mathbf{x}-\mathbf{x}'\|$ can be the Euclidean distance between $\mathbf{x}$ and $\mathbf{x}'$, and σ is a free parameter. Compared with SVM with an RBF kernel, SVM with a linear kernel is a parametric model and has reduced complexity and training cost especially when the data size is large. Other kernel functions include polynomial kernel, sigmoid kernel, etc.

Let $X = \mathbf{x}_1,\mathbf{x}_1,...,\mathbf{x}_L$ and $Y = \{y_1,y_2,...,y_L\}$ denote a group of patterns and the corresponding labels, in which $\mathbf{x}_j \in \mathrm{R}^d$, d is the dimension of feature vectors, and $y_j \in 1,2,...,M$ denotes the class label associated with $\mathbf{x}_j$. Then for the classification problem $F_M = (X,Y)$, support vector machines determines the separating boundary between classes that minimizes the classification errors. Specifically, the parameters of the hyperplane can be represented as a weight vector $\mathbf{w}$ and a bias b, which can be learned through the constrained optimization as follows (Cortes and Cortes 1995)

$$\min_{\mathbf{w}\in\mathrm{R}^d,b\in\mathrm{R}} \quad \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_j^L \xi_j$$

$$\text{s.t.} \quad y_j(\mathbf{w}^T\mathbf{x}_j + b) \geq 1 - \xi_j, \quad \xi_j \geq 0, \quad \forall j = 1,...L$$
(13)

where ξ_j is the slack variable that allows misclassification in case of non-separable data, C is a parameter that controls the trade-off between the misclassification rate of training examples and the complexity of the decision boundary, and b is the bias parameter. In this way, SVM transforms the problem of model parameter learning to a quadratic optimization problem.

One of the valuable features of SVM is that the model parameter determination is a convex optimization problem, which means that the global optimum can be reached by any local solution. Additionally, SVM has advantages over other classifiers regarding generalization, data quantity requirement, robustness against the error of models, and computational efficiency. SVM has been widely used in structural health monitoring and damage identification (Ni *et al.* 2005, Oh and Sohn 2009, Worden and Lane 2001, Li and Burgue 2010).

4.2 Results and discussions

This section presents the results of bridge scour identification using SVM classifiers with the selected features. One important parameter of the SVM classifier is the type of kernel function that projects the data into the corresponding space where the samples can be linearly separable to the largest extent. This study tries two types of kernel functions, the linear kernel and the Gaussian radial basis kernel, and optimizes their corresponding parameters using a grid search strategy. Additionally, the ten-fold cross-validation ensures the generality of the trained model and avoids the over-fitting problem.
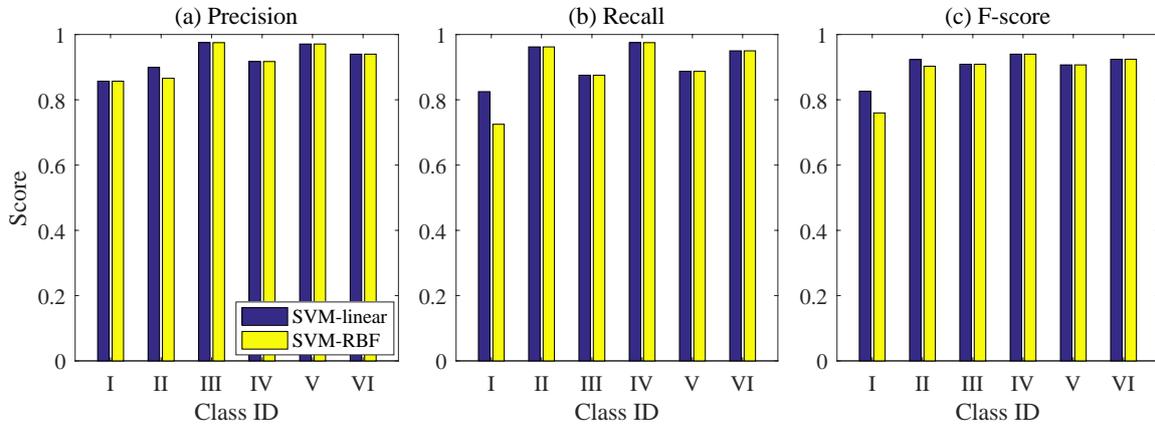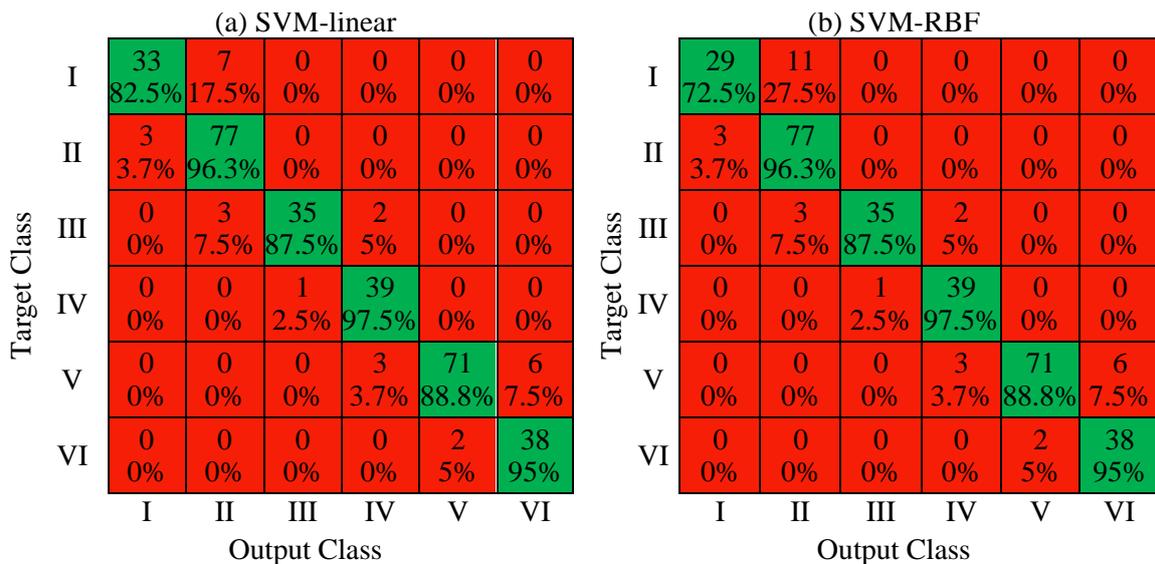
Fig. 9 Classification scores

Fig. 10 Confusion matrices

For bridge scour detection, the kernel function selection for SVM does not yields significant difference in the classification performances. For example, projecting the training data using the Gaussian radial basis kernel function does not remarkably improve their linear separability. Therefore, the linear kernel function is sufficient for the bridge scour identification using SVM classifiers. The discussion below are based on the results of SVM classifiers with linear kernels. Fig. 9 compares the classification scores of parametrically optimized SVM classifiers. As shown in Fig. 9, the SVM classifier yields good performance when distinguishing the scour levels with most evaluation scores above 0.8. For example, all f-scores above 0.9 except for Class I which corresponds to the scour in the top layer.

Additionally, Fig. 10 shows the mean confusion matrices of SVM classification. Differing from the overall classification metrics such as f-score, confusion matrices present more information about the classification for each class. In the confusion matrix, the diagonal elements denote the number of correctly labeled test samples on average, while off-diagonal elements denote the number of samples that are classified incorrectly. Taking Class I in Fig. 10(a) as an example, among the samples belonging to Class I, 82.5% are recognized successfully with 17.5% of them tagged with the wrong label (Class II); on the other hand, 7.5% samples labeled as Class I while in in fact they are samples of Class II. After examining the figures, one can find that all the incorrectly classified samples are recognized as the adjacent classes of the target class. This result relieves the consequence of misclassification when identifying bridge scour levels, as incorrect identification does not significantly under-estimate or overestimate the scour severity.

In summary, the support vector machines provide an effective machine learning approach to identify the bridge scour level using the selected features extracted from structural modal properties. The classifiers yields evaluating scores above 0.85 for most cases with all the cases labeled incorrectly categorized into its adjacent classes.

5. Conclusions

This paper studies the effectiveness of a machine learning method, support vector machines, for vibration based bridge scour prediction. A simplified numerical model is developed to simulate the structural response of the bridge substructure. The displacement responses of the bridge pier are collected and used for extracting the structural modal information in an output-only manner. In practical implementation, bridge displacements can be measured using vision-based sensors in a noncontact manner. Feature extraction and selection derives from the modal properties the most efficacious physical attributes for quantifying the bridge scour progression. The accuracy of scour prediction validates the effective representativeness of the selected features as well as the distinguishing capability of support vector machines when used for supervision pattern recognition. Future work will focus on identifying scour development using the structural dynamic responses under service load, such as traffic loading, or environmental effects like wind loading. Field validation of the established method will also be included in the future study. In reality, the acceleration data are usually more accessible than the displacement data. Then the operational modal analysis (OMA) can be conducted to obtain the structural modal properties and the corresponding modal-based features. Candidate OMA methods include the frequency domain decomposition (FDD) and stochastic subspace identification (SSI) methods. On the other hand, considering the limited data that can be collected from a certain bridge with various scour levels, an accurate finite element model should be necessary to generate sufficient amount of reliable numerical data, so that the trained machine leaning model has satisfactory classification accuracy and generality. The finite element model can be obtained from finite element model updating. In addition, other classifiers will be applied for bridge scour evaluation and their performance will be evaluated and compared with that of SVM.

Acknowledgments

References

Adams, R.D., Cawley, P., Pye, C.J. and Stone, B.J. (1978), "A vibration technique for non-destructively assessing the integrity of structures", *J. Mech. Eng. Sci.*, **20**(2), 93-100.

Allemang, R.J. and Brown, D.L. (1982), "A correlation coefficient for modal vector analysis", *Proceedings of the 1st international modal analysis conference*, Orlando: Union College Press.

Azhari, F., Scheel, P.J. and Loh, K.J. (2015), "Monitoring bridge scour using dissolved oxygen probes", *Struct. Monit. Maint.*, **2**(2), 145-164.

Bao, T., Swartz, R.A., Vitton, S., Sun, Y., Zhang, C. and Liu, Z. (2017), "Critical insights for advanced bridge scour detection using the natural frequency", *J. Sound Vib.*, **386**, 116-133.

Barthorpe, R.J., Manson, G. and Worden, K. (2017), "On multi-site damage identification using single-site training data", *J. Sound Vib.*, **409**, 43-64.

Bishop, C.M. (2016), Pattern Recognition and Machine Learning. Springer-Verlag New York.

Bolduc, L.C., Gardoni, P. and Briaud, J.L. (2008), "Probability of exceedance estimates for scour depth around bridge piers", *J. Geotech. Geoenviron. Eng.*, **134**(2), 175-184.

Briaud, J.L., Hurlebaus, S., Chang, K.A., Yao, C., Sharma, H., Yu, O.Y., Darby, C., Hunt, B.E. and Price, G.R. (2011), "Realtime monitoring of bridge scour using remote monitoring technology", Austin, TX.

Briaud, J.L., Yao, C., Darby, C., Sharma, H., Hurlebaus, S., Price, G.R., Chang, K.A., Hunt, B.E. and Yu. O.Y. (2010), "Motion sensors for scour monitoring: laboratory experiments and numerical simulations", *Transportation Research Board (TRB) 89th Annual Meeting*, 10-14.

Casas, J.R. and Aparicio, A.C. (1994), "Structural damage identification from dynamic-test data", *J. Struct. Eng.*, **120**(8), 2437-2450.

Cawley, P. and Adams, R.D. (1979), "The location of defects in structures from measurements of natural frequencies", *J. Strain Anal. Eng. Des.*, **14**(2), 49-57.

Chen, Y., Joffre, D. and Avitabile, P. (2018), "Underwater dynamic response at limited points expanded to full-field strain response", *J. Vib. Acoust.*, **140**(5), 051016.

Chen, Y., Logan, P., Avitabile, P. and Dodson, J. (2019), "Non-model based expansion from limited points to an augmented set of points using Chebyshev polynomials", *Exp. Techniques*, 1-23.

Cortes, C. and Cortes, V. (1995), "Support-vector networks", *Machine learning*, **20**(3), 273-297.

David Martin Powers (2011), Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation.

Deng, L. and Cai, C.S. (2009), "Bridge scour: Prediction, modeling, monitoring, and countermeasures", *Practice periodical on structural design and construction*, **15**(2), 125-134.

Doebling, S.W., Farrar, C.R., Prime, M.B. and Shevitz, D.W. (1996), Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review.

Elsaid, A. (2012), "Vibration based damage detection of scour in coastal bridges", PhD thesis, North Carolina State University.

Elsaid, A. and Seracino, R. (2014), "Rapid assessment of foundation scour using the dynamic features of bridge superstructure", *Constr. Build. Mater.*, **50**, 42-49.

Farrar, C.R. and James III, G.H. (1997), "System identification from ambient vibration measurements on a bridge", *J. Sound Vib.*, **205**(1), 1-18.

Farrar, C.R. and Worden, K. (2012), *Structural health monitoring: a machine learning perspective*, John Wiley & Sons.

Foti, S. and Sabia, D. (2010), "Influence of foundation scour on the dynamic response of an existing bridge", *J. Bridge Eng.*, **16**(2), 295-304.

He, X., Cai, D. and Niyogi, P. (2006), "Laplacian score for feature selection", *Adv. Neural Inform. Process. Syst.*, 507-514.

Huang, C. and Nagarajaiah, S. "Experimental study on bridge structural health monitoring using blind source separation method: arch bridge", *Struct. Monit. Maint.*, **1**(1), 69-87.

Ju, S.H. (2013), "Determination of scoured bridge natural frequencies with soil–structure interaction", *Soil Dynam. Earthq.Eng.*, **55**, 247-254.

Kaminski, P.C. (1995), "The approximate location of damage through the analysis of natural frequencies with artificial neural networks", *Proceedings of the Institution of Mechanical Engineers, Part E: J. Process Mech. Eng.*, **209**(2), 117-123.

Kim, J.H., Jeon, H.S. and Lee, C.W. (1992), "Applications of the modal assurance criteria for detecting and locating structural faults", *Proceedings of the International Modal Analysis Conference*, pages 536-536. SEM Society for Experimental Mechanics Inc.

Kim, J.T., Ryu, Y.S., Cho, H.M. and Stubbs, N. (2003), "Damage identification in beamtype structures: frequency-based method vs mode-shape-based method", *Eng. Struct.*, **25**(1), 57-67.

Kong, S. and Cai, S.C. (2016), "Scour effect on bridge and vehicle responses under bridge–vehicle–wave interaction", *J. Bridge Eng.*, **21**(4), 04015083.

Lai, C., Reinders, M.J.T. and Wessels, L. (2006), "Random subspace method for multivariate feature selection", *Pattern Recogn. Lett.*, **27**(10), 1067-1076.

Lam, H.F., Ko, J.M. and Wong, C.W. (1998), "Localization of damaged structural connections based on experimental modal and sensitivity analysis", *J. Sound Vib.*, **210**(1), 91-115.

Lee, Y.S. and Chung, M.J. (2000), "A study on crack detection using eigenfrequency test data", *Comput. Struct.*, **77**(3), 327-342.

Li, L., Liu, G., Zhang, L. and Li, Q. (2018), "Deep learning-based sensor fault detection using s-long short term memory networks", *Struct. Monit. Maint.*, **5**(1), 51-65.

Li, Z. and Burgue, R. (2010), "Using soft computing to analyze inspection results for bridge evaluation and management", *J. Bridge Eng.*, **15**(4), 430-438.

Lv, Y., Liu, Z. and Li, Z. (2015), "Seismic response analysis for unequal height pier bridges considering soil-pile interaction", *J. Vib. Shock*, **35**(23), 114.

May, R.W.P., Ackers, J.C. and Kirby, A.M. (2002), Manual on scour at bridges and other hydraulic structures, **551**. Ciria London.

Melville, B.W. and Coleman, S.E. (2000), Bridge scour, Water Resources Publication.

Nagarajaiah, S. and Chen, B. (2016), "Output only structural modal identification using matrix pencil method", *Struct. Monit. Maint.*, **3**(4), 395-406.

Ng, A.Y. (2004), "Feature selection, l 1 vs. l 2 regularization, and rotational invariance", *Proceedings of the 21st international conference on Machine learning*, ACM.

Ni, Y.Q., Hua, X.G., Fan, K.Q. and Ko, J.M. (2005), "Correlating modal properties with temperature using long-term monitoring data and support vector machine technique", *Eng. Struct.*, **27**(12), 1762-1773.

Oh, C.K. and Sohn. H. (2008), "Damage diagnosis under environmental and operational variations using unsupervised support vector machine", *J. Sound Vib.*, **325**(1-2), 224-239.

Pandey, A.K., Biswas, M. and Samman, M.M. (1991), "Damage detection from changes in curvature mode shapes", *J. Sound Vib.*, **145**(2), 321-332.

Prendergast, L.J., Hester, D. and Gavin, K. (2016), "Determining the presence of scour around bridge foundations using vehicle-induced vibrations", *J. Bridge Eng.*, **21**(10), 04016065.

Prendergast, L.J., Hester, D., Gavin, K. and OSullivan, J.J. (2013), "An investigation of the changes in the natural frequency of a pile affected by scour", *J. Sound Vib.*, **332**(25), 6685-6702.

Qin, Y., Dong, M., Zhao, F., Langari, R. and Gu, L. (2015), "Road profile classification for vehicle semi-active suspension system based on adaptive neuro-fuzzy inference system", *Proceedings of the Decision and Control (CDC), 2015 IEEE 54th Annual Conference on*.

Rytter., A. (1993), Vibrational based inspection of civil engineering structures, PhD thesis, Aalborg

University.

Saha, B., Goebel, K., Poll, S. and Christophersen, J. (2009), "Prognostics methods for battery health monitoring using a bayesian framework", *IEEE T. Instrum. Meas.*, **58**(2), 291-296.

Satish Nagarajaiah and Kalil Erazo. (2016), "Structural monitoring and identification of civil infrastructure in the United States", *Struct. Monit. Maint.*, **3**(1), 51-69.

Sen, D., AAghazadeh, A., Mousavi, A., Nagarajaiah, S. and Baraniuk, R. (2019), "Sparsitybased approaches for damage detection in plates", *Mech. Syst. Signal Pr.*, **117**, 333-346.

Sen, D., Nagarajaiah, S. and Gopalakrishnan, S. (2017), "Harnessing sparsity in lamb wave-based damage detection for beams", *Struct. Monit. Maint.*, **4**(4), 381-396.

Shirole, A.M. and Holt, R.C. (1991), "Planning for a comprehensive bridge safety assurance program", *Transportation Research Record*, **1290**, 39-50.

Shirole, A.M. and Holt, R.C. (2009), Monitoring Scour Critical Bridges-A NCHRP, Synthesis of highway, practice, Washington, DC.

Wahab, M.M.A. and Roeck, G.D. (1999), "Damage detection in bridges using modal curvatures: application to a real damage scenario", *J. Sound Vib.*, **226**(2), 217-235.

Worden, K. and Lane, A.J. (2001), "Damage identification using support vector machines", *Smart Mater. Struct.*, **10**(3), 540.

Yang, Y. and Nagarajaiah, S. (2012), "Time-frequency blind source separation using independent component analysis for output-only modal identification of highly damped structures", *J. Struct. Eng.*, **139**(10), 1780-1793.

Yao, R. and Pakzad, S.N. (2012), "Autoregressive statistical pattern recognition algorithms for damage detection in civil structures", *Mech. Syst. Signal Pr.*, **31**, 355-368.

Yu, H., Zhu, H.P., Weng, S., Gao, F., Luo, H. and Ai, De M. (2018), "Damage detection of subway tunnel lining through statistical pattern recognition", *Struct. Monit. Maint.*, **5**(2), 231-242.

Zhao, J. and Zhang, L. (2012), "Structural damage identification based on the modal data change", *Int. J. Eng. Manufact. (IJEM)*, **2**(4), 59.

SN