J Civil Struct Health Monit (2017) 7:153–162
DOI 10.1007/s13349-017-0222-y

ORIGINAL PAPER

Damage detection of a truss bridge utilizing a damage indicator from multivariate autoregressive model

Yoshinao Goi[1] · **Chul-Woo Kim**[1]

Received: 16 October 2016 / Revised: 23 March 2017 / Accepted: 12 April 2017 / Published online: 2 May 2017

Abstract This study proposes a damage indicator automatically derived from a set of multivariate autoregressive models estimated from ambient vibration of bridges. The damage indicator evaluates a stochastic distance between a set of reference data (healthy bridge data) and unknown test data. Statistical hypothesis testing based on a probability distribution of the damage indicator was applied for damage detection. A field experiment conducted on an actual steel truss bridge with truss members that were artificially severed was conducted to assess the damage detection efficacy of the proposed damage indicator. Experimentally obtained results showed that the proposed damage indicator enables detection of three damage patterns. The proposed damage indicator effectiveness was also assessed by comparison to the damage sensitive feature from a univariate autoregressive model using experimental data of the same bridge. This comparison also demonstrated the efficacy of the proposed damage indicator obtained from a multivariate linear system.

Keywords Damage indicator · Damage detection · Field application · Multivariate autoregressive model · Hypothesis testing

1 Introduction

Management of aging infrastructure is a crucially important issue confronting civil engineering professionals. To reduce the potential risk of structural failure as well as life cycle costs, an efficient inspection method is desired for preventive maintenance. Techniques of structural health monitoring (SHM) based on vibration measurements have been attracting bridge owners because of their efficient inspection processes [1–3]. Changes in structural integrity of bridges engender changes in their modal properties such as natural frequencies, damping ratios, and mode shapes that are identifiable from vibration data [4–8]. Consequently, vibration-based SHM is a useful technique if a bridge can be excited easily and effectively. Two commonly used means to excite bridges are forced vibration tests and ambient vibration tests. Bridges are excited using a shaker, impact hammer or other means in the forced vibration test. However, ambient vibration tests use various natural excitations such as wind and ground motion.

For bridge health monitoring, the ambient vibration test is regarded as much more convenient than the forced vibration test because natural excitations require no traffic control, whereas the forced vibration test requires traffic control during the test. However, in ambient vibration tests conducted on actual bridges, the identified modal properties are contaminated by unknown noises. In other words, deciding whether changes in modal properties are caused by damage or noise is difficult. To reduce noise and to improve identification of modal properties, this study was undertaken to develop a damage indicator (DI) that can be defined directly using a multivariate linear system model estimated from measured vibration data.

For existing studies using a coefficient of an autoregressive model as a damage-sensitive feature, Nair's damage-sensitive feature (NDSF), which consists of univariate autoregressive (AR) coefficients, has been investigated as a damage indicator for a model building [9]. A laboratory experiment conducted on a model bridge shows the feasibility of NDSF for SHM of bridges based on traffic-induced vibrations [10]. A damage detection method using the

✉ Chul-Woo Kim
kim.chulwoo.5u@kyoto-u.ac.jp

[1] Department of Civil and Earth Resources Engineering, Kyoto University, Kyoto, Japan

Mahalanobis distance (MD) [11, 12] was proposed to evaluate the statistical distance between groups of samples of NDSF. The validity of this damage detection method has also been verified from results of field experiments conducted on real steel truss bridges with artificial damage to its truss members [13]. It is noteworthy that the NDSF includes information about vibrations of each sensor, but information about correlation between sensors is not included because it is derived from the univariate AR model.

Chang and Kim [14] investigated the MD sensitivity to modal parameters attributable to damage. They concluded that considering all the modal parameters including modal frequency, damping ratio, and the modal assurance criterion from the mode vector demonstrated that the inclusion of more parameters in outlier analysis might engender more sensitive features. That is because any statistical change in those features would be reflected in the MD, which indicates that the mode vector, which is a kind of correlation between sensors, would provide useful information related to the bridge health condition. However, it is not straightforward to decide damage-sensitive modes as well as relevant mode vectors for healthy bridges. Therefore, aiming at a pattern recognition approach to estimate damage-sensitive modes or damage locations, a novel DI is derived directly from a multivariate linear system to include spatial information between sensors.

This study proposes a novel DI based on the MD of damage-sensitive features that are extracted directly from a multivariate autoregressive (MAR) model using principal component analysis (PCA) [15]. The DI evaluates stochastic changes in modal information contained in the MAR model, whereas the NSDF is merely based on univariate AR model. Therefore, the DI is expected to detect changes in modal properties related to correlation among measurement points and improve damage detection performance. This study also proposes a damage detection approach based on the hypothesis test to enable stochastically appropriate decision-making. Acceleration data from the field experiment in [14] are examined to validate the feasibility of the proposed method. In this experiment, truss members on an actual steel truss bridge were artificially severed. This study investigates the efficacy of the presented DI and hypothesis testing for three damage patterns considered in the experiment. The NDSF is also examined using the field experiment data and compared to the proposed DI.

2 Methodology

2.1 Damage detection flow

This study investigates an unsupervised change detection method for changes between a reference dataset obtained from a bridge under healthy condition and a newly observed dataset monitored from the same bridge under artificial damage. Based on machine learning approaches [16], the former dataset is designated as the "training set". The latter is the "test set". An outline of the calculation of the novel DI and damage detection is presented in Fig. 1. Detailed procedures for each process are as follows:

- As the first step, parameter vectors related to modal properties of the structure are estimated from measured acceleration data using the MAR model (see Sect. 2.2).
- The obtained parameter vectors are classified as a training set $\boldsymbol{C}$ and a test set $\boldsymbol{C}'$ (see Sect. 2.3).
- Damage-sensitive features and the corresponding subspace are extracted by application of PCA to the training set (see Sect. 2.3).
- The proposed DI is defined using MD between the training set and the test set with respect to the damage-sensitive features (see Sect. 2.4).
- Damage detection is formulated as a hypothesis test based on the probability distribution of DIs (see Sect. 2.5).

2.2 Linear system model

Let $\boldsymbol{y}(k)$ denote a column vector of the discrete time series of measured acceleration data for which the components, respectively, correspond to M measurement points (or M sensors). The time series obtained from a linear structural system excited by white noise can be modeled as a MAR model of sufficient model order P as [17, 18]

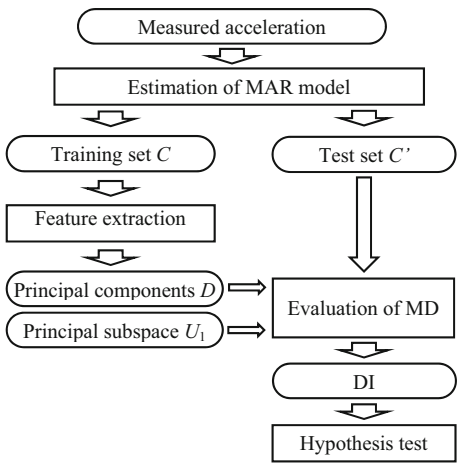

Fig. 1 Flow of the proposed damage detection

$$\boldsymbol{y}(k) = \sum_{i=1}^{P} \boldsymbol{A}_i \boldsymbol{y}(k-i) + \boldsymbol{e}(k), \tag{1}$$

where $\boldsymbol{A}_i$ denotes the ith AR coefficient matrix and $\boldsymbol{e}(k)$ denotes a white noise vector.

Letting $E[\cdot]$ represent expectation and letting $[\cdot]^{\mathrm{T}}$ stand for a transposed matrix or vector, the autocorrelation matrix of the measured data $\boldsymbol{y}(k)$ is defined as $\boldsymbol{R}(s) = E[\boldsymbol{y}(k)\ \boldsymbol{y}(k-s)^{\mathrm{T}}]$. It provides the Yule–Walker equation as follows:

$$\boldsymbol{R}(s) = \sum_{i=1}^{P} \boldsymbol{A}_i \boldsymbol{R}(-i) + \boldsymbol{W}(s=0) \tag{2}$$

$$\boldsymbol{R}(s) = \sum_{i=1}^{P} \boldsymbol{A}_i \boldsymbol{R}(s-i)\ \ (s = 1, 2, \ldots) \tag{3}$$

Therein, $\boldsymbol{W} = E[\boldsymbol{e}(k)\boldsymbol{e}(k)^{\mathrm{T}}]$.

From a measured time series of finite data length N, autocorrelation matrices are estimated as shown below:

$$\hat{\boldsymbol{R}}(s) = \frac{1}{N} \sum_{k=s}^{N} \boldsymbol{y}(k)\boldsymbol{y}(k-s)^{\mathrm{T}} (s = 0, 1, \ldots, P) \tag{4}$$

$$\hat{\boldsymbol{R}}(s) = \hat{\boldsymbol{R}}(-s)^{\mathrm{T}} (s < 0) \tag{5}$$

Therefore, Eq. 3 can be summarized as the form of following equations using autocorrelation matrices estimated from Eqs. 4 and 5:

$$[\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_P]\boldsymbol{T} = \left[\hat{\boldsymbol{R}}(1), \hat{\boldsymbol{R}}(2), \cdots, \hat{\boldsymbol{R}}(P)\right] \tag{6}$$

$$\boldsymbol{T} = \begin{bmatrix} \hat{\boldsymbol{R}}(0) & \hat{\boldsymbol{R}}(1) & \cdots & \hat{\boldsymbol{R}}(P-1) \\ \hat{\boldsymbol{R}}(-1) & \hat{\boldsymbol{R}}(0) & \cdots & \hat{\boldsymbol{R}}(P-2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\boldsymbol{R}}(1-P) & \hat{\boldsymbol{R}}(2-P) & \cdots & \hat{\boldsymbol{R}}(0) \end{bmatrix} \tag{7}$$

The AR coefficient matrix $\boldsymbol{A}_i$ can be estimated from an actual measurement time series by solving Eq. 6. Autocorrelation matrices defined in Eqs. 4 and 5 provide the positive definite matrix $\boldsymbol{T}$ for which the inverse matrix is assured to exist to solve Eq. 4.

Using the z-transform, Eq. 1 is transformed in z-domain as shown below [19]:

$$\boldsymbol{Y}(z) = \boldsymbol{H}(z)\boldsymbol{E}(z) \tag{8}$$

$$\boldsymbol{H}(z) = \left(\boldsymbol{I}_M - \sum_{i=1}^{P} z^{-i}\boldsymbol{A}_i\right)^{-1}, \tag{9}$$

where, $\boldsymbol{Y}(z)$ and $\boldsymbol{E}(z)$, respectively, denote z-transforms of $\boldsymbol{y}(k)$ and $\boldsymbol{e}(k)$. Also, $\boldsymbol{I}_M$ denotes the identity matrix of M-order. Matrix $\boldsymbol{H}(z)$ in Eqs. 8 and 9 is the transfer function of the linear system shown in Eq. 1. The conjugated pairs of the poles of the transfer function $\boldsymbol{H}(z)$ are related to the modal characteristics of the structure, as shown in Eq. 10.

$$\lambda_i, \lambda_i^* = \exp\left(\left(-\xi_i \pm j\sqrt{1-\xi_i^2}\right)\omega_i \Delta t\right) \tag{10}$$

In that equation, the following variables are used: ω_i denotes the natural angular frequency of the ith mode; ξ_i stands for the damping ratio of the ith mode; Δt signifies the sampling time of the discrete time series; and j represents an imaginary unit. These poles are obtainable by solving the eigenvalue problem with respect to z as

$$|\boldsymbol{I}_l z - \boldsymbol{A}| = 0 \tag{11}$$

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_{P-1} & \boldsymbol{A}_P \\ \boldsymbol{I}_M & 0 & \cdots & 0 & 0 \\ 0 & \boldsymbol{I}_M & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{I}_M & 0 \end{bmatrix}, \tag{12}$$

where $|\cdot|$ denotes determinant of a matrix and $\boldsymbol{I}_l$ denotes the identity matrix of l-order in which $l = MP$.

2.3 Feature extraction

The left-hand-side of Eq. 11 can be rearranged to the following polynomial with the appropriate polynomial coefficients c_i:

$$z^l + c_1 z^{l-1} + \cdots + c_{l-1} z^1 + c_l. \tag{13}$$

The polynomial of Eq. 13 produces its roots as shown in Eq. 10. Therefore, each of the polynomial coefficients is related to modal properties of the structure. In this study, a coefficient vector $\boldsymbol{c} = [c_1\ c_2 \cdots c_l]^{\mathrm{T}}$ is assumed to be a vector of random variables that follow a multivariate Gaussian distribution produced from an independent observation of the bridge vibration. These coefficients presumably change their statistical characteristics when damage on a structure causes any change in the modal properties. Therefore, the proposed damage detection method is formulated to detect changes in these coefficients.

To formulate the change detection method, the coefficient vectors $\boldsymbol{c}$ obtained from measurement data are classified into a training set and a test set. Let $\boldsymbol{c}_i$ ($i = 1, 2 \ldots, n$) and $\boldsymbol{c}'_i$ ($i = 1, 2 \ldots, n_t$), respectively, denote the training samples and the test samples of the coefficient vectors. Each set is described as a matrix comprising the sample vectors, i.e. the training set and the test set are described as follows:

$$C = [\boldsymbol{c}_1, \boldsymbol{c}_2, \cdots, \boldsymbol{c}_n] \tag{14}$$

$$C' = \left[\boldsymbol{c}'_1, \boldsymbol{c}'_2, \cdots, \boldsymbol{c}'_{nt}\right] \tag{15}$$

Damage-sensitive features are extracted using PCA with respect to the training set. The PCA is definable as the orthogonal projection of the data onto a lower dimensional linear space, known as the principal subspace, such that the

variance of the projected data is maximized [15]. Kim et al. [20] reported that the damage sensitivity of the NDSF can be improved by application of PCA to the estimated AR coefficients of a univariate AR model, so that the PCA is expected to extract damage-sensitive features from the training set. The orthonormal basis of the principal subspace is defined by the m eigenvectors corresponding to the m largest eigenvalues of the data covariance matrix. The PCA is performed using the singular value decomposition (SVD) to a matrix comprising deviation vectors of the training samples, as shown in Eq. 16:

$$\boldsymbol{C}_{\text{dev}} = [\tilde{\boldsymbol{c}}_1, \tilde{\boldsymbol{c}}_2, \cdots, \tilde{\boldsymbol{c}}_n] \tag{16}$$

In that equation, $\tilde{\boldsymbol{c}}_i$ represents the deviation vectors defined as Eq. 17:

$$\tilde{\boldsymbol{c}}_i = \boldsymbol{c}_i - \frac{1}{n}\sum_{k=1}^{n} \boldsymbol{c}_k (i = 1, 2, \ldots, n) \tag{17}$$

The basis of the principal subspace is provided as the sub-matrix $\boldsymbol{U}_1$ using the SVD:

$$\boldsymbol{C}_{\text{dev}} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\text{T}} = [\boldsymbol{U}_1\boldsymbol{U}_2]\boldsymbol{S}\boldsymbol{V}^{\text{T}} \tag{18}$$

Therein, $\boldsymbol{S}$ denotes the $l \times n$ diagonal matrix with diagonal entries that are non-negative real numbers listed in descending order. Also, $\boldsymbol{U}$ and $\boldsymbol{V}$ stand for the corresponding orthogonal matrices. $\boldsymbol{U}_1$ represents the sub-matrix which consists of the first m columns of $\boldsymbol{U}$, and $\boldsymbol{U}_2$ represents the sub-matrix excluded $\boldsymbol{U}_1$ from $\boldsymbol{U}$.

Using the basis of the principal subspace $\boldsymbol{U}_1$, the principal components of training set $\boldsymbol{C}$ are formulated as

$$\boldsymbol{D} = \boldsymbol{U}_1^{\text{T}}\boldsymbol{C}. \tag{19}$$

In that equation, each column vector included in $\boldsymbol{D}$ corresponds to principal components of the training data in $\boldsymbol{C}$.

2.4 Novel damage indicator

The MD is traditionally used to detect an anomalous sample referring to a set of normal samples [12]. Using a set of reference sample vectors $\boldsymbol{X} = [\boldsymbol{x}_1\ \boldsymbol{x}_2 \cdots\ \boldsymbol{x}_n]$ and another sample vector $\boldsymbol{x}'$, the MD in squared units is defined as

$$\text{MD}(\boldsymbol{x}', \boldsymbol{X}) = (\boldsymbol{x}' - \bar{\boldsymbol{X}})^{\text{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}' - \bar{\boldsymbol{X}}), \tag{20}$$

where $\bar{\boldsymbol{X}}$ and $\boldsymbol{\Sigma}$, respectively, signify the sample mean and the covariance matrix estimated from $\boldsymbol{X}$ as defined in Eqs. 21 and 22.

$$\bar{\boldsymbol{X}} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x}_i \tag{21}$$

$$\boldsymbol{\Sigma} = \frac{1}{n-1}\sum_{i=1}^{n} (\boldsymbol{x}_i - \bar{\boldsymbol{X}})(\boldsymbol{x}_i - \bar{\boldsymbol{X}})^{\text{T}} \tag{22}$$

Assuming that each of the samples $\boldsymbol{x}_1$, $\boldsymbol{x}_2$,…, $\boldsymbol{x}_n$ and $\boldsymbol{x}'$ are independent and identically distributed (i.i.d.) Gaussian random vectors and the number of the reference samples n is much larger than the length of the sample vectors m, MD is known to be approximately χ^2-distributed with m degrees of freedom. The probability density function (PDF) of the χ^2-distribution with v degrees of freedom is given as shown below:

$$\chi^2(x|v) = \frac{1}{2\Gamma(v/2)}\left(\frac{x}{2}\right)^{v/2-1}\exp\left(-\frac{x}{2}\right), \tag{23}$$

where $\Gamma(\cdot)$ denotes the gamma function. The expectation and the variance of this distribution are known, respectively, as v and $2 \times v$.

Using MD with a test sample of the coefficient vectors $\boldsymbol{c}_i'$ the novel DI proposed in this study is defined as

$$\text{DI}(\boldsymbol{c}_i') = \frac{1}{m}\text{MD}(\boldsymbol{U}_1^{\text{T}}\boldsymbol{c}_i', \boldsymbol{D}), \tag{24}$$

where $\boldsymbol{U}_1$ and $\boldsymbol{D}$, respectively, stand for the matrices defined in Eqs. 18 and 19.

2.5 Hypothesis test

Assuming that the test sample vectors included in $\boldsymbol{C}'$ and the training sample vectors included in $\boldsymbol{C}$ are i.i.d. Gaussian vectors and $m \ll n$, the MD in Eq. 24 follows X^2-distribution with m degrees of freedom as mentioned in Sect. 2.4; hence, the $m\text{DI}(\boldsymbol{c}'_i)$ is presumed to have the following distribution:

$$m\text{DI}(\boldsymbol{c}_i') \sim \chi^2(\boldsymbol{x}|m), \tag{25}$$

where $m\text{DI}(\boldsymbol{c}_i') = \text{MD}(\boldsymbol{U}_1^{\text{T}}\boldsymbol{c}_i', \boldsymbol{D})$ from Eq. 24 and $\sim$ denotes that the variable in the left-hand-side takes the distribution as the PDF in the right-hand-side.

Therefore, assuming that the monitored bridge is healthy, the expectation of the DI is presumed to be 1. For damage detection, it is reasonable to carry out the right-sided test with respect to the expectation of the DI because the DI shows the stochastic distance between a test sample and the training set, and anomalies in the test set will provide larger DIs. The null hypothesis H_0 and alternative hypothesis H_1 for the hypothesis test are formulated as follows:

$$\text{H}_0 : E\left[\text{DI}(\boldsymbol{c}_i')\right] = 1 \tag{26}$$

$$\text{H}_1 : E\left[\text{DI}(\boldsymbol{c}_i')\right] > 1 \tag{27}$$

As for test statistics, the sample mean of DI from the test set C' shown in Eq. 28 is examined.

$$\overline{\text{DI}} = \frac{1}{n_t}\sum_{i=1}^{n_t} \text{DI}(\boldsymbol{c}_i') \tag{28}$$

The reproductive property of the χ^2-distribution shows that the summation of the $m\mathrm{DI}(\boldsymbol{c}')$ also follows χ^2-distribution as shown in Eq. 29 because the PDF of the $m\mathrm{DI}(\boldsymbol{c}')$ follows the χ^2-distribution.

$$\sum_{i=1}^{n_t} m\mathrm{DI}(\boldsymbol{c}_i') \sim \chi^2(\boldsymbol{x}|mn_t) \tag{29}$$

Therefore, using significance level α for the hypothesis test, the rejection region of the sample mean is definable as shown below:

$$\overline{\mathrm{DI}} > \frac{1}{mn_t} F^{-1}(1-\alpha|mn_t), \tag{30}$$

where $F^{-1}(p|v)$ is the inverse function with respect to x of the following cumulative distribution function:

$$p = F(x|v) = \int_0^x \chi^2(t|v)\mathrm{d}t. \tag{31}$$

Equation 29 provides the variance of the sample mean inversely proportional to n_t as shown below:

$$E\left[\left(\overline{\mathrm{DI}}-1\right)^2\right] = \frac{2}{mn_t} \tag{32}$$

Equation 32 shows that providing more test samples will yield stricter critical values of the rejection region. The critical value of the rejection region given by Eq. 30 converges to expectation 1 as n_t goes to infinity. If a sufficient amount of data is available, this method could be useful for making a decision for damage or anomalies of the bridges. On the other hand, the stricter the critical values are, the easier it becomes to reject the null hypothesis even though bridges are healthy. For practical use, therefore, it is recommended to examine validity of the critical values by applying the proposed method to a data set from the bridge in healthy condition independently from the training set as presented in Sect. 3.

3 Field experiment

Field experiments were conducted with a moving vehicle on an actual bridge. The target bridge for the field experiment is a single lane simply supported by a through-type steel Warren truss bridge as presented in Fig. 2a, b. The bridge has 59.2 m span length, 8 m maximum height, and 3.6 m width. The vehicle used for the experiment is a two-axle recreation vehicle (Serena; Nissan Motor Co. Ltd.) with total weight of about 21 kN. During the experiment, all traffic except the load vehicle was prohibited.

Eight uniaxial accelerometers were installed on the deck of the bridge to measure vertical vibrations as presented in Fig. 2b. Two optical sensors were installed at the respective ends of the bridge to detect the vehicle entrance and exit. The sampling rate of each sensor was set as 200 Hz. Five scenarios were considered in this study, as shown in Table 1 and Fig. 3. Initially, The INT scenario represents the intact bridge with no damage. For the DMG1 scenario, a half-cut damage was applied to the vertical truss member at the midspan (see Fig. 2b and Fig. 3), and for the DMG2 scenario, a full-cut damage was applied to the same member. After examining DMG2, the damaged member was repaired, which is denoted as the RCV scenario. Finally, for DMG3 scenario, full cut was applied in a vertical member at 5/8th-span (see Fig. 2b) after examining the RCV scenario. Each experiment was conducted under the vehicle running at about 40 km/h. The bridge vibration under the passing vehicle was measured 10 times except DMG1. For DMG1, bridge vibrations were measured 12 times.

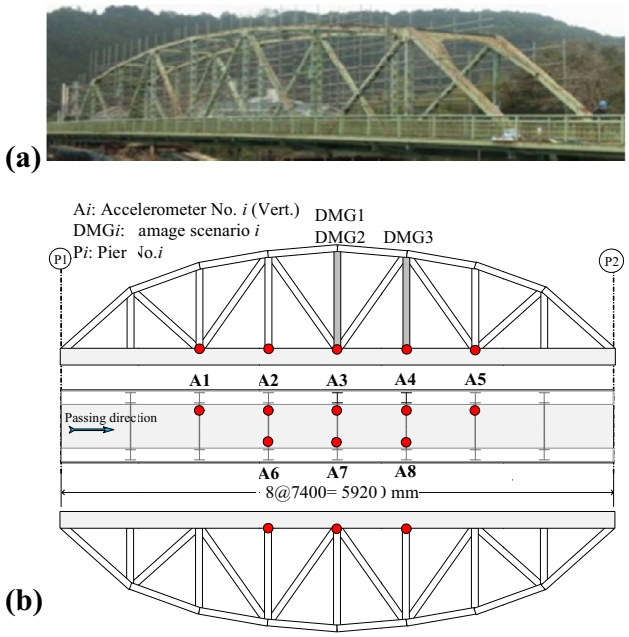

Fig. 2 Target bridge **a** photograph and **b** sensor deployment

Figure 4 presents the time series and PSD curves of sensor A3 of INT, DMG1, and DMG2 scenarios acquired

Table 1 Scenarios considered in the field experiment

Scenario	Description
INT	Intact bridge
DMG1	Half cut in vertical member at midspan
DMG2	Full cut in vertical member at midspan
RCV	Recovery of the cut member
DMG3	Full cut in vertical member at 5/8th-span

from the experiment. According to the record of the optical sensors (vertical dashed lines in Fig. 4), the time series is separable at the moment of the vehicle exit from the bridge. The time series after the exit of the vehicle can be interpreted as a free vibration data because no external loading significantly affected the bridge vibrations. Empirically, free vibration data are known to enable stable modal parameter identification because external loading can be a source of uncertainty for identification. However, the whole time series including excitation produced using a passing vehicle is still convenient for practical application because it requires no knowledge about the exit times of vehicles. Therefore, this study examines both free vibration (FV) data and whole vibration (WV) data.

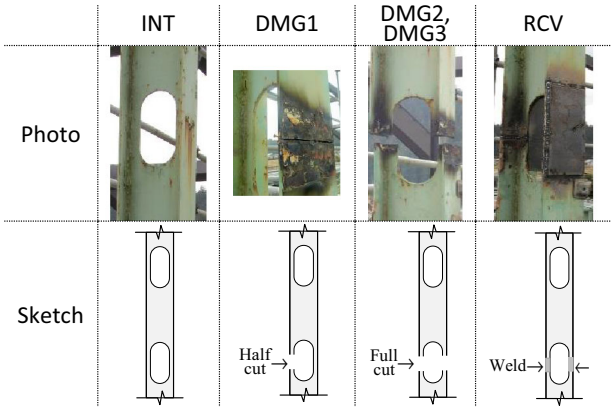

Fig. 3 Photographs and sketches of artificial damage

PSD curves in Fig. 4 demonstrated the difficulty of making a decision about the possibility of damage in the bridge from the Fourier spectra. The DI is, therefore, adopted for each scenario. The INT scenario was adopted as the training scenario for DMG1 and DMG2. For DMG3, the RCV scenario was adopted as the training scenario because the modal characteristics of the bridge might be changed after repairing the damaged member. Considering

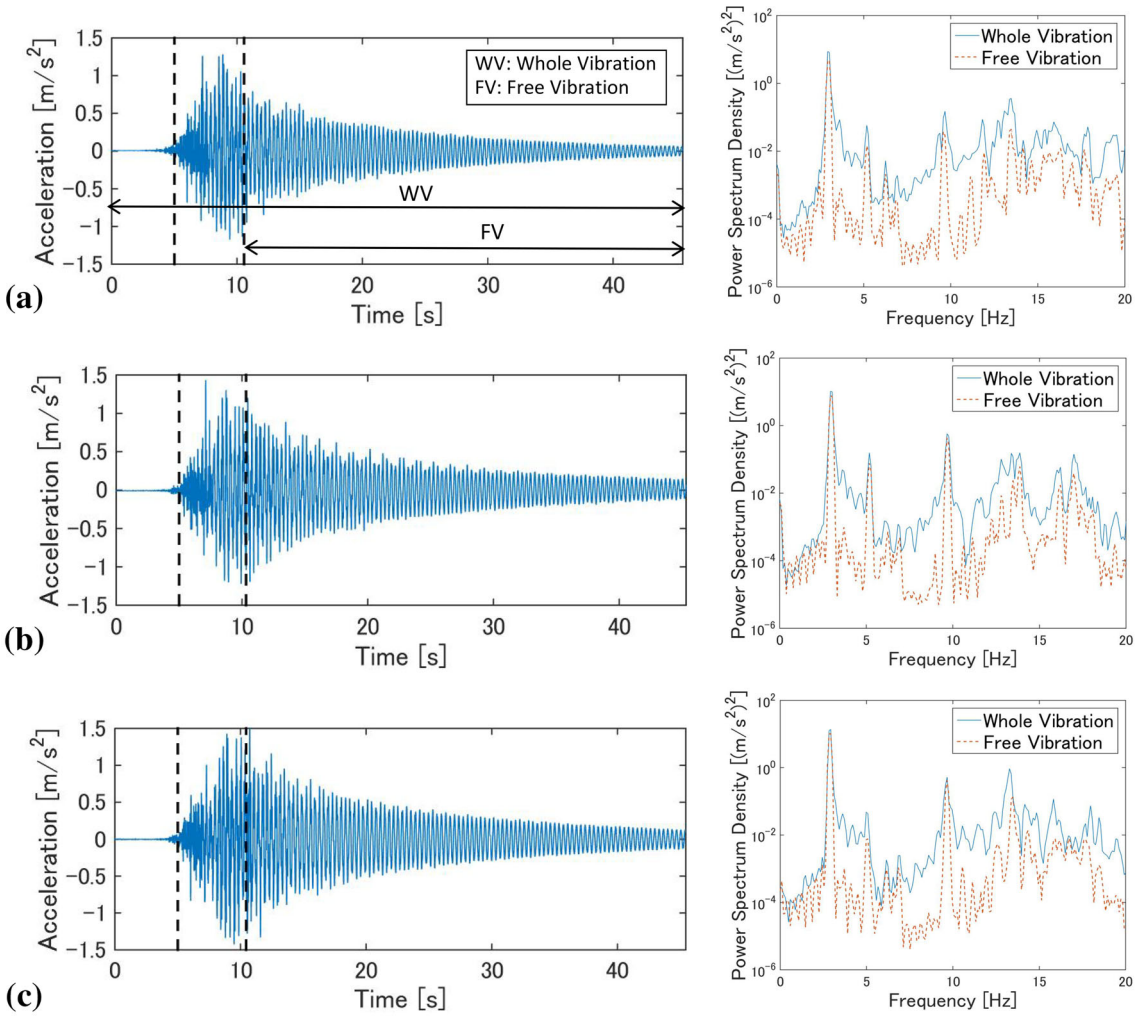

Fig. 4 Examples of acquired time series of A3 sensor and PSD curves: **a** Intact, **b** DMG1, and **c** DMG2

over-fitting in statistical models, the performance of a stochastic model on the training set is known to not always satisfy predictive performance on newly observed data. One effective approach used to confirm the validity of the stochastic model is the use of validation samples that are independent from training samples and have identical distributions with those of the training data. However, in this experiment, the training data are limited, so leave-one-out cross validation (CV) technique [16] is applied to assess the DI validity, i.e. a CV sample of the DI is evaluated using each of the samples in the training set as test data and the remaining samples as a training set. The validity of the DI is confirmed by checking whether the random distribution of the CV samples is identical to the distribution given in Eq. 25, or not. In this study, the sample means of the CV samples are checked if they are not in the rejection region to ensure the validity of damage detection.

4 Damage detection

4.1 Feasibility of damage detection using the proposed damage indicator

Before starting damage detection, the order of the AR model should be ascertained because the higher the AR order is the more precise the dynamical model would be, but requires more parameters to describe the model. Especially when the number of training samples is limited, a feature extraction by the PCA will be difficult for a higher order AR model as the dimension of the vector space to be investigated is higher than that of the training samples. Nevertheless, finding the suitable AR order before monitoring of the bridge is troublesome. In this study, sensitivity analysis with respect to the AR order is performed first. However, it is difficult to extend this approach to other bridges without information about damage. Therefore, for the practical health monitoring, it is recommended to examine several DIs according to different AR orders.

Figure 5 presents the sample means of the DI of the test samples and the CV samples obtained from the FV data of each scenario with respect to the AR order, where the dimension of the principal subspace m was fixed to 2, and demonstrated that the AR model that comprises 5–11 order is suitable for the damage detection. Therefore, in the following discussions, the AR order is fixed to 8, which gives maximum sample means of DIs for every damage scenario (see the vertical dashed line in Fig. 5).

Plots of DIs of the CV samples and test samples and their sample means in a logarithmic scale are depicted in Fig. 6, where m is fixed to 2. Those samples are obtained from the WV data. The expectation given as $E\left[\mathrm{DI}\left(\boldsymbol{c}_i^t\right)\right] = 1$ (see Eq. 26) and the critical values of the rejection regions given as $\frac{1}{mn_t}F^{-1}(1-\alpha|mn_t)$ (see Eq. 30) are also presented in the figure for the hypothesis test, where significance level α is set to 1%. For every damage scenario, the sample means are located in the rejection region. Damage is clearly detectable. For both scenarios of INT and RCV, the sample means of CV samples are not in the rejection region, but they are located closer to the critical values than to the expected value.

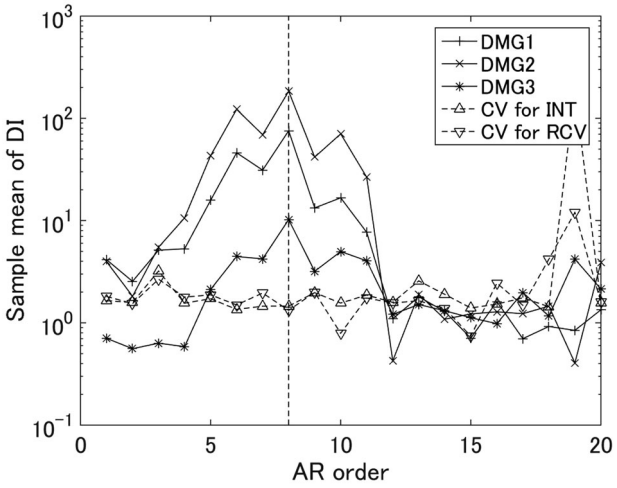

Fig. 5 Sample means of DI of FV data with respect to AR orders

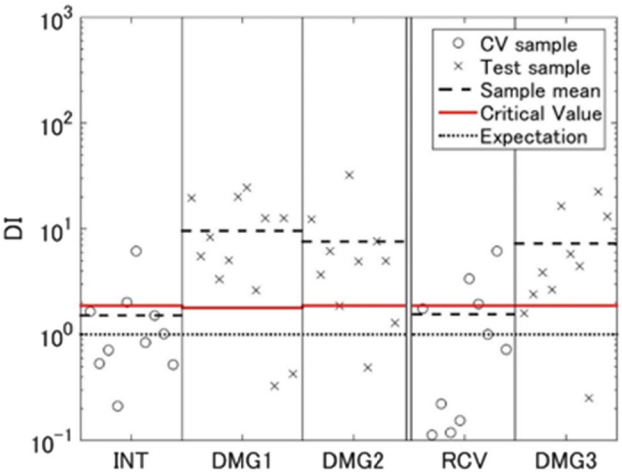

Fig. 6 DIs and test statistics obtained from WV data

Figure 7 depicts the DIs and the test statistics for the damage detection estimated from FV data. Parameters m and α were also fixed to 2 and 1%, respectively, in the same way considered in Fig. 6. For DMG1 and DMG2, the sensitivity of the DIs was improved considerably. It is noteworthy that the vertical scale is larger than that of Fig. 6. The sensitivity of the DIs for DMG3 was also improved slightly. This observation showed that using free vibration data improves performance of the proposed damage detection method.

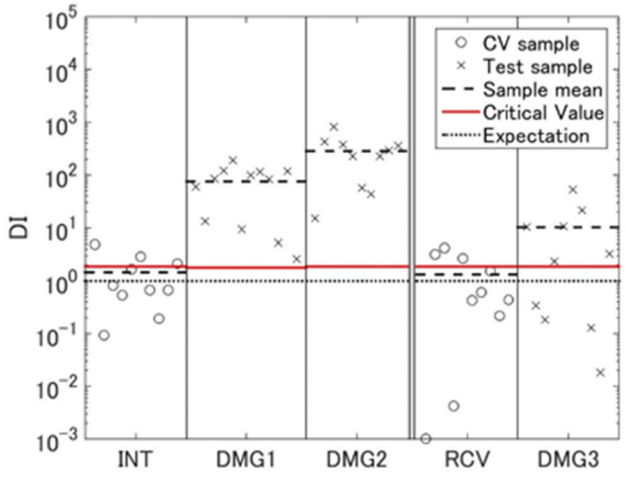

Fig. 7 DIs and test statistics obtained from FV data

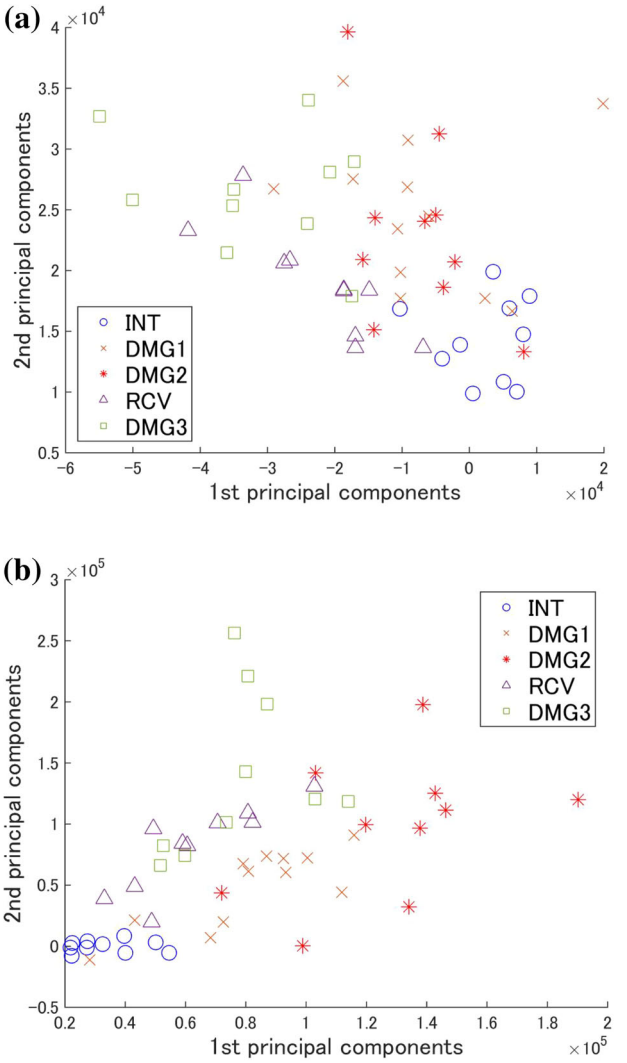

Fig. 8 Principal components of the polynomial coefficients: **a** from WV data and **b** from FV data

The first and second principal components of the polynomial coefficients that are obtained from both of WV and FV data are shown as presented in Fig. 8, where the dataset obtained from the INT scenario is referred as the training set and the dataset from the other scenarios are referred as the test set. Apparently, the principal components of each scenario formed groups. For example, the blue circles plotted in Fig. 8, which is relevant to INT scenario, were distinguished from every other scenario for both of WV data and FV data, which demonstrated that the changes in the modal properties attributable to severing and recovering the members of the bridge are well extracted using the MAR model and PCA. Especially, the FV data caused smaller variation of the principal components from the INT scenario and more sensitive DIs for DMG1 and DMG2 than the WV data. Furthermore, apparently, the sample groups are mutually separated, especially in the FV data. It is noteworthy that the data measured from bridges under traffic during testing are still useful to detect damage even though the free vibration will provide better accuracy.

4.2 Comparison with existing indicator

A brief description about the NDSF is given as follows in advance of making a comparison between the proposed DIs and NDSFs. Using a time series of local acceleration at l-th measurement point denoted as $y_l(k)$, vibration data can be modeled using a univariate AR model described as the following equation, which is a univariate version of Eq. 1:

$$y_l(k) = \sum_{i=1}^{P} a_i^{(l)} y_l(k-i) + e_l(k) \tag{33}$$

In that equation, $a_i^{(l)}$ denotes the ith AR coefficient with respect to the lth measurement point. Also, e_l represents a white-noise time series. The AR coefficients are identified by solving the Yule–Walker equations (see Eqs. 6, 7) derived from the univariate AR model. Each of the AR coefficients, especially the low order AR coefficients, are presumed to include the modal information of the target structure. The NDSF defined in Eq. 34 [4, 5] has, therefore, been investigated to detect bridge damage:

$$d_l = \mathrm{abs}(a_1^{(l)}) \Big/ \sqrt{\sum_{i=1}^{3} a_i^{(l)}} \tag{34}$$

where, abs(·) denotes the absolute value, and d_l stands for the NDSF with respect to the lth measurement point.

The NDSF is obtained independently from each measurement point. Moreover, it is likely to be uncorrelated with the NDSFs at the other measurement points. Existing research showed that the sample means of CV samples of MDs calculated from the NDSFs of the monitored bridge

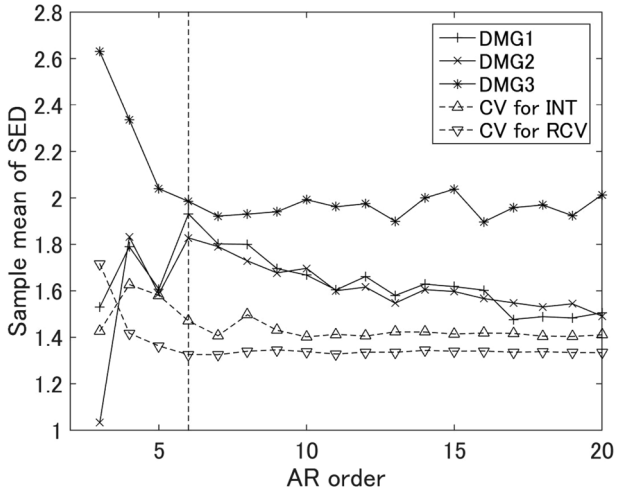

Fig. 9 Sample means of SEDs from FV data with respect to AR orders

are tend to be much larger than expected [7], possibly because the uncorrelated NDSFs caused over fitting problem. Therefore, to validate the hypothesis test discussed in Sect. 2.5, the standard Euclidean distance (SED) in squared units is given as the following equation, which is applied instead of MD for damage detection:

$$\mathrm{SED} = \frac{1}{M}\sum_{l=1}^{M}(d_l - \hat{\mu})^2/\hat{\sigma}_l^2, \tag{35}$$

where M denotes number of the measurement points; $\hat{\mu}$ and $\hat{\sigma}_l^2$, respectively, denote the expectation and the variance estimated from the training data at each measurement point. Assuming independent Gaussian distributions, $M \times \mathrm{SED}$ is presumed to be a χ^2-distribution with M degrees of freedom. Therefore, a hypothesis test for damage detection can be formulated according to the discussion presented in Sect. 2.5.

Figure 9 shows the sample means of the SED obtained from the FV data of each scenario with respect to the AR order in the manner discussed in Fig. 5. For the DMG1 and DMG2, the AR model considered up to 6th order provides the maximum sample mean. For DMG3, the AR model considered up to 3rd order provides the maximum value but the sample means of both CV samples appear to be unstable with the AR order. In the following discussions, therefore, the AR order is fixed at 6 (see the vertical dashed line in Fig. 9).

The SEDs of the NDSFs are shown as presented in Fig. 10, in which those SED and NDSF are obtained from the five sensors deployed on the damage-introduced side of the bridge, i.e. sensors A1–A5 presented in Fig. 2a. Sample means and the critical values for the hypothesis test are also considered, where the significant level α is also fixed to 1%. The vertical axis is scaled linearly in Fig. 10. Figure 10a presents anomalies detected in DMG1 and DMG2 scenarios,

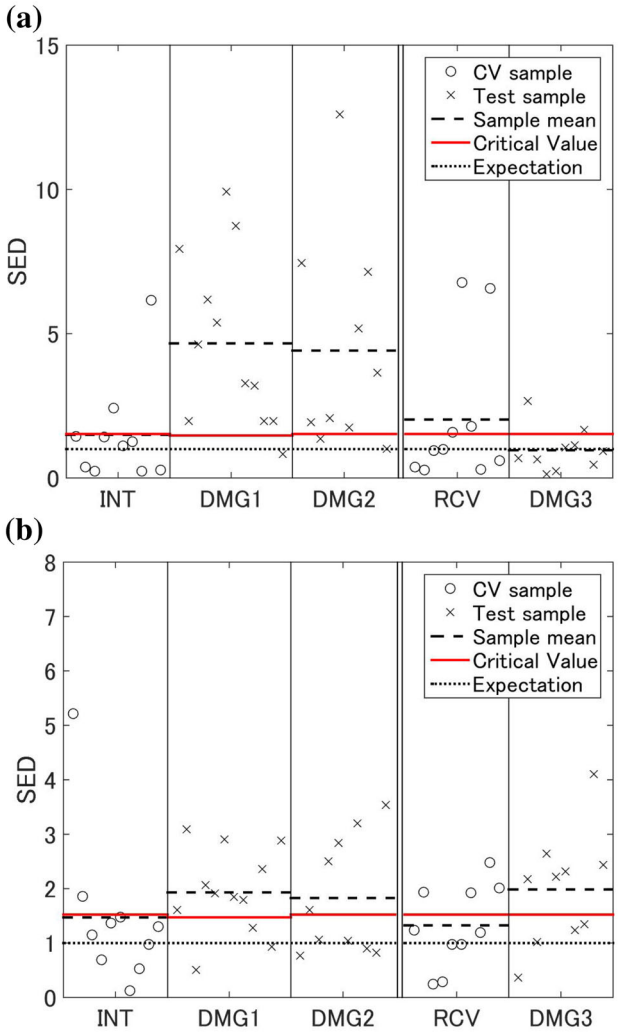

Fig. 10 SEDs of NDSF **a** from WV data and **b** from FV data

but the anomalies were not detected using the test statistic in the DMG 3 scenario. This observation corresponds to those from a previous study considering the MD of the NDSFs [13]. The sample means of CV samples were located close to the critical values because of its outliers, e.g. NDSF seems not precisely Gaussian distributed. Figure 10b shows that anomalies of three damage scenarios are detected using the hypothesis test for FV data. However, the difference between CV samples and the test samples was not significant. Comparing these results from NDSF to results from the novel DI, one observes that the novel DI can improve damage detection performance.

5 Concluding remarks

This study proposed a damage detection method for a bridge structure using the novel damage indicator (DI) obtained from ambient vibrations of bridges. The novel DI

was derived from a multivariate autoregressive (MAR) model. The DI evaluated the stochastic distance of the principal components between a set of reference samples obtained from healthy bridge and unknown test samples using the Mahalanobis distance (MD). Principal component analysis (PCA) was also applied to extract damage-sensitive features from the MAR model. Statistical hypothesis testing based on a probability distribution of the DI was applied for damage detection. Field experiments were conducted on an actual steel truss bridge with truss members that were severed artificially. The feasibility of the proposed DI for damage detection was investigated.

Observations from sensitivity analysis showed that the DI performance depends on the AR order. The proposed damage detection method detected three damage patterns clearly; especially damage detection performance was improved using the free vibration data measured after the vehicle exited the bridge. However, the forced vibration data are still useful because damage was also detected clearly.

The presented DI was compared with the existing damage indicator, named Nair's damage-sensitive feature (NDSF). The hypothesis test for the NDSF is formulated in the same way as the proposed method. The comparison between the NDSF and the novel DI demonstrated that the modal information included in a multivariate system might help to improve damage detection performance by the proposed DI.

Acknowledgements This study is partly sponsored by a Japanese Society for the Promotion of Science (JSPS) Grant-in-Aid for Scientific Research (B) under Project No. 16H04398. That financial aid is gratefully acknowledged. The authors wish to express their gratitude to Nara Prefecture for support in carrying out field experiments.

References

1. Doebling SW, Farrar CR, Prime MB, Shevitz DW (1996) Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review. Los Alamos National Laboratory report, LA-13070-MS
2. Ko JM, Ni YQ (2005) Technology developments in structural health monitoring of large-scale brides. Eng Struct 27(12):1715–1725
3. Peeters B, De Roeck G (2011) One-year monitoring of the Z24-Bridge: environmental effects versus damage events. Earthq Eng Struct Dyn 30(2):149–171
4. Zhang QW (2007) Statistical damage identification for bridges using ambient vibration data. Comput Struct 85(7–8):476–485
5. Deraemaeker A, Reynders E, De Roeck G, Kullaa J (2007) Vibration-based structural health monitoring using output only measurements under changing environment. Mech Syst Signal Process 22(1):34–56
6. Gul M, Catbas FN (2009) Statistical pattern recognition for structural health monitoring using time series modeling: theory and experimental verifications. Mech Syst Signal Process 23:2192–2204
7. Dilena M, Morassi A (2011) Dynamic testing of damaged bridge. Mech Syst Signal Process 25:1485–1507
8. Farrar CR, Baker WE, Bell TM, Cone KM, Darling TW, Duffey TA, Eklund A, Migliori A (1994) Dynamics characterization and damage detection in the I-40 Bridge over the Rio Grande. Los Alomos National Laboratory report: LA-12767-MS
9. Nair KK, Kiremidjian AS, Law KH (2006) Time series-based damage detection and localization algorithm with application to the ASCE benchmark structure. J Sound Vib 291(1–2):349–368
10. Kim CW, Isemoto R, Sugiura K, Kawatani M (2013) Structural fault detection of bridges based on linear system parameter and MTS method. J JSCE 1(1):32–43
11. Mahalanobis PC (1936) On the generalized distance in statistics. Proc Natl Inst Sci India 2(1):49–55
12. Taguchi G, Rajesh J (2000) New trends in multivariate diagnosis. Indian J Stat (Series B) 62(2):233–248
13. Kim CW, Chang KC, Kitauchi S, McGetrick PJ (2016) A field experiment on a steel Gerber-truss bridge for damage detection utilizing vehicle-induced vibrations. Struct Health Monit 15(2):174–192
14. Chang KC, Kim CW (2016) Modal-parameter identification and vibration-based damage detection of a damaged steel truss bridge. Eng Struct 122:156–173
15. Jolliffe IT (2002) Principal component analysis. Springer, New York
16. Bishop CM (2006) Pattern recognition and machine learning. Springer, New York
17. He X, De Roeck G (1997) System identification of mechanical structures by a high order multivariate autoregressive model. Comput Struct 64(1–4):341–351
18. Kim CW, Kawatani M, Hao J (2012) Modal parameter identification of short span bridges under a moving vehicle by means of multivariate AR model. Struct Infrastruct Eng 8(5):459–472
19. Heylen W, Lammens S, Sas P (1997) Modal analysis theory and testing. KU Leuven, Leuven
20. Kim CW, Kitauchi S, Sugiura K, Kawatani M (2014) Utilizing reproduced autoregressive model for damage detection of real truss bridges. In: 4th international symposium on life-cycle civil engineering. 16–19 November 2014, Tokyo, pp 369–376