

Long-term health monitoring for deteriorated bridge structures based on Copula theory

Yi Zhang^{*1}, Chul-Woo Kim^{1a}, Kong Fah Tee^{2b}, Akhil Garg^{3c} and Ankit Garg^{3d}

¹Department of Civil and Earth Resources Engineering, Graduate School of Engineering, Kyoto University, Japan

²Department of Engineering Science, University of Greenwich, United Kingdom

³Department of Mechatronics Engineering, Shantou University, Shantou, China

(Received April 20, 2017, Revised July 25, 2017, Accepted August 9, 2017)

Abstract. Maintenance of deteriorated bridge structures has always been one of the challenging issues in developing countries as it is directly related to daily life of people including trade and economy. An effective maintenance strategy is highly dependent on timely inspections on the bridge health condition. This study is intended to investigate an approach for detecting bridge damage for the long-term health monitoring by use of copula theory. Long-term measured data for the seven-span plate-Gerber bridge is investigated. Autoregressive time series models constructed for the observed accelerations taken from the bridge are utilized for the computation of damage indicator for the bridge. The copula model is used to analyze the statistical changes associated with the modal parameters. The changes in the modal parameters with the time are identified by the copula statistical properties. Applicability of the proposed method is also discussed based on a comparison study among other approaches.

Keywords: copula; structural health monitoring; bridge structure; long-term assessment

1. Introduction

Performance of deteriorated bridges is a key concern in retaining the stability of land transport. Maintenance strategies and repair works are usually required in order to prevent failure of such important infrastructures. However, developing an effective maintenance strategy is not easy as it needs timely evaluations on the health conditions of bridges. As most bridges have been functioned for several decades, certain potential risks are expected due to the structural strength reduction in long term. Moreover, following the fast economic development nowadays, the need for public transportation is increasing. This leads to more number of vehicles and higher load burden to the bridge which even worsen the situation. Therefore, an economical and efficient structural monitoring system for the bridges would be useful for ensuring its serviceability. Structural health monitoring of old deteriorated bridges has become of great significance in civil engineering community in past decades. Many advanced techniques

have been utilized and been attempted to evaluate the structural health condition. Park *et al.* (2007) have implemented the laser scanning technology to detect the displacement of the structure. The use of piezoceramic patches in the health monitoring of concrete bridges have been reported by Soh *et al.* (2000). Li *et al.* (2016) have provided an overview of the development in the field of structural health monitoring by using the optical sensors. Until recently, the advances and achievements in ultrasonic wave structural health monitoring have been discussed in Mutlib *et al.* (2014). Among these, the use of vibration data is still one of the most common and promising technologies (Feng *et al.* 2017, Hoag *et al.* 2017, Zhang *et al.* 2017a, b).

The idea is to use the structural dynamic parameters, such as the frequency, damping and modal shape, as an indicator for the health condition of bridges. Fundamentally, if a change in the parameters is observed, the health condition of bridge structures is considered changed. Typical applications include, for example, Chang and Kim (2016) have identified the modal frequencies and mode shapes of the bridge by using a stabilization diagram-aided multivariate autoregressive analysis of vehicle-excited bridge vibrations. Dilena *et al.* (2011) have estimated the modal parameters of a damaged bridge from frequency response measurements using harmonic excitations. A key step in the damage identification process is to find anomalies out from the observations. Many works have been devoted to this task based on statistical properties of observed variables. However, significant research conducted was focussed on individual statistical properties whilst the dependencies are simply ignored. Since vibration data is collected from a huge structural monitoring sensor network, such ignorance would lead to a large waste of

*Corresponding author, JSPS Research Fellow

E-mail: zhang_yi87@163.com

^a Professor

E-mail: kim.chulwoo.5u@kyoto-u.ac.jp

^b Reader

E-mail: k.f.tee@greenwich.ac.uk

^c Associate Professor

E-mail: akhil@stu.edu.cn

^d Associate Professor

E-mail: ankit@stu.edu.cn

information in the analysis.

An alternative way of improving the current identification techniques is to utilize the statistical properties of dependencies between the observed data. The canonical correlation analysis is perhaps one of these approaches that have drawn the most attention (2007). The idea of using the Mahalanobis Distance to detect the anomalies in the multivariate time series data has also been noticed by Kim *et al.* (2013) and Tee *et al.* (2013). However, all these approaches are having a general assumption that the data from different sources are linearly dependent. If the actual dependency is nonlinear, then the applied approach may only offer quite coarse approximations. Unfortunately, the nonlinear dependency is commonly existed between the sensor observations. The characterization of nonlinear dependency among the sensor network remains challenging and requires robust multivariate models to detect structural damage from the partial available information.

To remedy this problem, this study aims to provide a methodological contribution by developing copula-based multivariate models for the modal parameters of bridge structures, and an associated identification approach to detect structural damage. The paper is organized as follows. Section 2 will firstly present the fundamental modal parameter-based bridge diagnosis method using time series models. Then the copula-based damage identification method is described in Section 3. Section 4 then demonstrates the use of the developed method for a real old deteriorated bridge. Detailed benefits of utilizing the copula concept in the analysis are discussed. The concluding remarks are summarized in Section 5.

2. Structural damage identification from AR model

There are many techniques developed nowadays on the health monitoring of bridge structures based on damage-sensitive features. One of the common methods is focusing on changes in system frequencies and structural damping constants by utilizing a time series model such as the autoregressive (AR) model.

2.1 Damage indicator from AR coefficients

In a linear dynamic system, the output can be idealized by an AR model as

$$y(k) = \sum_{i=1}^p a_i y(k-i) + e(k) \quad (1)$$

where $y(k)$ represents the output of the structural dynamic system at time k , a_i is the i th order AR coefficient, and $e(k)$ denotes an error term. However, there are several key points to be addressed when using the time series model in modal parameter-based bridge diagnosis. For instance, the number of order terms used in a time series model for characterizing the vibration responses of bridge structures should be selected with caution. An inappropriate selection may result to spurious modal parameters and thus lead to unreliable results of system frequencies and damping constants.

A possible way to determine the optimal AR order can be obtained by the means of Akaike Information Criterion (AIC). The standard AIC for a time series model can be given as

$$AIC = n \log(2\pi E^2) + 2(m+1) + n \quad (2)$$

where n is the number of data, m is the number of orders in AR model, and E^2 is the square of prediction errors. AIC takes into account both the complexity and the goodness-of-fit of the model. A smaller value of AIC indicates a better model.

Once the order of AR model is confirmed, the AR coefficients can be transformed and linked to the modal parameters. By utilizing a z -transformation, the AR model with order of p can be expressed in the z -transform as

$$y(k) = \frac{1}{H(z)} e(k) = \frac{1}{1 + \sum_{i=1}^p a_i z^{-i}} e(k) \quad (3)$$

where $H(z)$ represents the transfer function of the system in the discrete-time complex domain, and z^{-i} is a forward shift operator. The denominator of the transfer function is the characteristic equation of the dynamic system and

$$z^n + a_1 z^{n-1} + a_2 z^{n-2} + \dots + a_{n-1} z + a_n = 0 \quad (4)$$

Therefore, the frequency and damping constant of each structural mode can be determined from the complex conjugate poles as follows.

$$z_k = \exp\left(h_k \omega_k \pm j \omega_k \sqrt{1-h_k^2}\right) \quad (5)$$

where j represents the imaginary unit, h_k and ω_k are the damping constant and circular frequency of the k th mode of the structural system, respectively. The function $H(z)$ is also defined as AR polynomial of the model transfer function which relates the input to output. The poles z_k in Eq. (4) can be estimated by finding the roots of the AR coefficient polynomial in the denominator of $H(z)$. As the coefficients of $H(z)$ are real, the roots should be real or complex conjugate pairs. The number of poles in z -plane must equal to the number of orders in AR model. Therefore, damages in the structure (e.g., a change in frequency and damping ratio) will correspond to changes in the AR model coefficients. Based on this concept, the AR model based damage indicator (DI) can be defined as

$$DI = \frac{|a_1|}{\sum_{i=1}^n \sqrt{a_i^2}} \quad (6)$$

where DI stands for the damage indicator, and also changes due to damage. Usually, in computing the damage indicator, it is not necessary to include all the order terms in the AR model. For example, the higher order terms in the AR model are corresponding to high frequency modal parameters which are insignificant in the structural dynamics. Therefore, the following equation is suggested

for a simplified damage indicator

$$DI = \frac{|a_1|}{\sqrt{a_1^2 + a_2^2 + a_3^2}} \quad (7)$$

where only the first three AR coefficients are utilized.

It was shown by Nair (2006) that these AR coefficients are the most significant ones in the AR model when there are damages occurred in the structure. The conclusion was also agreed by Kim *et al.* (2013), who observed from a bridge-moving vehicle laboratory experiment that *DI*, considering up to the third order of AR coefficients, is a promising parameter in bridge health monitoring. Feasibility of the *DI* for real bridge application is examined in (2016) which concluded that the *DI* is superior to modal parameters, with the advantages of preventing potential human errors.

2.2 Removing environmental and operational influences

Before performing any further damage identifications based on the calculated modal parameters, it should be realized that the environmental and operational changes could have significant influences on the modal parameters. For example, the temperature could be quite an influential factor that can greatly affect the structural frequency values.

This is because the Young's modulus is usually temperature dependant. Theoretically, for the same structural system, natural frequencies are proportional to the Young's modulus of the system. Therefore, a change in the temperature would eventually lead to a change in the structural natural frequency. In this sense, certain data pretreatment is required to remove the associated influences in order to have more independent modal parameter dataset.

To achieve this, Kalman filter and autoregressive exogenous (ARX) model (Varouchakis 2016, Omenzetter and Brownjohn 2006, Yan *et al.* 2004) are applied in this study. Following are details regarding application.

2.2.1 Kalman filter

Kalman filter can be utilized to remove the noises that might be induced by the environmental and operational conditions. In combination with ARX model, the disturbances caused by environmental and operational factors can be well identified and eliminated. Another advantage of applying Kalman filter is its updating procedure. Since it is an online updating algorithm, the structural health monitoring can be conducted in a more efficient way. The converged model parameters in Kalman filter after updating can be considered as an optimal model parameters. Therefore, the noises can be removed by applying the identified model parameters in the time series data analysis. The detailed procedures can be described in following steps.

The state space model for the Kalman filter is given by

$$\mathbf{x}_t = \mathbf{F}\mathbf{x}_{t-1} + \mathbf{G}\mathbf{v}_t \quad (8)$$

$$\mathbf{y}_t = \mathbf{H}_t\mathbf{x}_t + \mathbf{w}_t \quad (9)$$

where \mathbf{x}_t and \mathbf{y}_t are the state matrix and observation at time t , respectively, \mathbf{v}_t and \mathbf{w}_t denote noise at time t . For a particular time series model, \mathbf{F} , \mathbf{G} and \mathbf{H}_t can be estimated from the collected data. In the present structural health monitoring, \mathbf{x} could be referring to the observed damage indicators, e.g., *DI*.

Therefore, the future state can be predicted by means of the Kalman filter as follows (Kitagawa and Gersch 1984)

$$\mathbf{x}_{t|t-1} = \mathbf{F}\mathbf{x}_{t-1|t-1} \quad (10)$$

$$\mathbf{V}_{t|t-1} = \mathbf{F}\mathbf{V}_{t-1|t-1}\mathbf{F}^T + \mathbf{G}\mathbf{Q}\mathbf{G}^T \quad (11)$$

where \mathbf{x} and \mathbf{V} denote the predicted mean matrix and covariance matrix at time t given the value at time $t-1$, respectively. \mathbf{Q} stands for the covariance matrix of \mathbf{v}_t .

The filtered state can also be calculated from the following equations

$$\mathbf{x}_{t|t} = \mathbf{x}_{t|t-1} + \mathbf{K}_t (\mathbf{y}_t - \mathbf{H}_t\mathbf{x}_{t|t-1}) \quad (12)$$

$$\mathbf{V}_{t|t} = (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\mathbf{V}_{t|t-1} \quad (13)$$

where \mathbf{K}_t stands for the Kalman gain defined by

$$\mathbf{K}_t = \mathbf{V}_{t|t-1}\mathbf{H}_t^T (\mathbf{H}_t\mathbf{V}_{t|t-1}\mathbf{H}_t^T + R)^{-1} \quad (14)$$

where R denotes the variance of w_t .

2.2.2 ARX model

Although Kalman filter could remove a significant part of the noises, the influences induced by environmental and operational factors can still exist in the observed data. In order to further eliminate these effects, the ARX model could be adopted. Theoretically, it could express the environmental and operational factors discretely as exterior disturbances (Bishop 2006). To remove the environmental and operational influences from the observed *DI*, this can be obtained by

$$\mathbf{DI}^{(t)} = \sum_{i=1}^q \alpha_i \mathbf{DI}^{(t-i)} + \sum_{i=1}^r \beta_i u^{(t-i)} + w_t \quad (15)$$

where q and r are the orders of the time series model, α_i and β_i are the associated order coefficients, $u^{(t)}$ denote the observed environmental and operational factors at time t and w_t is the white noise. Furthermore, by connecting Eqs. (11) and (15), the relationships among the parameters and matrices can be given by

$$\mathbf{y}_t = \mathbf{DI}^{(t)} \quad (16)$$

$$\mathbf{F} = \mathbf{I} (\text{Identity matrix}) \quad (17)$$

$$\mathbf{G} = \mathbf{0} \text{ (Null matrix)} \quad (18)$$

$$\mathbf{X}_t = [\alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_r]^T \quad (19)$$

$$\mathbf{H}_t = [\mathbf{DI}^{(t-1)}, \dots, \mathbf{DI}^{(t-q)}, u^{(t-1)}, \dots, u^{(t-r)}] \quad (20)$$

$$\mathbf{R} = \mathbf{I}$$

While the initial mean and covariance matrices can be assumed to be

$$\mathbf{x}_{00} = \mathbf{0} \text{ (Null matrix)} \quad (21)$$

$$\mathbf{V}_{00} = \mathbf{I} \text{ (Identity matrix)} \quad (22)$$

The optimal model parameters are obtained by satisfying

$$|\mathbf{x}_{t|t} - \mathbf{x}_{t-1|t-1}| \cong 0 \quad (23)$$

Therefore, the environmental and operational effects can be greatly removed by characterizing the dependences in the ARX model.

2.3 Damage identification from multiple observations

Once the data corresponding to calculated modal parameters is preconditioned, the damage identification can be conducted. For instance, the DI could be utilized as a reference in observing the monitored structure. A large deviation with the observations of structure at its intact condition would be considered as a change in the structural condition. The damage indicator based structural health monitoring method is easy and convenient. However, it is usually criticized for its inefficiency in data usage. The construction of a time series model usually requires too much information. This normally needs a very long observation time. Moreover, the stability of the damage index in single time series is not reliable enough to judge the safety of the whole structure. The use of only one parameter to identify the health condition of bridge structure is usually too bias and unreliable. For instance, the identification accuracy of frequencies from the vibration data is higher than that of the damping constants. Based on this concern, the frequencies are more preferred in detecting the structural damage. However, , damping constants are more sensitive to the structural damage. The detection of changes in damping constants would be more meaningful compared to the frequencies. Therefore, many studies attempt to utilize the monitoring data from multiple sources. Among these, the Mahalanobis Distance (**MD**) is one of the widely adopted approaches (Chang and Kim 2016).

The Mahalanobis Distance is generally a measure of the distance between the observations and the theoretical values. In detecting the structural damages, it is usually utilized to

calculate the distance of an observation $\mathbf{X} = (x_1, \dots, x_n)^T$ from their theoretical mean $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^T$. This can be expressed as

$$\mathbf{MD}(\mathbf{X}) = \sqrt{(\mathbf{X} - \boldsymbol{\mu})^T \mathbf{S}^{-1} (\mathbf{X} - \boldsymbol{\mu})} \quad (24)$$

where \mathbf{S} is the covariance matrix for the n-dimensional data \mathbf{X} . The theoretical mean $\boldsymbol{\mu}$ can be obtained from the monitored data of the bridge at an early stage which is set as a reference. Therefore, a large deviation from the theoretical mean would imply a structural damage in the bridge.

Furthermore, the Bayes factor can also be introduced to make judgment of the value changes. Under the Bayesian theory, the probability of a hypothesis conditionally on observed data can be calculated. For example, if the null hypothesis (H_0) is indicating a ‘healthy’ condition and the alternate hypothesis (H_1) is defined as a ‘damage’ condition, the posterior odds can be obtained by utilizing priors and marginal likelihoods as follows

$$B = \frac{p(\mathbf{D} | H_1)}{p(\mathbf{D} | H_0)} = \frac{\int p(\mathbf{D} | \theta_1, H_1) p(\theta_1 | H_1) d\theta_1}{\int p(\mathbf{D} | \theta_0, H_0) p(\theta_0 | H_0) d\theta_0} \quad (25)$$

where \mathbf{D} refers to the data obtained during monitoring, $p(\mathbf{D} | H_0)$ and $p(\mathbf{D} | H_1)$ are the marginal likelihoods, and θ_0, θ_1 are parameters under hypothesis H_0 and H_1 respectively. Bayes factor is in fact the ratio of likelihood of the ‘healthy’ scenario and ‘damage’ scenario. By utilizing a similar concept of Mahalanobis distance, the hypothesis H_0 and H_1 can be defined as

$$\mathbf{H}_0 : x_i = \mu_i \quad (26)$$

$$\mathbf{H}_1 : x_i \neq \mu_i$$

which measures how much deviation between the observed data and the reference data.

If the Bayes factor is greater than 1, it indicates the data favor hypothesis H_1 and hence suggests that there is a structural damage. If Bayes factor is less than 1, it generally implies a ‘healthy’ structural condition.

Although **MD** and Bayes factor are both simple concepts that can be efficiently applied, the drawbacks of these two are also obvious. Mahalanobis distance might not offer an accurate measure of structural damage if the observations are nonlinearly dependent. The Bayes factor only measures the data mean. The change of dependency structure could not be captured by the Bayes factor. Therefore, with the aim of advancing the field of structural health monitoring, there is a strong need for establishing a more robust damage identification approach that could utilize the relationships between multiple observations.

3. Copula theory

To achieve a more efficient use of multivariate data, the concept of copula could be introduced in the structural health monitoring analysis (Zhang *et al.* 2015). Copula is a powerful tool for modeling multivariate data which has

been widely applied in finance and economics (Cherubini *et al.* 2014), offshore engineering (Zhang 2015, Zhang and Lam 2015), reliability theory (Zhang and Lam 2016a, b) as well as in hydrology and environmental sciences (Vanem 2016).

3.1 Basic formula

Copula is a model which could connect univariate marginal distributions to a multivariate distribution. In practice, a multivariate distribution function can be constructed by combining the marginal distributions with specific dependence structures where the latter is called a copula function. This copula function can be expressed as

$$C: [0,1]^n \rightarrow [0,1] \quad \text{and} \quad H(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)) \quad (27)$$

where $H(\cdot)$ is the cumulative joint distribution function and $F_i(\cdot)$ is the individual cumulative marginal distribution function for the i th variable. Specifically, copula C is a cumulative distribution function which connects the one-dimensional probability distributions $F_1(x_1), \dots, F_n(x_n)$ to a multivariate probability distribution $H(\cdot)$.

Generally, the most important characteristic in a copula model is the dependence structure. A copula model can describe various kinds of dependencies which include association concepts such as concordance, linear correlation and the related dependence measures. It is much more flexible than traditional concepts for characterizing multivariate dependencies. More detailed theoretical review can be found in (Nelson 2006, Genest and Favre 2007, Genest and Neslehova 2007, Joe 2015). Typical examples of copulas are shown in Appendix A.

3.2 Copula based damage detection method

In the damage identification problem, the structural damage can be detected by a change-point detection algorithm from the dependence structure in the multivariate series. For an n -dimensional data series $\mathbf{X} = (x_1, \dots, x_n)^T$, the joint distribution of the multivariate variables at time t can be formulated by a copula function by

$$H(\mathbf{X}_t) = C(u_t = (u_{1,t}, \dots, u_{n,t}) | \theta_t) \quad (28)$$

where u_i is the marginal probability vector of x_i , e.g., $u_i = F_i(x_i)$; $C(\cdot)$ is the copula function and θ_t stands for the copula parameter at time t .

Assuming there is no structural damage at $t+1$, then the type of copula should be generally constant. Thus, the null hypothesis of no change in the dependence structure of \mathbf{X} can be given as follows

$$\mathbf{H}_0: \theta_t = \theta_{t+1} \quad (29)$$

$$\mathbf{H}_1: \theta_t \neq \theta_{t+1}$$

If the null hypothesis H_0 is rejected, it indicates a dependence structural change in the multivariate series which implies an event of structural damage. The structural damage may not occur at a particular moment, and therefore

the copula parameter values can be compared between any two time points. That is, the hypothesis in Eq. (29) can be changed to $\mathbf{H}_0: \theta_{t_1} = \theta_{t_2}$ and $\mathbf{H}_1: \theta_{t_1} \neq \theta_{t_2}$.

The null hypothesis should be rejected when a small value of the test statistic of copula-based likelihood ratio is noticed. The copula-based likelihood ratio can be expressed by

$$\Lambda_{t_1, t_2} = \frac{L_{t_1+t_2}(X_{t_1+t_2})}{L_{t_1}(X_{t_1})L_{t_2}(X_{t_2})} = \frac{\prod c(u_{t_1+t_2} | \theta_{t_0})}{\prod c(u_{t_1} | \theta_{t_1}) \prod c(u_{t_2} | \theta_{t_2})} \quad (30)$$

where $L_{t_1}(\cdot)$, $L_{t_2}(\cdot)$ and $L_{t_1+t_2}(\cdot)$ are the likelihood functions for the multivariate data during time periods t_1 , t_2 and t_1+t_2 , respectively. On the condition that a structural damage is occurred between t_1 and t_2 , the hypothesis H_0 will be rejected. This can be tested by the following statistic (Xiong *et al.* 2015)

$$Z_{t_1, t_2} = -2 \ln \Lambda_{t_1, t_2} \quad (31)$$

If Z_{t_1, t_2} is large, it implies there is a structural damage.

The former work has derived the asymptotic distribution of $Z^{1/2}$ for the hypothesis test. This is given by

$$P(Z^{1/2} \geq z) = \frac{z^k \exp(-z^2/2)}{2^{k/2} \Gamma(k/2)} \times \left[\ln \frac{(1-h)(1-l)}{hl} - \frac{k}{z^2} \ln \frac{(1-h)(1-l)}{hl} + \frac{4}{z^2} + O\left(\frac{1}{z^4}\right) \right] \quad (32)$$

where k is the number of copula parameters, $\Gamma(\cdot)$ is the gamma function, $h=l=[\ln(n)]^{3/2}/n$ and n is the number of data. Thus, if the p value calculated in Eq. (32) is less than a criteria value such as the significance level, an inference can be made that the structure is damaged. This provides an opportunity that this concept can be implemented into the structural damage identification.

4. Framework of copula based long term structural health monitoring

Based on the theoretical concepts introduced in Sections 2 and 3, the framework of copula indicator based long term structural health monitoring is developed herein. The whole process can be divided into three stages which can be described as follows.

Step 1: Data collection and structural modal analysis

The initial step in the structural health monitoring is to extract the key structural information from the observed dynamic data. In this first step of the proposed framework, the measured data in the monitored structure including the accelerations, displacement and temperatures from the associated sensors are necessarily gathered. This group of data will serve as a reference for judging structural condition changes in a future observation. The accuracy of the data collected at the site is quite essential for the monitoring. Following the structural modal analysis procedures given in Section 2.1, the modal parameters including frequency and damping ratio are identified. Based on the results, the damage indicator DIs are calculated for

the key modal parameters. Usually, the structural damages are directly related to the structural modal frequencies. Therefore, DI for the structural frequency is employed in this case.

Step 2: Apply Kalman filter and ARX model

The second step is to remove the environmental and operational influences in the time series data. However, it should be recognized that the influences of environmental and operational conditions in the modal parameters are quite different. For example, the temperature might have significant influence on structural frequencies, but may not have large effect on damping ratio. In this proposed framework, such influences will be treated for different modal parameters separately.

In this study, the key elements regarding the impacts of environmental and operational conditions on bridge structures are believed to be the temperature and the vehicle weight. The establishing of ARX model considers both effects on the bridge structural modal parameters. As such, the number of variable in the regression of ARX model are two. It should also be noted the Kalman filter need to be validated in every step of data updating. Values of the model parameters have to be recalculated when new observed data is analyzed.

Step 3: Calculate the copula based damage indicator and evaluate the structural health condition

The last step is the calculation of the copula based damage indicator. During the monitoring period, the dynamic data is collected from the bridge structure. The modal parameters are calculated following the same procedures in Step 1. Based on the constructed ARX model in Step 2, the environmental and operational effects are removed from the modal parameters. From the filtered time series data, a pair of modal parameters can be utilized to compute the copula based damage indicator. In this sense, the most key modal elements should be selected. The value of copula damage indicator will be updated when new data is withdrawn from the measurement.

In the present study, the first mode frequency and damping ratio is chosen for calculating the copula damage indicator. This is mainly because many former works have indicated that these two modal parameters are the most important factors governing the structural safety condition. Once the result of copula based damage indicator is calculated, it can be compared with a marginal value to judge whether the structural condition is changed or not. In this sense, the value obtained in the first step serves as the reference and the new updated data would represent the current structural condition. In other words, the reference data should be taken from the time when the structure is at its intact condition. A detailed flowchart of the whole calculation procedures is highlighted in Fig. 1. However the applicability and suitability of these developed procedures for bridge structural health monitoring have not been examined. From the different indicators of structural damage, the type of assessment best describes the structural condition has to be made and this is elucidated in the following section.

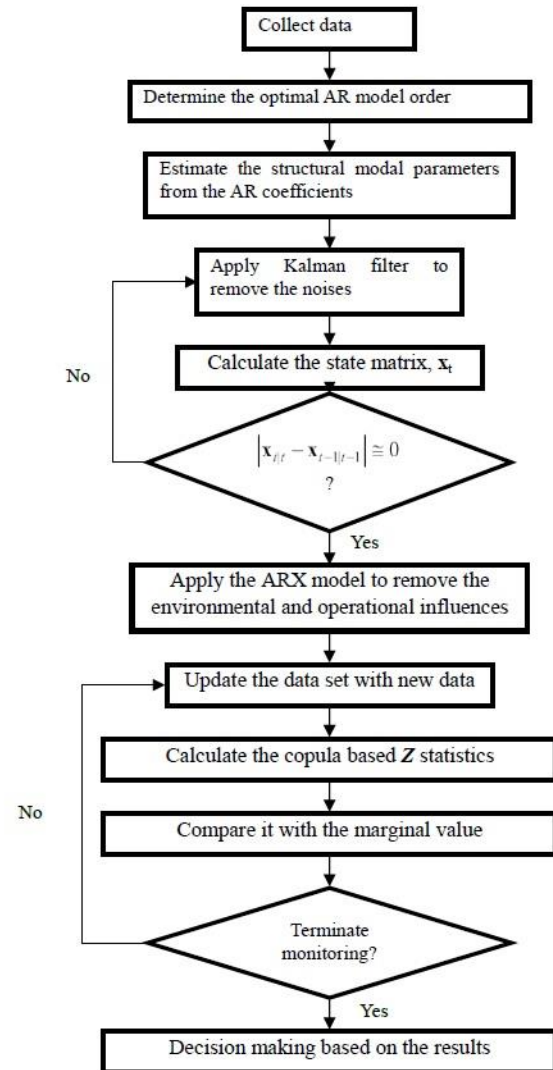


Fig. 1 Framework of copula based structural health monitoring procedures

5. Case study: long-term monitoring for an in-service steel plate-girder bridge

A steel plate-girder bridge with Gerber system which is located in Himeji, a city of Japan, is analyzed in this study. It has a width of 8 meters and length of 142 meters with seven spans. The bridge was constructed in 1960 and was operated since then. A picture of this bridge can be seen from Fig. 2. The basic information of this bridge is also provided in Table 1.

The bridge experienced fatigue damage due to high traffic volumes of heavy trucks. Since 2008 after full reinforcement against fatigue failure, the bridge has been monitored by the dynamic sensors. The sensors are put at four locations in the middle and the end of the bridge. Two are put at the up lane and remaining two are put at the down lane. A plan view of these sensors is plotted in Fig. 3. The marked red squares UA-1, UA-2, DA-1 and DA-2 stand for accelerometers that are used to measure acceleration

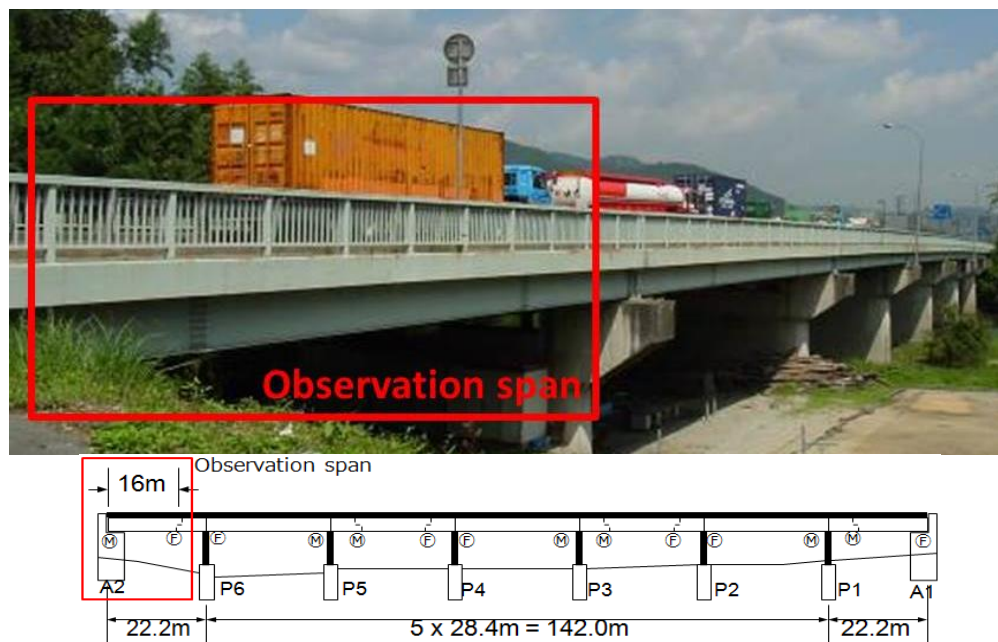


Fig. 2 General picture of the bridge

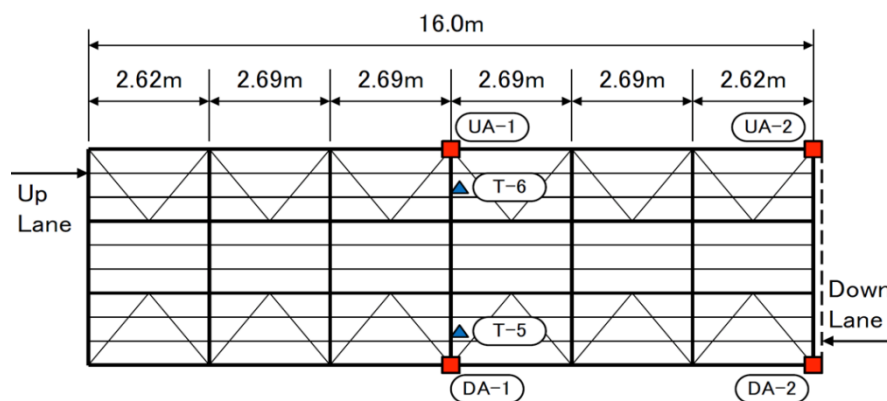


Fig. 3 Location of sensors on the bridge

Table 1 The basic information of the bridge

Construction year		1960
Bridge length (m)		142
Span length (m)	Hanging girder	16.0
	Anchorage girder	6.2+28.4+6.2
Width (m)		8.0

responses of steel girders on up and down lanes respectively. The sampling rate of the accelerometers is 200 Hz. Another two thermometers (T-5 and T-6) are also placed in the middle of the bridge to accompany the accelerometers. The temperature is measured and recorded once per hour. The monitoring system was working from 2008 to 2016 but however suffered several system errors and stopped functioning for some years.

This study intends to detect the long-term structural damage associated with the bridge from the vibration characteristics. As discussed in Section 2, the use of AR model in the structural health monitoring is based on a comparison between the observed structural condition and the intact structural condition. For this reason, the data collected at the earliest time is utilized as the reference for intact structural conditions. This covers the data recorded from 6th August 2008 to 21st June 2009. It was reported that

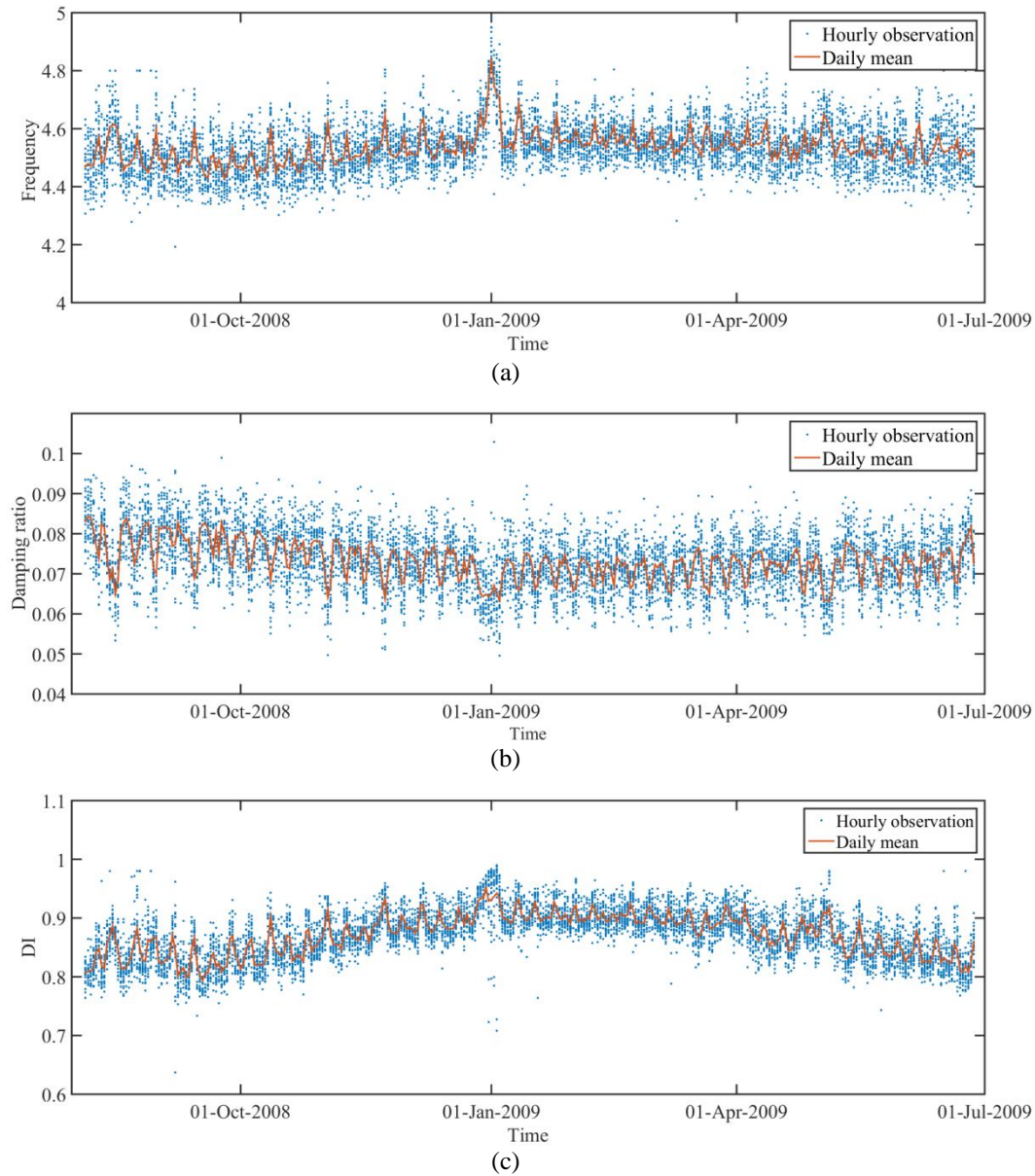


Fig. 4 Observed (a) frequency, (b) damping ratio and (c) DI from sensor DA-1 from July 2008 to July 2009

the bridge is in a perfect condition with no damage and deterioration during this period. Unfortunately, a car accident occurred in 2010. The bridge suffered impact during this accident and believed having some minor influences although no structural damage was reported by an emergency visual inspection. Moreover the periodic bridge inspection in 2013 found peeled painted surface of bridge members which was suspected as a sign of fatigue crack. Therefore, the observations after the year 2010 will be analyzed and compared to the data corresponding to intact condition of structure. It is noteworthy that operational influences to the long-term bridge monitoring are not considered in this study as this study mainly focuses on feasibility of the copula-based damage detection.

In this study, the key modal parameters including frequency and damping ratio are utilized to detect any sign of bridge damage. The data from sensor DA-1 is utilized in this analysis. Since this is a short span bridge, the first mode is believed to be dominant and therefore is chosen for the investigation. Based on the AR model based structural identification algorithm, which was introduced in Section 2 (Sung *et al.* 2017), the first mode frequency and damping ratio are identified. It should be noted the lower modal parameters are quite sensitive to boundary conditions. As the inspected point is located at a short span in the bridge, the boundary conditions of this inspected part are believed to be fixed. This is because the size and scale of this part is quite small compared to the whole bridge. However, it

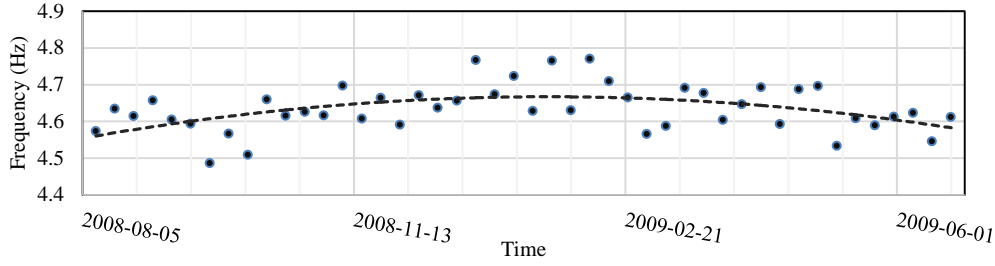


Fig. 5 Observed weekly first mode frequency and the fitted trend

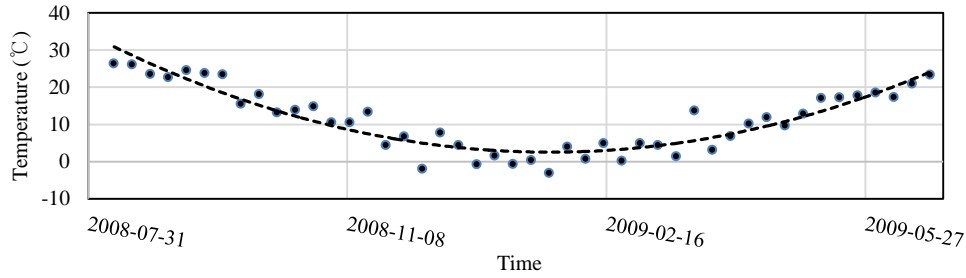


Fig. 6 Observed weekly temperature and the fitted trend

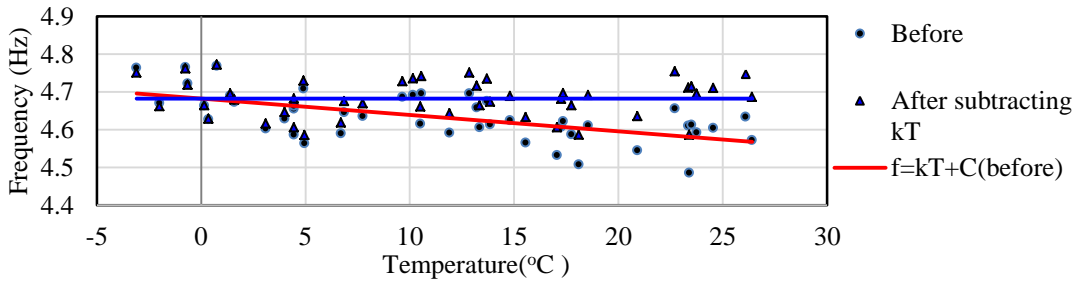


Fig. 7 Frequency before and after removing temperature's influence

should be realized the influences could be quite high if the boundary conditions are not fixed. For more references, see Li *et al.* (2010) and Li *et al.* (2018). A general plot of the estimated values of frequency and damping ratio is shown in Fig. 4. The calculated damage index as introduced in Eq. (7) for the damping ratio is also included in the figure. The optimal order of AR model in this case is determined to be 23. It can be seen the calculated damping ratio is quite high (>5%) which is larger than most of the concrete and steel bridges. The main reason is because of the structural deterioration which has caused a reduction in the critical damping.

Undoubtedly, the calculated modal parameters include some clear environmental influences. For the analyzed bridge in this study, the most significant environmental factor would be the temperature. For instance, the frequency values become very low when the temperature rises in the middle of the year. A more clearer view can be seen in Figs. 5 and 6 which include the trend terms. The trend of observed frequency values shows an obvious seasonal variation while a minimum value is spotted in summer and

a maximum value is spotted in winter. Therefore, the temperature factor has to be considered and a step of removing this effect is required. Following the developed procedures in Section 2.2, the Kalman filter and ARX model are applied in this case. A typical first order ARX model is adopted for the frequency. This is given as follows

$$f = kT + C, \quad (33)$$

where f is the frequency, T is the measured temperature, k is the coefficient and C is the frequency value that is independent of temperature. Therefore, the temperature effect is treated as a regression term in the ARX model. By subtracting the temperature term, kT , from the frequency data, the value of frequency is transformed to a temperature-independent one. An observation of this subtraction can be seen in Fig. 7 which compares the frequency value before and after the regression. The same procedures are applied to all the other modal parameters before a further analysis.

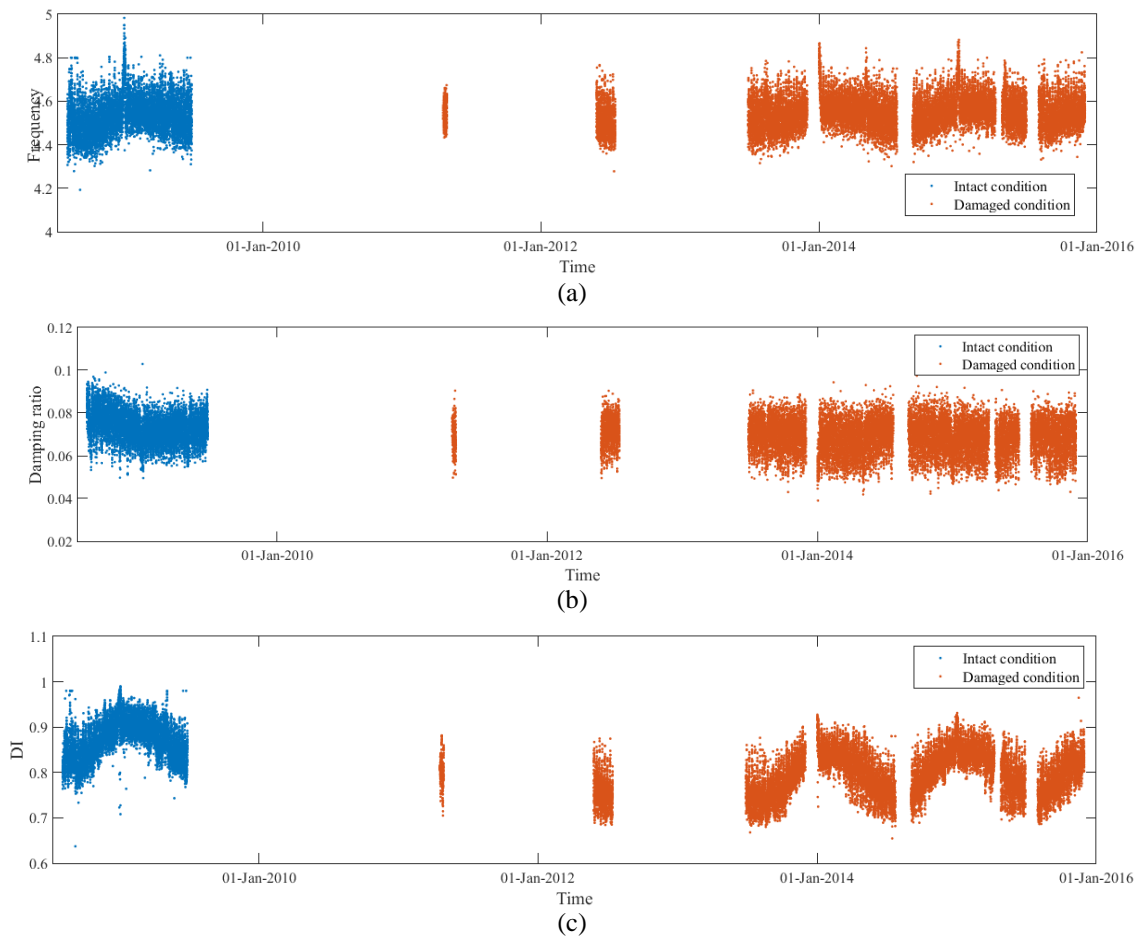


Fig. 8 Calculated (a) frequency, (b) damping ratio and (c) DI from sensor DA-1

After the ARX model is established, the data preconditioning is applied to both the observations before 2010 and after 2010. The results of the calculated structural frequency, damping ratio and **DI** are plotted in Fig. 8. As can be seen, the seasonal variations in these modal parameters are still obvious. This might be induced by some other environmental factors like transportation volume. However, it should be noted that the values of either frequency or damping ratio do not show large deviations after the year 2010. It is hardly to tell there is a structural change by only looking at the frequency and damping ratio values. Even by observing the damage index, the changes are still not obvious. In other words, this implies the structural performance assessment based on the pure calculated modal parameters is not easy.

Therefore, for the further detection of sign of structural damage, the indicators based on multiple observations are utilized. Herein, the traditional applied **MD**, Bayes factor and the newly proposed copula based indicators are used for the structural health assessment. The identified structural frequency and damping ratio are employed to serve as two data sources in this case. In other words, the dependences between the frequency and damping ratio is used in the damage identification. A general feel of the dependences changes between these two modal parameters can be seen in

Fig. 9 which includes the scatter plots frequency and damping ratio before and after 2010. It can be seen that the correlation between frequency and damping ratio decreases after 2010. This implies that there is a change in the dependence between the modal parameters.

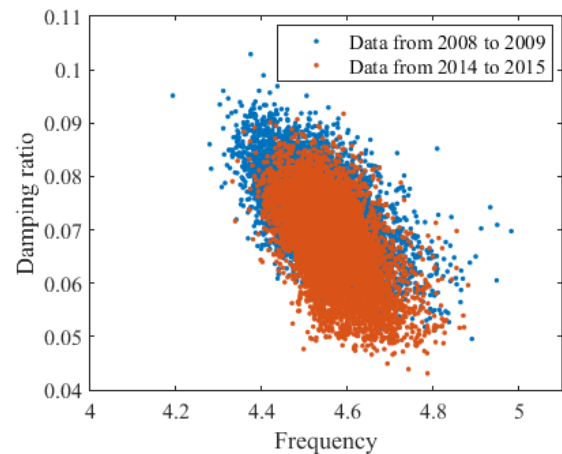


Fig. 9 Comparison of scatterplot of damping ratio and frequency for two periods

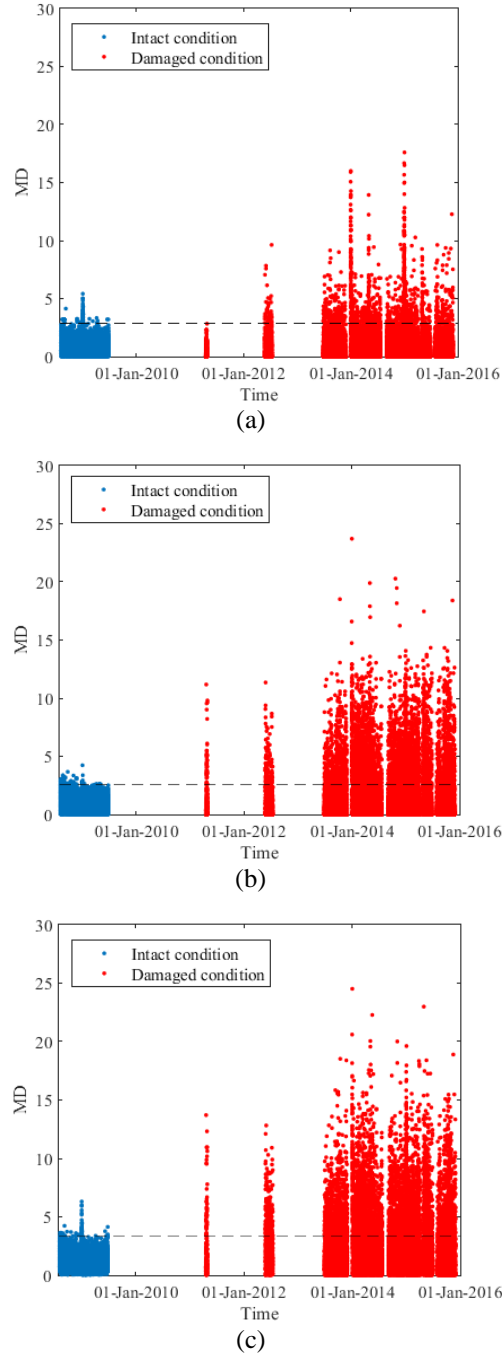


Fig. 10 Calculated MDs of (a) frequency, (b) damping ratio and (c) frequency and damping ratio for sensor DA-1 (black dash line indicates the 95 percentile of the data observed in intact condition)

Following the procedures introduced in Sections 2 and 3, the indicator values for the whole period are calculated and plotted in Figs 10-12. The MDs for individual frequency and damping ratio are also estimated. In order to observe the differences between the intact condition and the condition after the year 2010, the 95-percentile value of the data observed in intact condition is highlighted. For the copula based damage indicator, the commonly applied Gaussian copula is used to calculate the Z statistics. The critical Z statistic value corresponding to p -value of 0.05 is also estimated and highlighted in the figure.

As shown in the figures, **MD** shows very obvious differences between the intact condition and the condition after the year 2010. The Bayes factor is somehow less sensitive compared to **MD**. However, Bayes factor is more sensitive to the national holidays as denoted by the peaks. For example, the calculated Bayes factor for frequency has very large values around the New Year's holidays for bridge structure in both the intact condition and the damaged condition. This is more reasonable compared to **MD** results which do not show very obvious peak values. Compared with these two indicators, copula based Z statistics shows

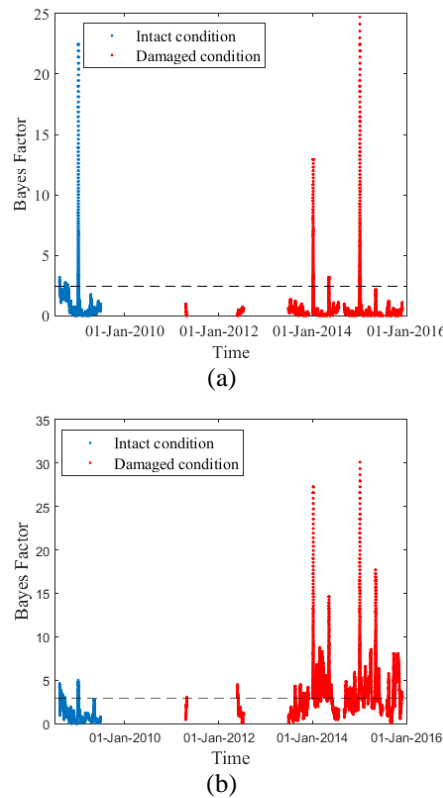


Fig. 11 Calculated Bayes factors of (a) frequency and (b) damping ratio for sensor DA-1 (black dash line indicates the 95 percentile of the data observed in intact condition)

outstanding features. As can be seen from Fig. 12, the values of Z statistics between intact condition and after year 2014 are quite different from each other. Moreover, Z statistics is also very sensitive to the traffic effects due to the national holidays as illustrated by the peak values in Fig. 12.

As shown in the figures, **MD** shows very obvious differences between the intact condition and the condition after the year 2010. The Bayes factor is somehow less sensitive compared to **MD**. However, Bayes factor is more sensitive to the national holidays as denoted by the peaks. For example, the calculated Bayes factor for frequency has very large values around the New Year's holidays for bridge structure in both the intact condition and the damaged condition. This is more reasonable compared to **MD** results which do not show very obvious peak values. Compared with these two indicators, copula based Z statistics shows outstanding features. As can be seen from Fig. 12, the values of Z statistics between intact condition and after year 2014 are quite different from each other. Moreover, Z statistics is also very sensitive to the traffic effects due to the national holidays as illustrated by the peak values in Fig. 12.

This phenomenon can also be observed in Table 2 which compares the percentage of data that is beyond the marginal values (dash lines in Figs 10-12) for different indicators. From the results, it can be seen that MD of bivariate modal parameters are more sensitive compared to MD for individual modal parameter. The results from Bayes factor

shows that damping ratio is more sensitive to the structural damages. Nevertheless, the copula based Z statistics is the most sensitive indicator which produces an anomalous data rate of 71%. The main reason that Z statistic can outperform the other approaches is because of the copula characteristics. Not like the traditional approaches, which can only detect changes in individual parameters (Bayes factor) or joint relationships (**MD**), the copula model can characterize both the individual behavior and the complex dependencies. On one hand, the marginal functions can handle the individual variable's behaviors. On the other hand, the copula function can feature the dependence characteristics. The copula based Z statistic indicator is a complete reflection of all the statistical properties of the multivariate data. Therefore, the copula based damage identification approach could provide a full utilization of the multivariate data.

These results provide us an opportunity to utilize the copula concept for the structural damage detection. Implementing such statistic in the long term structural health monitoring is not straightforward and is still in the development stage. The key features of the copula model can help to formulate a more sensitive damage indicator from the observed multivariate data, whilst the traditional approaches cannot. The interpretation of the copula based indicator result is analog to the interpretation of other damage sensitive indicators. An additional feature of the proposed copula approach is a direct reflection of the dependences contained in the variables.

Table 2 Comparison of different indicators in assessing the data after 2010

Indicator name		MD for frequency	MD for damping ratio	MD for frequency & damping ratio	Bayes factor for frequency	Bayes factor for damping ratio	Copula Z statistic
Detected data	anomalous rate	38%	55%	62%	9%	42%	71%

Conversely, acceptable Z statistic values can be directly determined based on specified structural damage scenarios. The future work will be focusing on the development of a real time damage detection approach based on the copula approach.

5. Conclusions

In this study, feasibility of using copula theory in identification for health monitoring of an old deteriorated bridge structure is investigated through a real case study. A copula based feature sensitive indicator is proposed to detect the structural changes from sensor data. The indicator is developed based on the statistical properties between different modal parameters. Based on this concept, the framework of copula based long term structural health monitoring is developed. This covers three steps including data collection, data pre-processing and damage assessment. The proposed framework is then applied to a real bridge located in Japan for detecting structural changes over eight years of time. The case study demonstrated that the changes in dominant frequency and damping ratio of the bridge were well identified using the measured data. It shows that the copula based damage indicator is the most robust one among all the indicators for detecting the anomalous data. Moreover, the copula based damage indicator is found to be able to detect a sign of dependence changes among the modal parameters. The identification of damages of individual modal parameters is here replaced by an identification of the dependences changes among the modal parameters. Feasibility of the analysis regarding the successful damage detection rate is given in all cases.

Acknowledgments

This study is partly sponsored by Japanese Society for the Promotion of Science (JSPS) for the Grant-in-Aid for Scientific Research (B) under the project No. 16H04398. The first author, Yi Zhang, is sponsored by "The JSPS Postdoctoral Fellowship for Foreign Researchers" Program. Such financial aids are gratefully acknowledged.

References

- Abdelghani, M. and Friswell, M.I. (2007), "Sensor validation for structural systems with multiplicative sensor faults", *Mech. Syst. Signal Pr.*, **21**(1), 270-279.
- Bishop, C.M. (2006), *Pattern recognition and machine learning*, Springer: New York.
- Chang, K.C. and Kim, C.W. (2016), "Modal-parameter identification and vibration-based damage detection of a damaged steel truss bridge", *Eng. Struct.*, **122**, 156-173.
- Cherubini, U., Luciano, E. and Vecchiato, W. (2004), *Copula Methods in Finance*, The Wiley Finance Series. Wiley.
- Dilena, M., Morassi, A. and Perin, M. (2011), "Dynamic identification of a reinforced concrete damaged bridge", *Mech. Syst. Signal Pr.*, **25**(8), 2990-3009.
- Genest, C. and Favre, A. (2007), "Everything you always wanted to know about copula modeling but were afraid to ask", *J. Hydraul. Eng. -ASCE*, **132**(4), 347-368.
- Genest, C. and Mackay, J. (1986), "The joy of copulas - bivariate distributions with uniform marginals", *Am. Statistician*, **40**(4), 280-283.
- Genest, C. and Neslehova, J. (2007), "A primer on copulas for count data", *Astin Bull.*, **37**(2), 475-515.
- Genest, C. and Rivest, L.P. (1993), "Statistical-inference procedures for bivariate archimedean copulas", *J. Am. Statist. Assoc.*, **88**, 423, 1034-1043.
- Hoag, A., Hoult, N.A., Take, W.A., Moreu, F., Le, H. and Tolikonda, V. (2017), "Measuring displacements of a railroad bridge using DIC and accelerometers", *Smart Struct. Syst.*, **19**(2), 225-236.
- Joe, H. (2015), *Dependence modeling with copulas*, CRC Press, New York.
- Kim, C.W., Iseimoto, R., Sugiura, K. and Kawatani, M. (2013), "Structural fault detection of bridges based on linear system parameter and MTS method", *J. JSCE*, **1**, 32-43.
- Kim, C.W., Chang, K.C., Kitauchi, S. and McGettrick, P.J. (2016), "A field experiment on a steel girder-truss bridge for damage detection utilizing vehicle-induced vibrations", *Struct. Health Monit.*, **5**(2), 174-192.
- Kitagawa, G. and Gersch, W. (1984), "A smoothness priors-state space approach to the modeling of time series with trend and seasonality", *J. Amer. Statist. Assoc.*, **79**: 378-389.
- Li, H.N., Li, D.S. and Song, G.B. (2004), "Recent applications of fiber optic sensors to health monitoring in civil engineering", *Eng. Struct.*, **26**(11), 1647-1657.
- Li, Z., Feng, D., Feng, M.Q. and Xu, X. (2017), "System identification of the suspension tower of Runyang Bridge based on ambient vibration tests", *Smart Struct. Syst.*, **19**(5): 523-538.
- Li, Z.L., Li, A. and Zhang, J. (2010), "Effect of boundary conditions on modal parameters of the Run Yang Suspension Bridge", *Smart Struct. Syst.*, **6**(8), 905-920.
- Li, Z.L., Feng, M., Luo, L., Feng, D. and Xu, X. (2018), "Statistical analysis of modal parameters of a suspension bridge based on Bayesian spectral density approach and SHM data", *Mech. Syst. Signal Pr.*, **98**, 352-367.
- Mutlib, N.K., Baharom, S.B., El-Shafie, A. and Nuawi, M.Z. (2016), "Ultrasonic health monitoring in structural engineering: buildings and bridges", *Struct. Control Health Monit.*, **23**(3), 409-422.
- Nair, K.K., Kiremidjian, A.S. and Law, K.H. (2006), "Time series-based damage detection and localization algorithm with application to the ASCE benchmark structure", *J. Sound Vib.*, **291**, 349-368.
- Nelson, R.B. (2006), *An introduction to Copulas*, Springer, New York.

- Omenzetter, P. and Brownjohn, J.M.W. (2006), "Application of time series analysis for bridge monitoring", *Smart Mater. Struct.*, **15**(1), 129-138.
- Park, H.S., Lee, H.M., Adeli, H. and Lee, I. (2007), "A new approach for health monitoring of structures: Terrestrial laser scanning", *Comput. -Aided Civil Infrastruct. Eng.*, **22**(1), 19-30.
- Soh, C.K., Tseng, K.K.H., Bhalla, S. and Gupta, A. (2000), "Performance of smart piezoceramic patches in health monitoring of a RC bridge", *Smart Mater. Struct.*, **9**(4), 533-542.
- Sung, H.J., Do, T.M., Kim, J.M. and Kim, Y.S. (2017), "Long-term monitoring of ground anchor tensile forces by FBG sensors embedded tendon", *Smart Struct. Syst.*, **19**(3), 269-277.
- van Overschee, P. and de Moor, B.L. (1996), *Subspace Identification for Linear Systems*, US: Springer.
- Vanem, E. (2016), "Joint statistical models for significant wave height and wave period in a changing climate", *Mar. Struct.*, **49**, 180-205.
- Varouchakis, E.A. (2016), "Modeling of temporal groundwater level variations based on a Kalman filter adaptation algorithm with exogenous inputs", *J. Hydroinform.*, DOI: 10.2166/hydro.2016.063
- Xiong, L., Jiang, C., Xu, C.Y., Yu, K. and Guo, S. (2015), "A framework of change-point detection for multivariate hydrological series", *Water Resources Res.*, **51**, 8198-8217.
- Yan, A.M., De Boe, P. and Golinvall, J.C. (2004), "Structural damage diagnosis by Kalman model based on stochastic subspace identification", *Struct. Health Monit.*, **3**(2), 103-119.
- Zhang, Y. (2015), "Comparing the robustness of offshore structures with marine deteriorations — a fuzzy approach", *Adv. Struct. Eng.*, **18**(8), 1159-1172.
- Zhang, Y. and Lam, J.S.L. (2015), "Reliability analysis of offshore structures within a time varying environment", *Stochast. Environ. Res. Risk Assess.*, **29**(6), 1615-1636.
- Zhang, Y. and Lam, J.S.L. (2016a), "A copula approach in the point estimate method for reliability engineering", *Q. R. Eng. Int.*, **32**(4), 1501-1508.
- Zhang, Y. and Lam, J.S.L. (2016b), "Estimating economic losses of industry clusters due to port disruptions", *T. Res. Part A: Policy Pract.*, **91**, 17-33.
- Zhang, Y., Beer, M. and Quek, S.T. (2015), "Long-term performance assessment and design of offshore structures", *Comput. Struct.*, **154**, 101-115.
- Zhang, Y., Kim, C.W. and Tee, K.F. (2017a), "Maintenance management of offshore structures using Markov process model with random transition probabilities", *Struct. Infrastruct. Eng.*, **13**(8), 1068-1080.
- Zhang, Y., Kim, C.W., Tee, K.F. and Lam, J.S.L. (2017b), "Optimal sustainable life cycle maintenance strategies for port infrastructures", *J. Cleaner Production*, **142**, 1693-1709.

Table A.1 Examples of Archimedean copulas

Copula	Bivariate Formula $C_\gamma(u, v)$	Generating Function $\phi_\gamma(t)$	$\gamma \in$
Gumbel	$\exp\left\{-\left[(-\ln u)^\gamma + (-\ln v)^\gamma\right]^{\frac{1}{\gamma}}\right\}$	$(-\ln t)^\gamma$	$[1, +\infty)$
Frank	$\frac{1}{\gamma} \ln \left(1 + \frac{(e^{u^\gamma} - 1)(e^{v^\gamma} - 1)}{e^\gamma - 1}\right)$	$\ln \frac{e^{\gamma t} - 1}{e^\gamma - 1}$	$(-\infty, +\infty)$
Clayton	$(u^{-\gamma} + v^{-\gamma} - 1)^{-\frac{1}{\gamma}}$	$t^{-\gamma} - 1$	$(1, +\infty)$

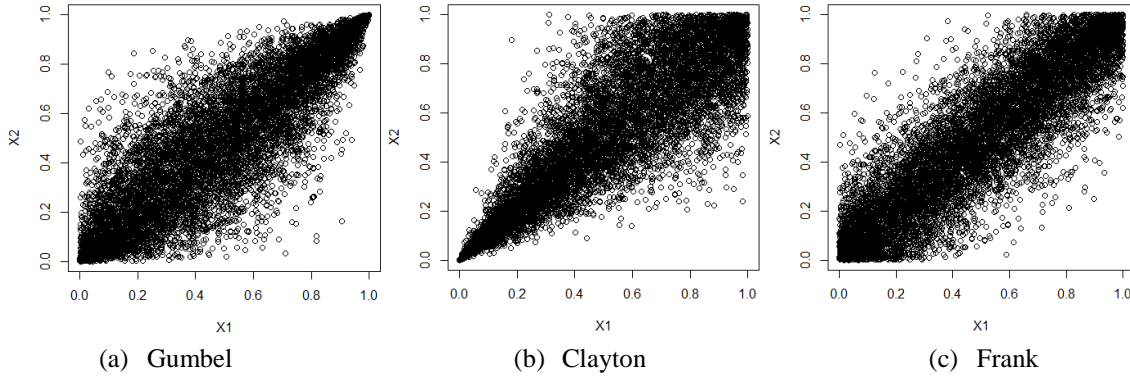


Fig. A.1 Comparison of different bivariate copulas for the same correlation coefficient equals to 0.7

Appendix A Examples of copulas

There are many copula families and classes available in the literature. Each one could characterize specific dependences of data. Most of these copulas are built for a bivariate case which may be extended to a multivariate model quite easily by the mathematical transformations. In most of the applications, the Archimedean copula are frequently used in practice (Genest and Mackay 1986).

Archimedean copulas

The family of Archimedean copulas includes a wide range of parametric copula classes. They are widely applied in data modeling because of their feature and ease of construction. Generally, an n -dimensional Archimedean copula is initially constructed from algebraic method utilizing the generating function $\phi(\cdot)$:

$$C_{\text{Archimedean}}(u_1, \dots, u_n; \theta) = \phi^{[-1]}(\phi(u_1; \theta) + \dots + \phi(u_n; \theta); \theta) \quad (\text{A1})$$

where $\phi: [0, 1] \times \Theta \rightarrow [0, \infty)$ is a strictly decreasing convex function with $\phi(1)=0$. θ is the copula function parameter with its domain Θ . $\phi^{[-1]}$ is the pseudo-inverse of ϕ which is defined by

$$\phi^{[-1]}(t; \theta) = \begin{cases} \phi^{-1}(t; \theta) & \text{if } 0 \leq t \leq \phi(0; \theta) \\ 0 & \text{if } \phi(0; \theta) \leq t \leq \infty \end{cases} \quad (\text{A2})$$

Therefore, the construction of a multivariate copula model is relying on the generating functions used in Eq.

(A.2). There are many well-known Archimedean copula families which can be found in the literature (Genest and Rivest 1993). The generating functions of the most popular single parameter Archimedean copulas including Gumbel, Frank and Clayton are given in Table A.1.

An illustration of these Archimedean copulas is provided by means of bivariate scatter plots as depicted in Fig. A.1. It can be seen from the plot each Archimedean copula characterizes a special tail dependency for the bivariate data. For example, the Gumbel copula is an excellent candidate model for data having stronger dependencies at large values compared to low values, whereas the Clayton copula is good at characterizing data exhibiting strong low value dependencies. On the other hand, the Frank copula is considered quite appropriate for data having relatively weak dependencies at both tails. For the structural health monitoring problems, in some modal parameters, stronger dependencies are observed at the extremes. For instance, frequency values of different modes usually show high correlations with each other when structural damages occurs, and thus one can expect the data to show stronger dependencies between the frequencies when there are some changes in the structural conditions. Hence, to be suitable, a copula model for this case is robust enough to capture these characteristics observed in the data.