

## Online damage detection using pair cointegration method of time-varying displacement

Cui Zhou<sup>1</sup>, Hong-Nan Li<sup>\*1</sup>, Dong-Sheng Li<sup>1</sup>, You-Xin Lin<sup>2</sup> and Ting-Hua Yi<sup>1</sup>

<sup>1</sup>*Faculty of Infrastructure Engineering, Dalian University of Technology, Dalian, China*

<sup>2</sup>*Guangdong Electrical Company, Guangzhou, China*

(Received July 12, 2012, Revised December 20, 2012, Accepted December 26, 2012)

**Abstract.** Environmental and operational variables are inevitable concerns by researchers and engineers when implementing the damage detection algorithm in practical projects, because the change of structural behavior could be masked by the conditions in a large extent. Thus, reliable damage detection methods should have a virtue of immunity from environmental and operational variables. In this paper, the pair cointegration method was presented as a novel way to remove the effect of environmental variables. At the beginning, the concept and procedure of this approach were introduced, and then the theoretical formulation and numerical simulations were put forward to illustrate the feasibility. The jump exceeding the control limit in the residual indicates the occurrence of damage, while the direction and magnitude imply the most potential damage location. In addition, the simulation results show that the proposed method has strong ability to resist the noise.

**Keywords:** structural health monitoring; damage detection; environmental variable; pair cointegration; time series; data-based model

### 1. Introduction

Advanced technologies in sensors, communication, data acquisition and processing have made it possible to implement continuous structural health monitoring (SHM) system, which could guide the damage detection, state evaluation and conditional maintenance of large-scale structures. The core of a SHM system is the diagnostic algorithms for detection of the presence, location and extent of the damage. However, development of SHM in civil engineering is constrained heavily by the poor performance of damage detection methods applied in the practical project.

Up to date, considerable research efforts have been made in order to enhance the validity of damage detection, in which the vibration-based damage identification techniques have shown to be the most promising methods (Doebeling *et al.* 1998, Fan and Qiao, 2011). But many serious challenges still stand when applied to real structures, for it is often impractical to excite full scale structures in a controlled way, and natural frequencies are inherently insensitive to local damages. Besides the problems aforementioned, mode shapes are easily spoiled by various sources of environmental and operational noise. There are also some alternative fingerprints to aid the

---

\*Corresponding author, Professor, E-mail: hnli@dlut.edu.cn

damage detection, such as the static displacement, which is immune to the local disturbance and easily accessible via high-resolution instruments. The measurement of displacements has become an indispensable component of SHM systems (Ghrib *et al.* 2012), especially for large-span structures. Wang *et al.* (2012) proposed Symmetrical Displacement Difference Index to identify the damage in bridges using static displacements. Lee and Eun (2008) utilized the displacement curvature to locate successfully a simple cantilever beam. Dewangan (2011) presented a row-echelon form of matrix based static displacement and force data, which was able to predict the damage existence in the structure with only few measurements.

Like the fundamental frequencies, the drawback of the displacement is also weak resistance to the variation induced by environmental and operational conditions, especially the dominant temperature. Based on a field test, Moorthy and Roeder (1992) reported that the bridge deck showed a significant expansion as temperature increased from both the analytical model and the measured values. Doebling and Farrar (1997) revealed that the first mode frequency of the Alamosa Canyon Bridge varied about 5% during one day. Askegaard and Mossing (1988) recorded approximately 10% seasonal changes for a three-span RC footbridge. Kim *et al.* (2003) published that the measured natural frequencies of a 46 m long simply supported plate girder bridge decreased up to 5.4% as a result of heavy traffic. For more detailed information, please refer to Sohn (2007).

In the continuous SHM system, structural static response signal increment can be roughly expressed as

$$\Delta S = S_T + S_p + S_c + S_D + S_R \quad (1)$$

where  $S_T$  is the temperature effect,  $S_p$  is the load effect,  $S_c$  is the concrete shrinkage and creep effect,  $S_D$  is the structural damage effect and  $S_R$  is the system test error. It is clear that the structural response signal increment is the superposition of multiple effects.

Many features extracted from displacements for their sensitivity to damage are also sensitive to changes caused in the structure by environmental and operational conditions, resulting in the damage information indecipherable. It can be inferred that the aforementioned displacement-based methods would be hassled by the problem of environmental variability. For practical application, it becomes an imperative task to divorce the effect of environment and live load from that caused by damage. Several methods have been proposed to divorce the environmental and operational variables from the damage indices. Peeters *et al.* (2000) used the ARX model to filter out the temperature effects from the measured frequencies. Fritzen *et al.* (2003) modified an existing subspace-based identification method for temperature compensation. Sohn (2007) made a comprehensive review on removing the environmental and operational variables from the identified eigenfrequencies. Figueiredo *et al.* (2010) compared four machine learning methods, including auto-associative neural network, factor analysis, Mahalanobis distance, and singular value decomposition, to detect the structural damage in the presence of operational and environmental variables. Loh *et al.* (2011) combined the singular spectrum analysis with the auto regressive model to form a damage identification algorithm based on the continuous monitoring of the dam static deformation. Zhou *et al.* (2011) presented a back-propagation neural network-based method for elimination the temperature effect in vibration-based structural damage alarm.

In current work, the pair cointegration approach is suggested as a suitable methodology for exploring the damage trace from the noise time series polluted by environmental and operational variables. This original idea has been applied by Chen *et al.* (2009) in statistical process control and by Cross *et al.* (2010) in removal of environmental trends from the identified eigenfrequencies

series, but they only addressed the simplest level of damage identification, i.e., whether damage is present or not. Here, this method will be extended to localize the damage using temperature-dependent displacement time series. After purging the influence of temperature by the pair cointegration of neighbored displacement series, the up and low control limit of the residue will be set up through the statistical analysis, and thereafter the early warning could be determined in the online SHM system.

## 2. Fundamental of cointegration

### 2.1 Outline

Engle and Granger (1987) first introduced the concept of cointegration method, and it has become an indispensable step in the analysis of non-stationary time series in econometrics. The underlying idea is that even if two variables or more are non-stationary, there can be a combination of them to create a new stationary one. This definition leads to an interesting interpretation as the variables can then be represented to have a stable relationship known as a long-term equilibrium, that is, they share common stochastic trend.

If a non-stationary process variable  $Y$  becomes stationary after differencing  $d$  times, it is said to be integrated of order  $d$ , denoted as  $Y \sim I(d)$  (Box *et al.* 1994), Engle and Granger considered a set of  $d$ -integrated non-stationary variables,  $Y^T = (Y_1, Y_2, \dots, Y_n) \in \mathbb{R}^n$ , that holds a long-term dynamic equilibrium if and only if

$$\beta_1 Y_1 + \beta_2 Y_2 + \dots + \beta_n Y_n = \varepsilon \quad (2)$$

where  $\varepsilon \in \mathbb{R}$  is the residual sequence and  $\beta^T = (\beta_1, \beta_2, \dots, \beta_n) \in \mathbb{R}^n$  is a cointegrating vector.

If the equilibrium is to describe the long-term relationship between the non-stationary variables, then the residual sequence must be stationary, i.e., integrated of order 0,  $\varepsilon \sim I(0)$ . Therefore, cointegration modeling targets to identify coefficient vectors such that the residual sequence becomes stationary.

There are generally two steps to find the coefficient vectors. First, ascertain the order of integration of the time series, because cointegration tests only work when the non-stationary variables stay in the same order of integration. For this procedure, unit root test, specifically the augmented Dickey-Fuller (ADF), is employed. Once the variables are integrated to the same order, it is ready to find out the cointegrating vector. According to the test object, there are two types of cointegration tests: one is based on the regression residual, such as the EG test and CRDW test (Gao 2009), and another is based upon the regression factor, such as the Johansen test. The Johansen test will be used here, owing to its increased sophistication.

### 2.2 Augmented dickey-fuller test

The unit root test is integrated with the Augmented Dickey-Fuller test (Dickey and Fuller 1979). If a time series has a unit root, it will be non-stationary. It is clearly expressed by the autoregressive model AR(1).

$$Y_t = \rho Y_{t-1} + e_t, t = 1, 2, 3, \dots, n \quad (3)$$

where  $\rho$  is a real number, and  $\{e_t\}$  represents a sequence of independent normal random variables with mean zero and variance  $\sigma^2$ . The time series  $Y_t$  converges to a stationary time series if  $|\rho| < 1$ . When  $|\rho| = 1$ , the time series develops into non-stationarity. As for  $|\rho| > 1$ , the time series is not stationary and the variance of the time series grows exponentially. The hypothesis that  $\rho = 1$  is of some interest in the application, because it corresponds to the hypothesis that it is appropriate to transform the time series by differencing (Dickey and Fuller 1981).

In more general case, consider AR( $p$ ) model

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t \quad (4)$$

Eq. (4) subtracts  $Y_{t-1}$  from both sides. With the method of adding and deleting terms, get

$$\Delta Y_t = \eta Y_{t-1} + \sum_{i=1}^{p-1} \beta_i \Delta Y_{t-i} + e_t \quad (5)$$

where

$$\eta = \sum_{i=1}^p \phi_i - 1 \quad \beta_i = - \sum_{j=i+1}^p \phi_j \quad (6)$$

After calculating the parameters in the model, the null hypothesis of the ADF test includes  $H_0: \eta = 0$ ,  $H_1: \eta \neq 0$ .

Assuming that  $H_0$  be accepted then  $Y_t$  has a unit root, and  $\Delta Y_t$  is a stationary sequence, i.e.,  $Y_t$  is an  $I(1)$  non-stationary sequence. If  $H_0$  is rejected, keep on testing  $\Delta Y_t$ . If it is accepted then,  $Y_t$  represents an  $I(2)$  non-stationary sequence. If  $H_0$  is rejected again, repeat the test for  $d = 3$ , etc. until  $H_0$  is accepted.

The ADF test procedure is, therefore, to estimate the parameters in Eq. (5) by the least-square methods and then test the null hypothesis  $\eta = 0$ . The test statistic term is

$$t_\eta = \frac{\hat{\eta}}{\sigma_\eta} \quad (7)$$

where  $\hat{\eta}$  is the least-square estimate of  $\eta$ , and  $\sigma_\eta$  is its standard deviation, which should be evaluated and compared with the critical values from Dickey–Fuller (DF) tables (Fuller 1996).

Additional hypotheses and test statistics are needed if the model is extended to include intercept or trends (or both). For the extended time series form

$$\Delta Y_t = \eta Y_{t-1} + \sum_{i=1}^{p-1} \beta_i \Delta Y_{t-i} + \mu + \gamma t + e_t \quad (8)$$

The null hypothesis for the time series to be  $I(1)$  should be extended to include  $\mu = 0, \gamma = 0$ . More details for these specific cases can be found in Ref. (Fuller 1996).

Having ascertained the order of non-stationarity of each process variable of interest, an attempt

to build a stationary residual through combination of those variables integrated to the same order becomes easy to implement. The Johansen procedure is outlined below for this purpose.

### 2.3 Johansen test

Johansen and Juselius proposed a cointegration test based on the vector auto-regression (VAR) which represents a multivariate AR time series model (Johansen and Juselius 1990). The test is only applicable to a set of non-stationary variables integrated of  $I(1)$ . In that case, it can be found that the independent cointegrating vectors create the most stationary linear combination. In general, a cointegration model of  $N$  non-stationary variables is given by the following linear combination

$$\varepsilon = \beta^T Y \quad (9)$$

where  $\beta^T Y$  could be interpreted as the “equilibrium” of a dynamic system, and  $\varepsilon$  as the vector of “equilibrium errors”.

Consider the VAR( $p$ ) model of order  $p$

$$Y_t = \sum_{i=1}^p \Phi_i Y_{t-i} + e_t \quad (10)$$

where  $\Phi$  is the parameter matrix of the multivariate AR model. A vector error correction model (VECM) can be obtained by subtracting  $Y_{t-1}$  from both side of Eq. (10)

$$\Delta Y_t = \Pi Y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + e_t \quad (11)$$

where

$$\Pi = \sum_{i=1}^p \Phi_i - I, \Gamma_i = -\sum_{j=i+1}^p \Phi_j \quad (12)$$

Through differencing procedure, the variables  $\Delta Y_t, \Delta Y_{t-i} (i=1, 2, \dots, p)$  in Eq. (11) are variables of  $I(0)$ , so only if  $\Pi Y_{t-1}$  are variables of  $I(0)$ , each variables in  $Y_{t-1}$  are cointegration and  $\Delta Y_t$  is stationary. Whether variables in  $\Pi Y_{t-1}$  are cointegration, it mainly depends on the rank of  $\Pi$ .

There are three possible cases:

- (1) Rank( $\Pi$ )= $n$ , i.e., the matrix  $\Pi$  has a full rank, implying that  $Y_t$  is stationary.
- (2) Rank( $\Pi$ )=0, i.e., the matrix  $\Pi$  is a null matrix and Eq. (11) corresponds to a traditional differenced equation.
- (3)  $0 < \text{Rank}(\Pi) = r < n$ , indicating that there are  $n \times r$  matrices  $\alpha$  and  $\beta$  to satisfy  $\Pi = \alpha \beta^T$ .

We can write Eq. (11) in the following form:

$$\Delta Y_t = \alpha \beta^T Y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + e_t \quad (13)$$

The residual vector  $\varepsilon_{t-1}$  can be obtained as follows

$$\varepsilon_{t-1} = (\alpha^T \alpha)^{-1} \alpha^T (\Delta Y_t - \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} - e_t) \quad (14)$$

Eq. (14) indicates that  $\varepsilon_{t-1}$  is integrated of  $I(0)$ , i.e., stationary sequences. Thus, the linear combination of  $\beta^T Y_{t-1}$  is  $I(0)$ , and the column vectors of  $\beta$  are cointegrating vector matrix, and  $r$  is the number of cointegrating vectors.

The keystone of Johansen procedure is to change the cointegration test of  $Y_t$  to that of the matrix  $\Pi$ . The rank of the matrix  $\Pi$  is equal to the number of its non-zero eigenvalues; therefore, the cointegration test can estimate the number of the non-zero eigenvalues.

Assuming that the eigenvalues of matrix  $\Pi$  are  $\lambda_1 > \lambda_2 > \dots > \lambda_r \neq 0$  in descending order, then  $r$  cointegration vectors could be obtained, and the other  $n-k$  non-cointegration combination, i.e.,  $\lambda_{r+1}, \dots, \lambda_k$ , should be zero. It is called the trace test. The null and alternative hypotheses are

$$\begin{aligned} H_{r0} : \lambda_{r+1} &= 0 \\ H_{r1} : \lambda_{r+1} &> 0, r = 0, 1, \dots, n-1 \end{aligned} \quad (15)$$

The corresponding test statistic is

$$\eta_r = -T \sum_{i=r+1}^n \ln(1 - \lambda_i), r = 0, 1, \dots, n-1 \quad (16)$$

where  $\eta_r$  is called eigenvalue trace statistic. Test the significance of these statistic: when  $\eta_0$  is not significant (i.e.,  $\eta_0$  is smaller than the critical value of Johansen distribution at some significant level) and  $H_{00}$  is accepted, indicating that  $n$  unit roots and non-cointegrating vector. Otherwise,  $\eta_0$  is significant and  $H_{00}$  is rejected, revealing that there are at least one cointegrating vector and the continued significant test of  $\eta_1$  should be proceeded.

### 3. Theoretical formulation

In the current section, a theoretical formulation is provided to illustrate how the proposed pair cointegration method works, i.e., not only to identify the existence of the damage, but also to localize the damage through the pair-cointegration of neighbored displacements. A hinged-hinged beam was used to demonstrate for simplicity.

The time-invariant analytical model of a both-ends hinged uniform beam is shown in Fig. 1. Assuming that the beam be loaded by a concentrated force  $P$ , and a crack exists at point  $j$  on the left side. The deflection of the overall beam before and after damage (Li 2009) can be expressed as

$$D_u(x) = \frac{Px}{48EI}(3L^2 - 4x^2), 0 \leq x \leq \frac{L}{2} \quad (17)$$

$$D_d(x) = \frac{P}{EI} \left\{ \frac{(2x-L)^3}{48} U(x - \frac{L}{2}) - \frac{1}{12} [x^3 + 6\alpha x_j(x - x_j)U(x - x_j)] \right. \\ \left. + \frac{1}{12} [\frac{3L^2}{4} + \frac{6\alpha x_j(L - x_j)}{L}] x \right\}, 0 \leq x \leq L \quad (18)$$

where  $D_u(x)$  and  $D_d(x)$  are the deflections in the undamaged and damaged case;  $L$  is the total length of the beam;  $x_j$  is the location of damage, and  $\alpha$  is the additional flexibility coefficient that describes the damage degree;  $U(x)$  indicates the unit step function, showing discontinuity at  $x_j$ , defined as  $U(x - x_j) = 0$  for  $x < x_j$ ,  $U(x - x_j) = 1$  for  $x \geq x_j$ .

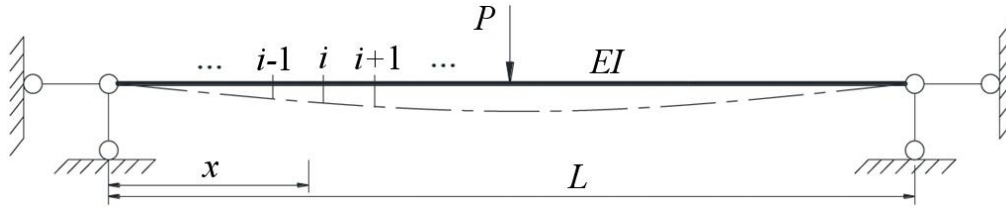


Fig. 1 A hinged-hinged supported beam

For pair cointegration method, the feature we truly care is that the displacements of adjacent positions. Assume that  $D_i$  and  $D_{i+1}$  are the static displacements at adjacent points  $i$ ,  $i+1$  respectively, which is called cointegration pair (CP) for convenience. Then, the pair cointegration residual is constructed by

$$\varepsilon_i = a_i D_i + b_i D_{i+1}, i = 1, 2, \dots \quad (19)$$

where  $a_i$  and  $b_i$  are the cointegration coefficients. Once  $a_i$  and  $b_i$  are determined, the damage could be identified through the relative change of  $\varepsilon_i$ . It is true that  $D_{u,i+1}$  is greater than  $D_{u,i}$  for the left half span. In this condition, assuming  $a_i < 0$ ,  $b_i > 0$  and  $|a_i| > |b_i|$ , i.e., the smaller displacement corresponds to negative and absolute large number. It is reasonable that only in this way can the residual be near zero.

When the damage arises, the displacement change is

$$\Delta D = D_d - D_u = \frac{P}{2EI} [-\alpha x_j (x - x_j) U(x - x_j) + \frac{\alpha x_j x (L - x_j)}{L}], 0 \leq x \leq \frac{L}{2} \quad (20)$$

From Eq. (20),  $\Delta D$  increases with the augmentation of  $x$  when  $x \leq x_j$ ; while  $x > x_j$ ,  $\Delta D_i$  decreases with the augmentation of  $x$ .

The jump series can be described as

$$\Delta \varepsilon_i = a_i \Delta D_i + b_i \Delta D_{i+1} \quad (21)$$

then the problem will be separated into three conditions. Firstly, if the damage is located at the right side of both sensors, mathematically  $x_j > x_{i+1}$ ,  $\Delta \varepsilon_i$  will be a distinct positive number; secondly, the damage lies between the sensors, i.e.,  $x_i < x_j < x_{i+1}$ ,  $\Delta \varepsilon_i$  might stay stationary or an undistinguishable positive or negative jump; in the last case, when  $x_j < x_i$ ,  $\Delta \varepsilon_i$  will be a distinct negative number. In addition, the closer to the damage, the more significant the amplitude of the jump becomes, which is also a feature to help the diagnosis.

In conclusion, CP residuals of two neighbored displacements sharing common trend is able to capture the existence of damage simply by a jump in the time series. In order to locate the damage, a distributed measuring scheme should be employed; then the opposite jump appearing in the residuals will indicate the location of damage, and the location on which the direction shift appears firstly is the most potential damaged area. For simplicity, the pair cointegration method is illustrated for hinged-hinged beam theoretically, but this method is not confined to this kind of structure; for whatever type the structure is, the neighbored displacements will always show a close correlation, which is the principle the pair cointegration algorithm adopts.

As illustrated in Eq. (1), the static displacement, which possesses nonlinear characteristics inherently, results from very complicated factors, including the possible damage. In order to eliminate all the influence except the damage, the pair cointegration is employed to identify the existence of damage, and to indicate the position where the damage located as well. The main procedures are summarized as follows:

(1) To achieve the feasibility of the monitoring of non-stationary process, training data sets that comprise multiple effects are employed to build the VAR models. These data sets as the reference should be sampled under the normal operation conditions.

(2) To further in the analysis, these monitored data sets are tested by an ADF test to ascertain whether they are of the same order of non-stationarity. Specially, each data set should be integrated order one in order to exert the Johansen procedure. It should be noted that most of SHM data satisfy this requirement (Cross *et al.* 2011). Especially for static displacement, it is inherently integrated order one, for the difference of displacement is structural velocity response which is a stationary series.

(3) When each data set meets the qualification, the Johansen procedure is applied to the monitored data to identify the cointegrating vectors. The eigenvector associated with the largest eigenvalue commonly have the best description of the equilibrium relationship, so that it is adopted as the coefficients that combine the data set.

(4) The residual sequence should be a stationary process if the model acts out the long-term equilibrium in normal condition. In the case of on-line monitoring, the dynamic equilibrium is distorted and do not recover when a disturbance (damage) is occurred. A jump would occur in the residual sequence, and this position indicates the instance when damage occurs.



#### 4. Numerical simulation

In order to illustrate the feasibility of the proposed pair cointegration algorithm for damage detection under varying environmental conditions, a numerical model of a prestressed concrete beam is established using the commercial finite element software ANSYS. Half span of the beam is depicted in Fig. 2 due to its symmetry.

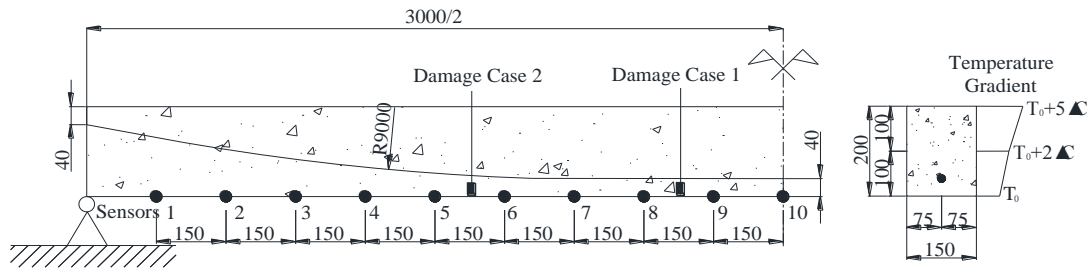


Fig. 2 Prestressed concrete model (units: mm)

The parameters of the beam are as follows: the area, density, and stretching force of the prestressed tendon are  $139 \text{ mm}^2$ ,  $7921 \text{ kg/m}^3$  and  $180 \text{ kN}$ , respectively; the density of the concrete is  $2300 \text{ kg/m}^3$ . Both ends of the beam are constrained by pin joints. The relationships of the elastic modulus of concrete and prestressed tendon versus temperature are shown in Fig. 3 (Yan *et al.* 2005). The concrete is simulated by the concrete element SOLID65, and the prestressed tendon is modeled as LINK8. Totally, there are 1798 elements and 2520 nodes.

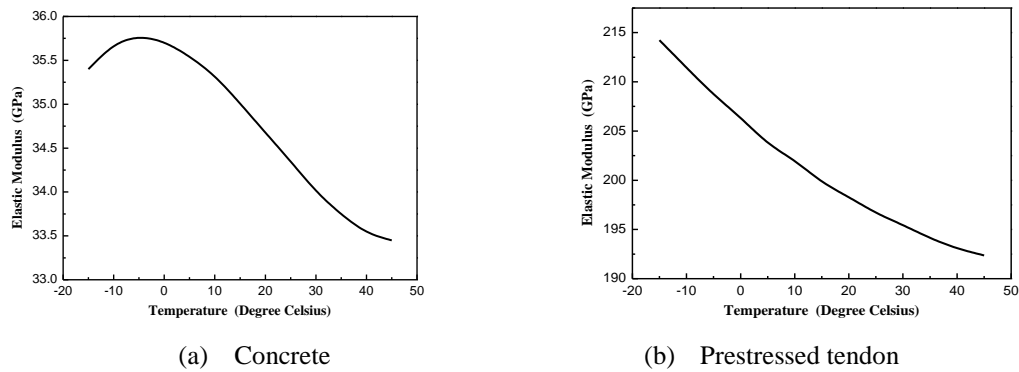


Fig. 3 Temperature-dependent elastic modulus

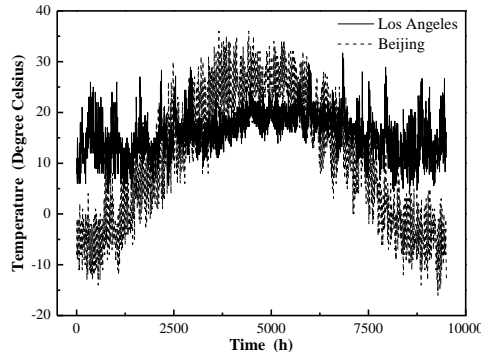


Fig. 4 Temperature history series (Jan. 1, 2011-Jan. 31, 2012)

The temperature dominates the variation of structural behavior via directly affecting the stiffness of a structure, although it can additionally change the boundary conditions of a structure, if foundations, etc. become frozen (Cross *et al.* 2011). In the numerical simulation, the effect of temperature is only considered on displacements calculated by static analysis. Two real temperature history series (shown in Fig. 4) are selected in the analysis: one is the temperature of Los Angeles (T-LA) from Jan. 1, 2011 to Jan. 31, 2012 recorded by a weather station near Los Angeles once per two hours; the other is a similar one in Beijing (T-BJ). The temperature-dependent elastic modulus of concrete exhibits distinct attributes below and up zero, as shown in Fig. 3(a), which is an important factor causing the nonlinearity. It is for this reason that these temperature series, ranging from 4°C to 31.7°C and from -15°C to 36°C, are applied. In the next part, we will show that our cointegration method performs well in spite of the nonlinearity. In addition to the uniform distributed temperature on the structural section, the temperature gradient case is also considered in the model shown in Fig. 2, in which  $T_0$  represented the recorded temperature series, and two linear gradients are assumed, a similar type as recommended by the concrete code (2006). Two temperature series and two kinds of temperature modes are combined to four situations with test the proposed method.

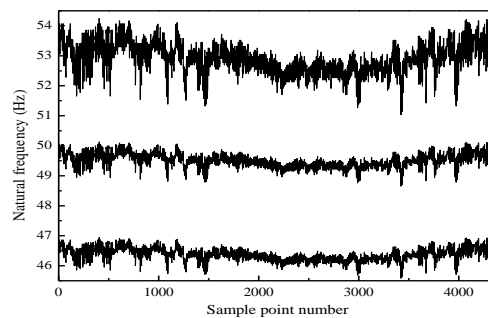


Fig. 5 Natural frequencies of the first three modes

Note: For better display, the second and third mode frequencies are translated.

Two damage cases simulated by killing the element (Fig. 2) are introduced. In damage case 1, the size of the element deactivated is 5 cm×5 cm×15 cm, while 5 cm×7 cm×15 cm for damage case 2. The displacements of sensor locations were extracted for the undamaged and damaged states in the computation. The training set was conducted from Jan. 1, 2011 to Dec. 31, 2011, containing 4380 samples. Damage was introduced on Jan. 1, 2012 and other 372 data points were recorded.

To better illustrate the environmental effect, the following pictures describe the change of the natural frequencies (Fig. 5) and displacements (Fig. 6) with the temperature. The nonstationary properties exhibit clearly, and the variation caused by the environmental factors would submerge the change due to the damage.

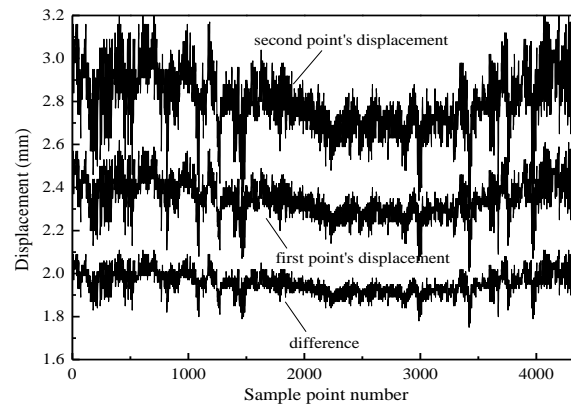


Fig. 6 Time varying displacements and their difference

Note: the second point's displacement is translated by 1.5 mm.

#### 4.1 Simulation 1

The first simulation was carried out in the case of T-LA uniformly distributed temperature, with the damage existing between sensors 8 and 9. The detection results of the damage were depicted in Fig. 7, in which the control limit was determined by  $\pm 3$  times of standard deviation of each training set residual.

It is clear that all the sequences have a quick jump when damage happens. From CP7 to CP9, an obvious opposite direction of jump appears, thus it can be inferred that the damage is located at CP8 as explained in the former section, and it is really where the damage lies. In addition, a gradually changing magnitude emerges in these figures, which can be used to locate the damage as well. Like the aforementioned, a more significant jump will arise near the damage spot. Therefore, the pair cointegration of displacement series could successfully identify when and where the damage occurs.

In previous simulation, the validity of the proposed identification procedure has been substantiated on the basis of exactly measured displacements. However, in the practical experiments or applications, they are all exposed to unmeasurable noise composing an important

source of the error. Hence, the immunity to noise is a required criterion to evaluate the damage detection method. It will be shown that the proposed pair cointegration approach performs well even on the condition of high level noise.

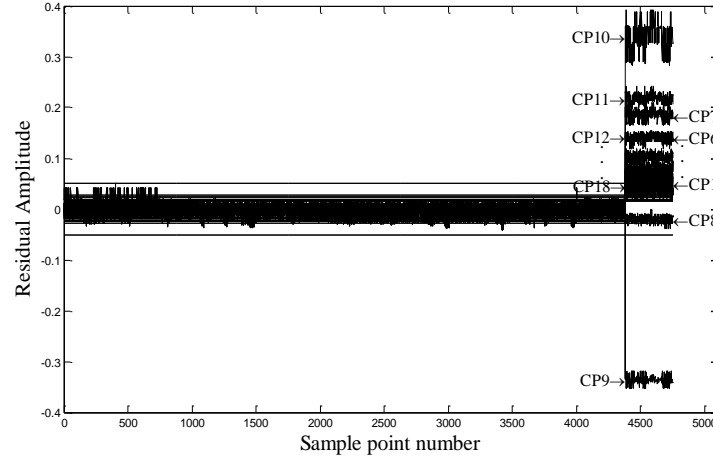


Fig. 7 Detection result of simulation 1 without noise

The noise model added is displayed in Eqs. (22) and (23)

$$S_n = S + \alpha \times A \times randn \quad (22)$$

in which

$$A = \frac{[rms(max) - rms(min)]}{2} \quad (23)$$

where  $S_n$  is the noise signal,  $S$  is the perfect signal,  $\alpha$  is the noise level, rms is short for the root mean square, max and min represent the maximum and minimum envelopes of the perfect displacement series and  $randn$  means normally distributed random variables independent of each other with zero mean and unit variance. Under the same situation, 2.5%, 5%, 7.5% and 10% levels of noise are added to the simulation displacement, respectively. In all the tests, the cointegration works well. For simplicity, only 10% noise results are shared here in Fig. 8. Compared with the cointegration result without simulation noise, the residual of each cointegration pair oscillates more intensely, and the control limit is enlarged correspondingly; however, the large level of noise does not mask the damage information. The pair cointegration of displacement series could still successfully diagnose the damage. Therefore, it is proper to say that the proposed approach has a great degree of noise resistance. However, when the noise level is above 15%, obvious jumps still could be observed but it is hard to distinguish pinpoint damage position.

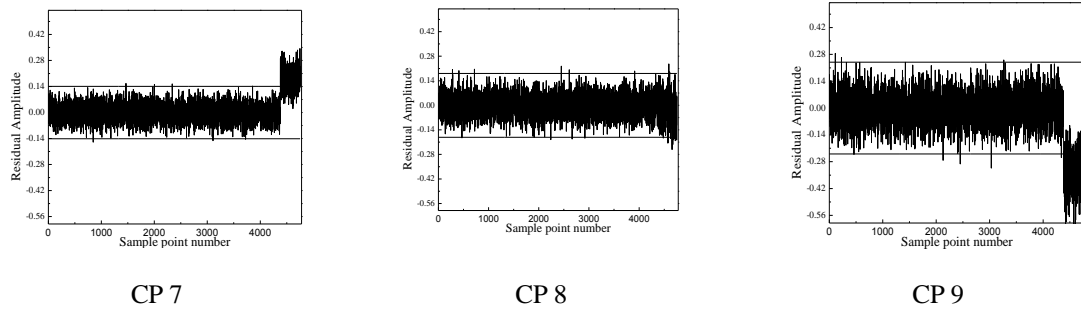


Fig. 8 Detection result of simulation 1 with 10% noise

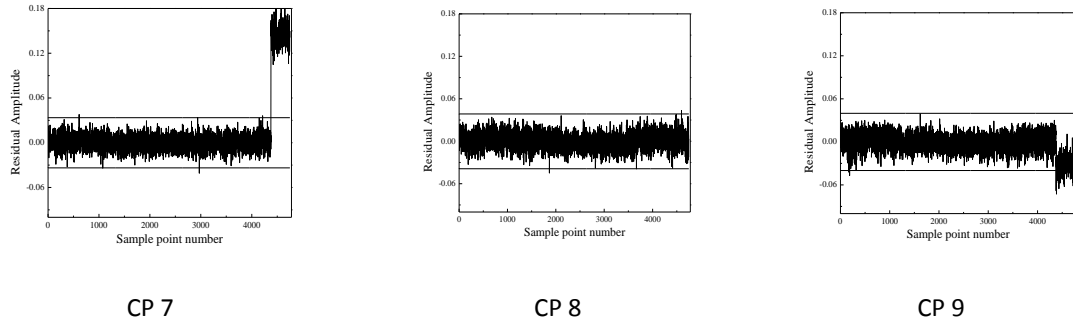


Fig. 9 Detection result of simulation 2 with 10% noise

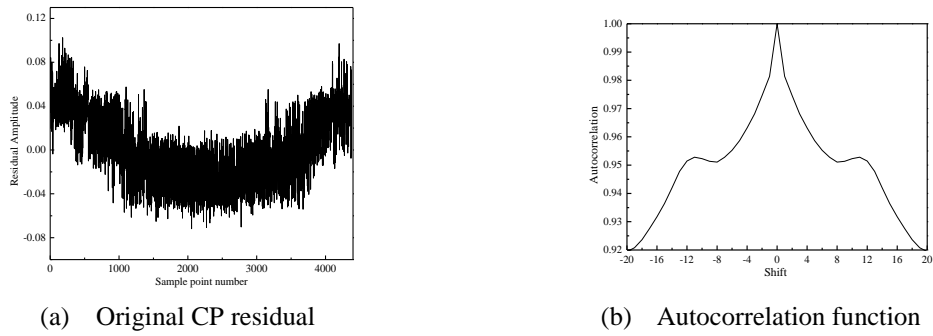


Fig. 10 CP residual and its autocorrelation function

#### 4.2 Simulation 2

The second simulation is carried out with T-LA applied in temperature gradient introduced before. This condition is quite similar to the real case. Damage case 1 is used in the computation of

displacements. Then the displacements are polluted by 10% noise. Only three pictures (Fig. 9) are given to describe the damage situation for simplicity, and similar results are obtained for this rather complicated case. The proposed approach still can distinguish when and where the damage occurs, although the results are not so well as the last simulation.

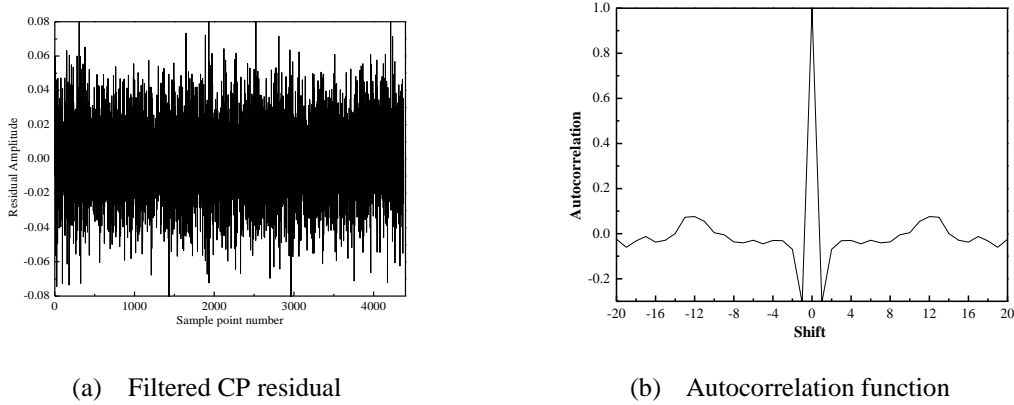


Fig. 11 Filtered residual sequences and its autocorrelation function

### 4.3 Simulation 3

The third one is applied with the T-BJ in uniform distribution through the section of the beam. The simulated displacement series are contaminated 10% noise, and damage case 2 is employed. In this situation, the CP residual exhibits strong autocorrelation, leading to the nonstationarity of the residual, as shown in Fig. 10. If the nonstationary variables (displacements) are auto-correlated and cross-correlated, the residual is auto-correlated (Chen *et al.* 2009). At the same time, the residual sequences are normally autocorrelated as a result of controller feedback and the presence of unmeasured disturbances which may be autocorrelated. In the former simulation, the displacements are also auto-correlated, but the degree is not strong enough to demonstrate the non-stationarity. However, when the noise level is high, it shows non-stationarity property in the residual sequences. To crack out this problem, the residual are filtered by an inverse AR filter (Monson 1996). In an AR model of order  $p$ , the current output is a linear combination of the past  $p$  outputs plus a white noise input. The weights on the  $p$  past outputs minimize the mean-square prediction error of the autoregression. Let  $e(t)$  be a stationary and serially uncorrelated sequence, the stationary autocorrelated residual sequence  $\varepsilon(t)$  can be determined as follows

$$e(t) = \varepsilon(t)A(z^{-1}) \quad (24)$$

where  $A(z^{-1}) = 1 - (a_1 z^{-1} + a_2 z^{-2} + \dots + a_p z^{-p})$ . Invert the AR filter to get

$$e(t) = \varepsilon(t) - \sum_{i=1}^p \varepsilon(t-i) \quad (25)$$

The filtered CP residual is shown in Fig. 11, showing that the autocorrelation is removed effectively. Now the filtered residual  $e(t)$  can be used for the online process monitoring.

The damage detection results are illustrated in Fig. 12. After postprocessing the CP residual, the result series looks perfect that the training set remains stationary while a clear jump could be observed when the damage occurs. Once again, the pair cointegration identifies the instant and the position of damage accurately. Through proper operating on the residuals, the pair cointegration can handle the nonlinearity caused by the elastic modulus dependent on the temperature. The inverse AR filter also improves the ability of resisting noise. Even when the noise level reaches 30%, ideal identification results could be obtained.

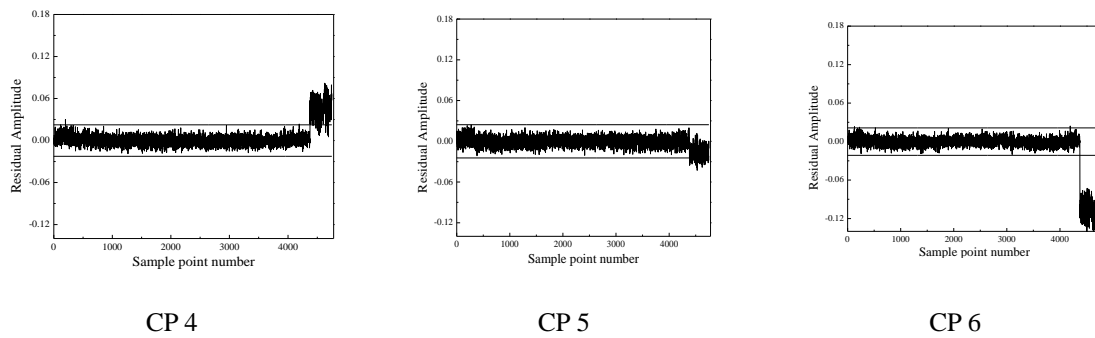


Fig. 12 Detection results of simulation 3 with 10% noise

#### 4.4 Simulation 4

The fourth case was used T-BJ with a temperature gradient across the section. Then, the displacement signals were added 10% noise. With the same damage location and the same data processing as Simulation 3, ideal results are obtained again, shown in Fig. 13, from which we can distinguish the damage temporally and spatially precisely. The proposed pair cointegration method does make a really good job even in this complicated model. It suggests that the pair cointegration method may be a prospective damage detection method to be applied in the real project.

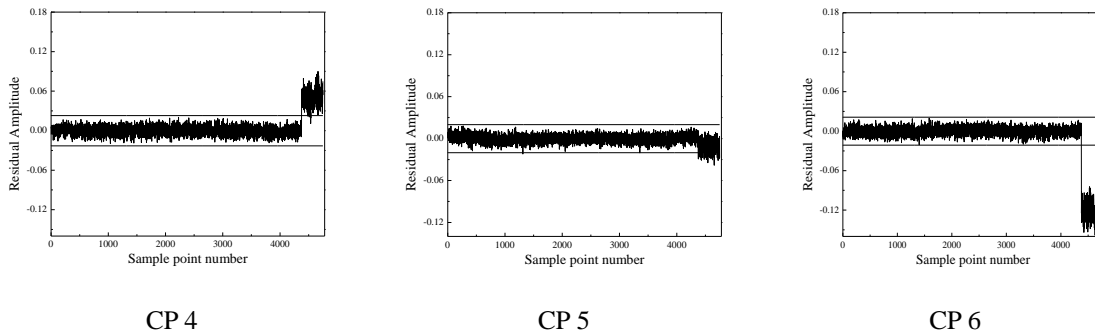


Fig. 13 Detection result of simulation 4 with 10% noise

## 5. Conclusions

The neighbored displacement series are combined to generate residuals by the cointegration approach, and it shows an excellent performance in detecting the damage spatio-temporally with environmental influence eliminated simultaneously. A prestressed concrete beam is adopted to test the feasibility of the proposed method. In all the four situations, the proposed detection method functions well and preserves the virtue of immunity to noise. Nonstationarity characteristics in the residual sequences could be amended by the autocorrelation filter that also can improve the anti-noise property. The pair cointegration method shows a well adaption for real application. It should be mentioned that the training sets should include all the possible environmental or operational conditions for the better performance of the pair cointegration method. Cointegration is a powerful mathematical tool to handle the time series sharing the similar trend, and as a result, it is not required that the variables cointegrated have the same unit.

## Acknowledgements

The authors are indebted to the editor and the referees, whose careful reading and comments helped improving an earlier version of the paper. This work was jointly supported by the Science Fund for Creative Research Groups of the NSFC (Grant No. 51121005), the National Natural Science Foundation of China (Grant No. 91315301, 51222806), the “973” Project (Grant No. 2011CB013605), and the Program for New Century Excellent Talents in University (Grant No. NCET-10-0287).

## References

- Askegaard, V. and Mossing, P. (1988), “Long term observation of RC-bridge using changes in natural frequencies”, *Nord. Concrete Federation*, **7**, 20-27.
- Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (1994), *Time series analysis-forecasting and control*, 3rd Ed, Prentice-Hall: Englewood Cliffs, NJ.
- Chen, Q., Kruger, U. and Leung, A. (2009), “Cointegration testing method for monitoring nonstationary processes”, *Ind. Eng. Chem. Res.*, **48**(7), 3533-3543.
- Cross, E.J., Worden, K. and Chen, Q. (2011), “Cointegration: a novel approach for the removal of environmental trends in structural health monitoring data”, *Proc. R. Soc. A.*, **467**, 2712-2732.
- Code for Design of Concrete Structures (GB 500102002), Beijing, 2002.
- Dewangan, U.K. (2011), “Structural damage existence prediction with few measurements”, *Int. J. Eng. Sci. Technol.*, **3**(10), 7587-7597.
- Dickey, D.A. and Fuller, W.A. (1979), “Distributions of the estimators for auto-regressive time series with a unit root”, *J. Am. Stat. Assoc.*, **74**, 427-431.
- Dickey, D.A. and Fuller, W. (1981), “Likelihood ratio statistics for autoregressive time series with a unit root”, *Econometrica*, **49**, 1057-1072.
- Doebling, S.W. and Farrar, C.R. (1997), “Using statistical analysis to enhance modal-based damage identification”, *Proceedings of the DAMAS 97: structural damage assessment using advanced signal processing procedures*, University of Sheffield, UK.
- Doebling, S.W., Farrar, C.R. and Prime, M.B. (1998), “A summary review of vibration-based identification



- methods", *Shock Vib.*, **30**(2), 91-105.
- Engle, R.F. and Granger, C.W.J. (1987), "Cointegration and error-correction: representation, estimation and testing", *Econometrica*, **55**, 251-276.
- Fan, W. and Qiao, P.Z. (2011), "Vibration-based damage identification methods: a review and comparative study", *Struct Health Monit.*, **10**(1), 83-111.
- Figueiredo, E., Park, G., Farrar, C.R., Worden, K. and Figueiras, J. (2010), "Machine learning algorithms for damage detection under operational and environmental variability", *Struct Health Monit.*, **10**(6), 559-572.
- Fuller, W. (1996), *Introduction to statistical time series*, Wiley-Interscience, New York.
- Fritzen, C.R., Mengelkamp, G. and Guemes, A. (2003), "Elimination of temperature effects on damage detection within a smart structure concept", *Proceedings of the 4th Int. Workshop on Structural Health Monitoring*, Stanford University, CA.
- Gao, T. M. (2009), *Econometrical analysis methods and modelling*, Qinghua Press, Beijing, China.
- Ghrib, F., Li, L. and Wilbur, P. (2012), "Damage identification of Euler-Bernoulli beams using static responses", *J. Eng. Mech. - ASCE*, **135**(5), 405-415.
- Johansen, S. (1988), "Statistical analysis of cointegrating vectors", *J. Econom. Dyn. Control*, **12**(2-3), 231-254.
- Johansen, S. and Juselius, K. (1990), "Maximum likelihood estimation and inference on cointegration with applications to the demand for money", *Oxford Bull. Econom. Stat.*, **52**(2), 169-210.
- Kim, C.Y., Jung, D.S., Kim, N.S., Kwon, S.D. and Feng, M.Q. (2003), "Effect of vehicle weight on natural frequencies of bridges measured from traffic-induced vibration", *Earthq. Eng. Eng. Vib.*, **2**(1), 109-115.
- Lee, E.T. and Eun, H.C. (2008), "Damage detection of damaged beam by constrained displacement curvature", *J. Mech. Sci. Technol.*, **22**(6), 1111-1120.
- Li, Y.H. (2009). *Study on structural modeling and damage identification methods of cracked beams*, M.S. thesis, Dalian Univ. of Technol., Dalian, China.
- Loh, C.H., Chen, C.H. and Hsu, T.Y. (2011), "Application of advanced statistical methods for extracting long-term trends in static monitoring data from an arch dam", *Struct Health Monit.*, **10**(6), 587-601.
- Monson, H. (1996), *Statistical digital signal processing and modeling*, John Wiley & Sons, Manhattan.
- Moorthy, S.S. and Roeder, C.W. (1992), "Temperature-dependent bridge movements", *J. Struct. Eng. - ASCE*, **118**(4), 1090-1105.
- Peeters, B. and De Roeck, G. (2000), "One year monitoring of the Z24-bridge: Environmental influences versus damage events", *Proceedings of the IMAC-XVIII, San Antonio, TX*.
- Sohn, H. (2007), "Effects of environmental and operational variability on structural health monitoring", *Philos. T. R. Soc. A.*, **365**, 539-560.
- Wang, Y.L., Liu, X.L. and Fang, C.Q. (2012), "Damage detection of bridges by using displacement data of two symmetrical points", *J. Perform. Constr. Fac.*, **25**(3), 300-311.
- Yan, A.M., Kerschen, G., De Boe, P. and Golinval, J.C. (2005), "Structural damage diagnosis under varying environmental conditions - part I: a linear analysis", *Mech. Syst. Signal Pr.*, **19**, 847-864.
- Zhou, H.F, Ni, Y.Q. and Ko, J.M. (2011), "Elimination temperature effect in vibration-based structural damage detection", *J. Eng. Mech. - ASCE*, **137**(12), 785-796.