

PAPER • OPEN ACCESS

A nonlinear cointegration approach with applications to structural health monitoring

To cite this article: H Shi *et al* 2016 *J. Phys.: Conf. Ser.* **744** 012025

View the [article online](#) for updates and enhancements.

Related content

- [Impact Damage Detection in Smart Composites Using Nonlinear Acoustics—Cointegration Analysis for Removal of Undesired Load Effect](#)
Phong B Dao, Andrzej Klepka, ukasz Pieczonka *et al.*
- [Cointegration approach for temperature effect compensation in Lamb-wave-based damage detection](#)
Phong B Dao and Wieslaw J Staszewski

Recent citations

- [Long-time monitoring of the G. Meazza stadium in a pattern recognition prospective](#)
Alessio Datteo *et al*

A nonlinear cointegration approach with applications to structural health monitoring

H Shi, K Worden and E J Cross

Dynamic Research Group, Department of Mechanical Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, UK

haichen.shi@sheffield.ac.uk

Abstract. One major obstacle to the implementation of structural health monitoring (SHM) is the effect of operational and environmental variabilities, which may corrupt the signal of structural degradation. Recently, an approach inspired from the community of econometrics, called *cointegration*, has been employed to eliminate the adverse influence from operational and environmental changes and still maintain sensitivity to structural damage. However, the linear nature of cointegration may limit its application when confronting nonlinear relations between system responses. This paper proposes a nonlinear cointegration method based on Gaussian process regression (GPR); the method is constructed under the Engle-Granger framework, and tests for unit root processes are conducted both before and after the GPR is applied. The proposed approach is examined with real engineering data from the monitoring of the Z24 Bridge.

1. Introduction

Structural Health Monitoring (SHM) is becoming increasingly important for industries like aerospace engineering, civil infrastructure, mechanical engineering, etc. The aim of SHM is to carry out structural health assessment, which often involves obtaining dynamic response measurements of a system, damage-sensitive feature extraction and statistical analysis of these features [1]. More specifically, for example, in order to perform damage detection of a bridge, which is the application of interest here, procedures including sensor instrumentation, data acquisition, signal processing, feature extraction and damage detection are implemented. Dynamic characteristics like natural frequencies, mode shapes and damping ratios are considered to be appropriate features because they are considered to be functions of a system's physical properties including mass, stiffness, damping and boundary conditions; damage-induced system changes can be defined as changes of these physical properties [2].

Vibration-based SHM of bridges is often significantly affected by operational and environmental conditions, including: temperature, wind, humidity, ambient loading conditions, mass loading and so forth. The measured signals are sensitive to system variation as well as the changes of operational and environmental conditions, which may lead to false-positive or false-negative alarms of damage where no corrective procedure has been attempted. The method of separating the effects of operation and environment from real structural change is called data *normalisation*. In the past, many papers have been published relating to the issue of data normalisation; a fairly recent and comprehensive literature



review can be found in [3]. In [3], different types of environmental and operational variability, e.g. temperature, boundary condition, mass loading and wind, were summarised with examples. The author also provided a good summary of several data normalisation methods including regression analysis, novelty detection and five other approaches that have been employed by the SHM community.

More recently, Cross et. al. [4], inspired by the econometrics literature, applied the concept of *cointegration* to the issue of removing common trends in SHM data. Cointegration is a property of some sets of nonstationary time series where a linear combination of the nonstationary series can produce a stationary residual, which can then be a potentially effective damage feature; the stationarity allowing use of the residual in a control chart [2]. The parameters of the linear combination are often referred to as a *cointegrating vector*. In [4], the cointegrating vectors were determined by a method from econometrics, specifically the *Johansen procedure*. The Johansen procedure is a maximum likelihood method, implemented as an eigenvalue problem, which seeks a white Gaussian residual from the cointegrating relation. An alternative procedure for estimating cointegrating relations is via the Engle-Granger (EG) two-step procedure [5]. The two steps in the EG procedure are as follows:

1. Given a set of measured time variables $\{y_i(t), i = 1, \dots, F\}$, one selects one of the variables of interest, without loss of generality, say $y_1(t)$, and estimates using ordinary least-squares methods, a regression on the other variables i.e.,

$$y_1(t) = a_2 y_2(t) + \dots + a_F y_F(t) + \varepsilon(t) \quad (1)$$

2. The regression parameters are then used to form an Error Correction Model (ECM), which is used to represent the long run relationships between variables.

It is not necessary here, to dwell on the nature of the ECM, because a simplified implementation of the EG two-step procedure is sufficient for SHM purposes. Assuming that the regression relationship in equation (1) has captured all of the long term relationships between the variables, the cointegrating residual $\varepsilon(t)$ will be a stationary sequence and can be used directly in the formation of control charts for monitoring [2]. (Readers curious about ECMs can consult [5] or any standard monograph on cointegration.)

However, the theory of cointegration is largely a linear theory, as manifested by the linear relationship in (1), and will not be applicable if the true relationships between the measured time series are nonlinear. While there has been a great deal of work on the problem, it is fair to say that there is currently no general theory of nonlinear cointegration. The book [6] covers much of the historical work carried out within the community of econometrics. In the context of SHM, the need for a nonlinear theory made itself clear quite quickly after the introduction of the method. Many engineering systems have nonlinear relationships between the measured features of interest, the Z24 Bridge data discussed later in this paper is a classic example. The first two attempts to formulate a nonlinear cointegration method for SHM purposes were [7] and [8]; in these references, variants of the Johansen and EG procedures were used in an evolutionary optimisation framework in order to estimate parameters for multinomial cointegrating relationships. The problem which occurred in these papers, was that the cointegrating residuals from the methods were *heteroscedastic* i.e. nonstationary in their variance. A recent study [9], made an interesting attempt to remove the problem; however, it appeared to do so by moving the nonstationarity from the variance of the residual sequence into its tail distribution. The problem is then that normal control chart thresholds would not be appropriate.

A potentially straightforward approach to nonlinear cointegration, as discussed in [6] (and partially adopted in [7,8]), is to simply extend step (1) of the EG procedure above, to a nonlinear regression, e.g.

$$y_1(t) = g(y_2(t), \dots, y_F(t)) + \varepsilon(t) \quad (2)$$

with g some regression function estimated in a parametric or nonparametric manner. Neural networks were suggested for the representation in [6], but any appropriate machine learning algorithm could be used in principle and this idea is where a proportion of the recent nonlinear cointegration literature has concentrated its attention. One recent paper uses least-squares support vector machines to good effect [10], and incorporated a Bayesian approach. Motivated by a desire to increase the Bayesian element in the nonlinear regression, the current paper will adopt *Gaussian Process Regression* (GPR) in order to learn the required relationship in equation (2) from data. This brings the immediate advantage of providing natural confidence intervals for the regression model. To illustrate the use of the GPR cointegration model, the paper will re-examine data from the well-known Z24 Bridge SHM benchmark exercise. Before proceeding to a discussion of the GPR approach and then to the case study, the next section will provide a little background on nonstationarity and cointegration.

2. Cointegration

As discussed above, in order to address the issue of environmental and operational variations (EOV) in SHM, cointegration, a method inspired from the discipline of econometrics, was introduced and initially employed to deal with EOVs in data collected from bridges and from ultrasonic NDE experiments. In the context of econometrics, cointegration was originally developed in order to model the underlying equilibrium relationships between nonstationary time series, which if they existed, implied that some linear combination of them would produce a stationary time series. Econometricians are normally more interested in testing for the presence of cointegration and estimating the cointegration relationships in their own right, but from the perspective of engineering applications, especially SHM, the motivations are slightly different. By building cointegration relationships of multivariate time series that may be affected by EOVs, one can project out the EOVs by acquiring a stationary series, which may serve as an effective damage indicator. Successful implementations in SHM were demonstrated by Cross et.al. in [5,7,8], in which cointegrating regression model errors were utilised as a damage indicator; EOVs were effectively eliminated in the stationary residual series, which, however, still maintained sensitivity to damage. Unfortunately, nature of the cointegration limited the applicability of the method to problems where variates were related linearly. In this paper a new approach based on nonlinear cointegration theory is proposed to address this issue. Although this paper proposes a nonlinear algorithm, it makes use of ideas from the linear theory and this section will discuss those ideas briefly before moving to the GPR.

2.1. Linear Cointegration

2.1.1 Stationarity and cointegration. Cointegration theory first arose in the field of econometrics to address the issue of “spurious regression”, which is due to the nonstationarity of some economic series. If two time series show monotonic trends, even if the trends are not causally related, ordinary least-squares regression will potentially find a spurious relationship. For example, two completely unrelated time series may be found to exhibit some regression. One of the original aims of cointegration was to help expose truly causal relationships.

The basic definition of cointegration given by Engle and Granger is that two or more nonstationary time series (i.e. time series whose statistical moments vary with time) are *cointegrated* if some linear combination of them is stationary. To give a measure of the nonstationarity of time series, the *order of integration* is defined as follows. If a nonstationary time series $x(t)$ can be transformed to be stationary after differencing d times, then $x(t)$ is said to be integrated of order d , which is denoted as $x(t) \sim I(d)$. So, an $I(1)$ series is one which becomes stationary series after differencing only once, an $I(0)$ series implies a stationary series. A more precise definition is now possible; a set of time series is cointegrated if they are all integrated of the same order and a linear combination exists which is integrated of a lower order [11]. Having established this definition, it is clearly important to have a means of estimating the order of integration; this can be accomplished using the ADF test.

2.1.2 The Augmented Dickey-Fuller (ADF) test. In order to determine the order of integration of a time series, a number of well-established statistical methods can be employed; one of the most commonly applied approaches is the Augmented Dickey-Fuller (ADF) test. Consider a time series $x(t)$, to test whether $x(t)$ is stationary, if not, to what order $x(t)$ is integrated to, $x(t)$ will be fitted to the ADF regression model, which has the form [11]:

$$\Delta x(t) = \rho x(t-1) + \sum_{j=1}^p \beta \Delta x(t-j) + c_1 t + c_2 + \varepsilon_t \quad (3)$$

where Δ is the difference operator, i.e. $\Delta x(t) = x(t) - x(t-1)$, p is the number of lagged terms, ε_t is the model residual, $c_1 t$ and c_2 represent a deterministic trend term and a shift respectively, these two terms are both optional choices which are dependent on the complexity of the real model. The procedures of the ADF test are: first, use a least square method to estimate the parameters in model (2), then to test the null hypothesis of $\rho = 0$. The t statistic on ρ is calculated [11]:

$$t_\rho = \frac{\hat{\rho}}{\sigma_\rho} \quad (4)$$

where $\hat{\rho}$ is the estimate of parameter ρ and σ_ρ is the variance of the estimated parameters. Because the limiting distribution of the estimator $\hat{\rho}$ is not a typical distribution (normal, student, etc), Dickey and Fuller (DF) investigated the asymptotic distribution of the t statistic, and constructed a table of critical values. Once the t statistic is calculated, it is compared with the DF table critical values, if $t_\rho < t_\alpha$, where t_α is the critical value at level α given by DF, one can accept the null hypothesis. Thus the time series x_t has a unit root and is an $I(1)$ series, otherwise $t_\rho > t_\alpha$, and the null will be rejected. Next $x(t)$ will be used to test the same null hypothesis, if accepted then $x(t)$ can be regarded as an $I(2)$ series, if rejected the same procedures will be repeated until the order of integration is determined. In order to properly implement cointegration, the order of the series needs to be examined both before and after the analysis has been carried out, in order to check that cointegration has successfully lowered the order of integration.

Regardless of whether one is attempting linear or nonlinear cointegration, the ADF test (or an appropriate alternative) is an important part of the analysis because it establishes the stationarity (or not) of the regression model residual if one is employing the EG procedure. As mentioned above, the regression model used here will make use of Gaussian processes.

3. Gaussian Process Regression

Gaussian process regression (GPR) is a Bayesian machine learning approach that can deal with nonlinear regression problems in such a way that confidence intervals for predictions are produced in a natural way. Generally speaking, a GP can be considered to be a distribution over functions, any finite samples from which are jointly Gaussian distributed [12]. Just as a Gaussian distribution can be fully specified by its mean and variance, analogously, a GP can be fully defined by its mean function $m(x)$ and covariance function $k(x, x')$. This means that for any given inputs, the corresponding outputs are normally distributed, which gives possibilities for predicting unknown observations. As with other Bayesian inference approaches, unknown parameters are treated as random variables, and predictions are made probabilistically.

In applying GPR, first of all, an appropriate prior assumption for the mean and covariance functions is applied. Then the joint distribution of known and unknown points is conditioned on training data, to give a predictive distribution for any new input observations. The final step is to integrate out parameters/hyperparameters in the model to give a posterior distribution. The detailed procedures are outlined as follows.

For a given input \mathbf{x} , a GP can be defined as [12],

$$f(\mathbf{x}) \sim GP(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')) \quad (5)$$

$$m(\mathbf{x}) = E[f(\mathbf{x})] \quad (6)$$

$$k(\mathbf{x}, \mathbf{x}') = \text{cov}(f(\mathbf{x}), f(\mathbf{x}')) = E[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))] \quad (7)$$

where \mathbf{x}' are neighbouring points of the input \mathbf{x} , E is the expectation operator. As normally little is known about the model, the first “guess” about the data, or the prior mean function is set to zero here; other possible choices of mean functions are linear functions, polynomial functions, etc. Therefore, the selection of covariance function is the key issue for GP. The covariance function defines the covariance of neighbouring data points as a function of the corresponding distance; this governs the smoothness of a process. The function used in this paper is called the squared exponential function, which is one of the most simple and commonly used covariance functions, it has the form [12],

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{1}{2l^2} |\mathbf{x} - \mathbf{x}'|^2\right) + \sigma_n^2 \delta \quad (8)$$

where σ_f^2 is the signal variance that dominates the magnitude of the process, σ_n^2 is the noise variance, l is the length scale parameter which controls the minimum length of two uncorrelated input variables. These parameters that control the behaviour of covariance functions are termed the *hyperparameters*. In the implementation of GPR used here, optimal values of the hyperparameters are estimated using a maximum likelihood approach [12].

In order to make predictions at a new input point \mathbf{x}^* , one conditions the prior distribution on a known training set. As mentioned above, under the framework of the GP, the prediction and training set can be treated as joint Gaussian distribution, thus the predictive distribution of $f(\mathbf{x}^*)$ can be estimated by conditioning the prior distribution on training data. The details of the analysis can be found in [12], only the summary formulae for the results will be given here.

Consider a training set based on vector inputs and scalar outputs, $X = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$ is the input matrix, $\mathbf{y} = [y_1, y_2, \dots, y_N]$ is the corresponding output vector. For new (test) observations one defines the input matrix $X_* = [\mathbf{x}_1^*, \mathbf{x}_2^*, \dots, \mathbf{x}_r^*]$ and the corresponding output vector $\mathbf{y}_* = [y_1^*, y_2^*, \dots, y_r^*]$. \mathbf{y} and \mathbf{y}_* are joint normally distributed, and according to the GP assumption, one can show that the predictions are distributed as [12],

$$\mathbf{y}_* | \mathbf{y} \sim N\left(k(X_*, X)k(X, X)^{-1}\mathbf{y}, k(X_*, X)k(X, X)^{-1}k(X, X_*)\right) \quad (9)$$

For which the predictive mean $m_{pred}(X_*)$ and the predictive variance σ_{pred} are expressed as follows:

$$m_{pred}(X_*) = k(X_*, X)k(X, X)^{-1}\mathbf{y} \quad (10)$$

$$\sigma_{pred}(X_*) = k(X_*, X_*) - k(X_*, X)k(X, X)^{-1}k(X, X_*) \quad (11)$$

where $k(X, X_*)$ is a convenient shorthand for the covariance matrix between input matrices X and X_* [12]. The predictive mean is the expected value of the output for a given test input and the predictive variance provides a natural confidence interval.

GPR can be used to estimate the nonlinear cointegrating relationship proposed in equation (2). A sequence of steps should be followed.

1. Select suitable monitored variables for the model and calculate the order of the variables, which should be integrated to the same order. The ADF test is implemented on the variables to determine the integration order.
2. Separate the data sets into two parts: a training and test data set. Training data are used to train the GP regression model, the test data set is employed for the aim of monitoring potential system variation. The training data should not contain any data corresponding to damage, but should as far as possible span the range of EOVs anticipated.
3. Use the training data to train the GP regression model, then apply the ADF again to testify whether the model residual series is integrated to a lower order than the original variables. Once this goal is achieved, one can say that the nonlinear cointegrating relationship is established successfully, the common trends are purged, therefore, the model residual series may be a potentially good indicator of damage induced variations.

4. Case study: SHM of the Z24 Bridge

3.1 Properties of natural frequencies

The Z24 Bridge in Switzerland is a well-known bridge in the SHM community, because the monitoring campaign introduced artificial damage during the one year monitoring study, more specifically from 11 November 1997 to 11 September 1998. The monitoring campaign also measured environmental parameters including air temperature, wind characteristics, humidity and so on. In order to obtain the dynamic behaviour of the bridge, the modal parameters such as natural frequencies were extracted from acceleration measurements [18]. Figure 1 shows the first four natural frequencies with respect to observation numbers, the black dotted line shows the day when the first of the damage scenarios was implemented. However, partly due to the influence of environmental conditions, the natural frequencies show no significant sign of structural degradation. Furthermore, after nine months of monitoring of the normally functioning bridge, six kinds of damage scenarios were artificially employed on different parts of the bridge. All damage scenarios, with the corresponding date and data points are listed in Table 1. For more details about the monitoring systems and the progressive damage tests, the interested reader can refer to [13].

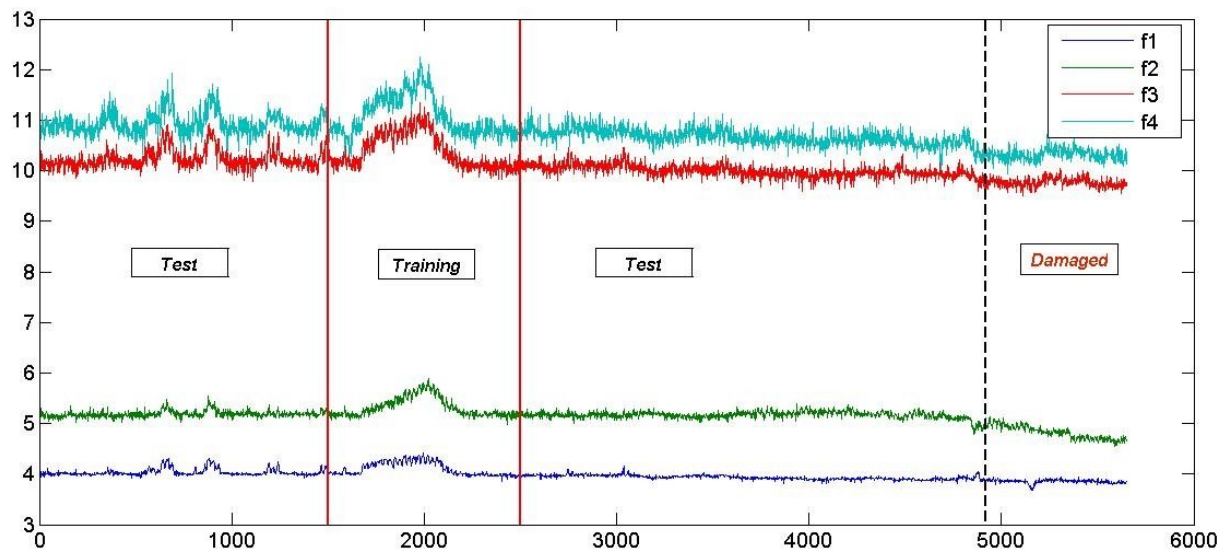


Figure 1. The first four natural frequencies identified from the Z24 Bridge, the black dotted line indicates damage introduction (at data point 4918), the shaded area separates the data set into two parts: training and test.

Date	Damage Scenario	Day Index	Data Points
10/08/98	Settlement of Pier, 20 mm	272	4918~4941
12/08/98	Settlement of Pier, 40 mm	274	4966~4989
17/08/98	Settlement of Pier, 80 mm	279	5084~5107
18/08/98	Settlement of Pier, 95 mm	280	5108~5131
19/08/98	Tilt of Foundation	281	5132~5155
20/08/98	3rd Reference Point	281	5156~5179

Table 1. Progressive damage test scenarios and test dates [19].

On further observation of Figure 1, it is obvious that the natural frequencies were affected by some external driver to a great extent, and this turned out to be temperature. See data points from 1800 to 2200 for example, the corresponding test dates were 28/01/1998 to 15/02/1998, air temperature was mostly below zero during this period. There emerged a large peak of the natural frequencies approximately at data number 1800-2200, this may be explained by the nonlinear behaviour of the natural frequencies caused by freezing of the asphalt [17]. It seems that the effect of temperature may result in nonlinear relations between natural frequencies. Figure 2 plots the relationships between each of the extracted frequencies. With visual inspection, it can be seen that the second frequency is not linearly related with the other frequencies, while f_1 , f_3 and f_4 appear to be mutually linearly related [17]. The results of the mutual relationships between frequencies are summarised in Table 2. Following the procedures in Section 2.3, the orders of integration for the time series of the four natural frequencies are determined using the ADF test; the results are given in Table 3, and show that these four natural frequency series are nonstationary at the 95% confidence level and integrated of order one, i.e. they are $I(1)$ series.

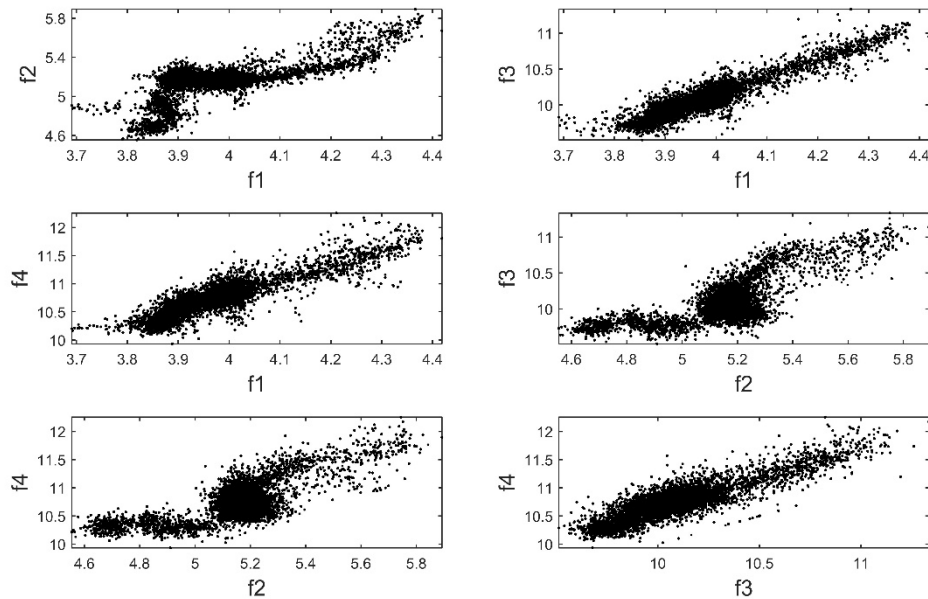


Figure 2. Mutual relationships of the first four natural frequencies.

	f_1	f_2	f_3	f_4
f_1		Nonlinear	Linear	Linear
f_2	Nonlinear		Nonlinear	Nonlinear
f_3	Linear	Nonlinear		Linear
f_4	Linear	Nonlinear	Linear	

Table 2. Relationships between the first four natural frequencies [17].

	ADF statistic(5% level)	5% critical value	Stationarity
f_1	-0.317481	-1.940880	N
Δf_1	-14.37213	-1.940880	Y
f_2	-0.741687	-1.940880	N
Δf_2	-17.09163	-1.940880	Y
f_3	-0.418102	-1.940880	N
Δf_3	-28.57986	-1.940880	Y
f_4	-0.423898	-1.940880	N
Δf_4	-26.46857	-1.940880	Y

Table 3. ADF test of the first four modal frequency time series.

3.2 Regression model of natural frequencies

It might be difficult to find the specific physical meaning of a natural frequency regression model; however, from the view of cointegration, once the cointegration relationship of variables is built, the equilibrium relationship may indicate that system is functioning under normal conditions, hence the model error may be sensitive to system variation induced by damage, if the model error tends to be stationary, the underlying common trends of the original data are successfully removed. Considering the nonlinear behaviour of natural frequencies during the conditions of low temperature, data points from data point 1500 ~ 2500 are selected as training data, for the reason that this part of the data contains both large peaks and small fluctuations caused by environmental conditions. The rest of the data set is used for testing, as is illustrated in Figure 1. Because it is necessary to select one of the four

natural frequencies to serve as regression target, there are four possible regression models, namely, $f_1 = g(f_2, f_3, f_4)$, $f_2 = g(f_1, f_3, f_4)$, $f_3 = g(f_1, f_2, f_4)$ and $f_4 = g(f_1, f_2, f_3)$. Following the proposed procedures, the ADF tests are applied to the training residual, the results are listed in Table 4, since all ADF test statistics are smaller than the critical value at 5% level, the residuals are stationary, and in other words, the nonlinear cointegrating relationships are successfully established. The prediction behaviours of the four possible models are quite similar, they all tend to have a good fit to the data. Due to the limited space here, only one of the four possible regression models is shown in Figure 3 which illustrates the prediction performance of GP regression model and also the model residual. In Figure 3(a), the regression model $f_2 = g(f_1, f_3, f_4)$ is used. Generally, the GP models follow the trend of the measurement data very well; the environmentally induced variations are also well predicted. Furthermore, by observing Figure 3(a), the residual becomes non-stationary after the damage is introduced. The upper and lower lines represent the training residual mean plus or minus three times the standard deviations. Combining the data of progressive damage tests of the monitoring campaign, Figure 3(b) correspond to the day index 270~271 (8~9 August, 1998). Recalling Table 1, it shows that around these dates, the damage scenario 'Settlement of Pier' was implemented. However in Figure 3, 'damage' is detected on day index 225, when no damage was introduced, similar detection results can be found in [19], the similar sudden frequency drop was interpreted in terms of a possible increase of mass of the bridge due to some heavy trucks standing on it.

Regression model	$f_1 = g(f_2, f_3, f_4)$	$f_2 = g(f_1, f_3, f_4)$
ADF test statistic	-10.451	-5.850
5% critical value	-1.941	-1.941
Stationarity	Y	Y
Regression model	$f_3 = g(f_1, f_2, f_4)$	$f_4 = g(f_1, f_2, f_3)$
ADF test statistic	-11.872	-9.156
5% critical value	-1.941	-1.941
Stationarity	Y	Y

Table 4. ADF test results of training residual.

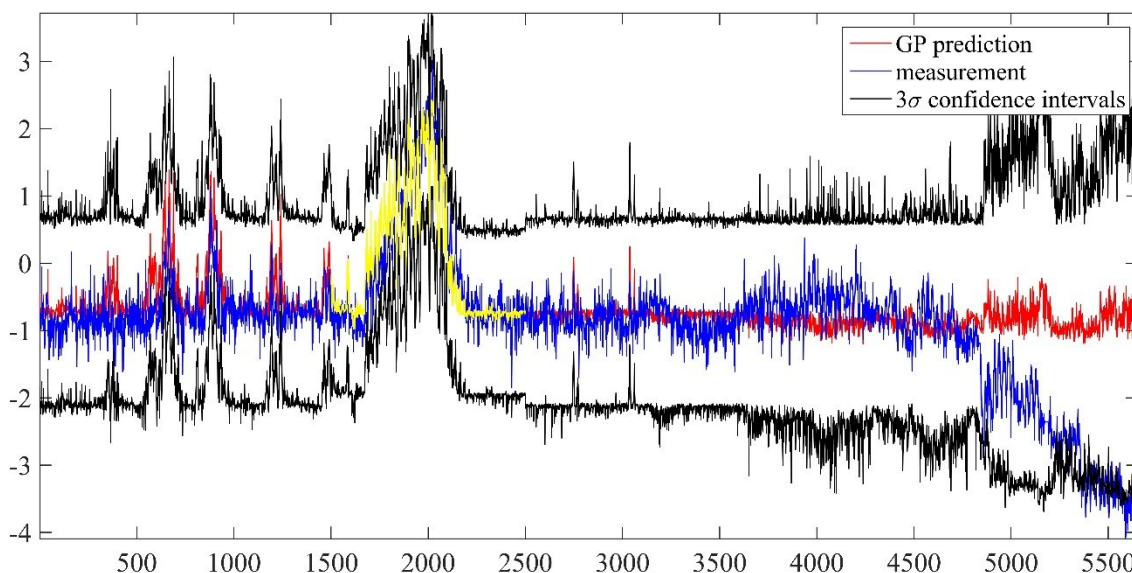


Figure 3(a). Gaussian process regression Model $f_2 = g(f_1, f_3, f_4)$ prediction performance comparing with real measurements.

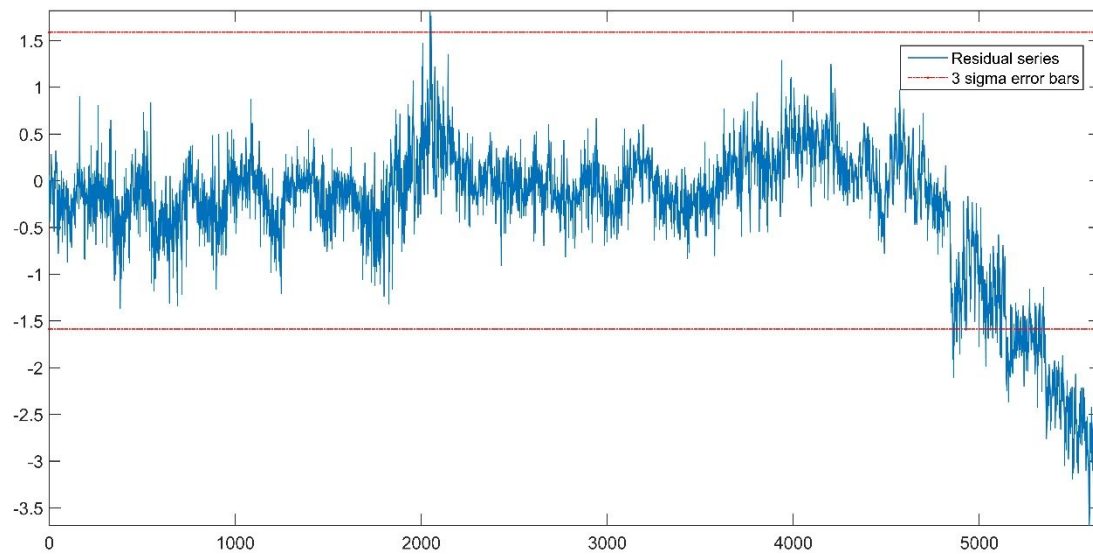


Figure 3(b). Gaussian process regression Model $f_2 = g(f_1, f_3, f_4)$ residual.

5. Discussion and conclusions

Cointegration, an emerging approach in SHM, has been successfully employed to deal with environmental and operational effects in SHM. More specifically, cointegration, used to model the long-run equilibrium relationships in economic series, can be utilised in SHM problems to find some linear combination of nonstationary variables sharing the same trend to construct a stationary residual purged of common trends, which may serve as a damage indicator. As an extension of previous work, a nonlinear cointegration approach based on the Engle-Granger procedure and Gaussian process regression is proposed in this paper, aiming to build nonlinear cointegrating relationships between variables. The idea is basically within the framework of linear cointegration, except for using GP regression as the cointegrating function. The ADF test is implemented both before and after the cointegrating regression to determine the order of integration, and furthermore to examine whether a nonlinear cointegration relationship has been built. Real data from the Z24 Bridge is used to validate this method, the results show that the model error series created from the model basically stays stationary during normal operating conditions, which means the trends induced by EOVs are effectively eliminated, yet the model series residuals still maintain sensitivity to structural damage.

One obvious advantage of implementing cointegration, comparing with other approaches like linear regression, is that it is clear that no direct measure of environmental or operational variations are necessary. This may provide great convenience for SHM, because the scale of structures of interest are growing massively, it is sometimes difficult to acquire measures of ambient variations accurately. However, one major limitation for this method would be that the GP assumes the model to be a smooth function, even though it works well under many circumstances, when discontinuities or an underlying regime switch occurs in structures, the GP may fail to model accurately. This perhaps can be explained by the different structural response under cold and warm temperatures. This issue may potentially be addressed by partitioning the data set into different regimes, and fitting multiple GPs into the model, this will be considered in the future work of the authors.

References

- [1] C R Farrar and K Worden 2013 *Structural Health Monitoring Structural: a Machine Learning Perspective* (John Wiley & Sons)
- [2] J Kullaa 2003 *Damage detection of the Z24 Bridge using control charts*. Mech. Sys. Sig. Proc. **17** pp.163–170
- [3] H Sohn 2007 *Effects of environmental and operational variability on structural health monitoring*. Phil. Trans. R. Soc. Ser. A **365** pp.539–560.

- [4] E J Cross, K Worden and Q Chen 2011 *Cointegration: a novel approach for the removal of environmental trends in structural health monitoring data*. Proc. R. Soc. Ser. A **467** pp.2712–2732.
- [5] A Banerjee, J Dolado, J W Galbraith and D.F. Hendry 1993 *Co-Integration, Error-Correction, and the Economic Analysis of Non-Stationary Data*. Oxford University Press.
- [6] G. Dufrénot and V Mignon 2002 *Recent Developments in Nonlinear Cointegration with Applications to Macroeconomics and Finance*. Springer.
- [7] E J Cross and K Worden 2011 *Nonlinear cointegration as a combinatorial optimisation problem*. Proceedings of International Workshop on Structural Health Monitoring, Palo Alto, CA, pp.87-94
- [8] E J Cross and K Worden 2011 *Approaches to nonlinear cointegration with a view towards applications in SHM*. J. Phys. Conf. Ser. **305** p.012069.
- [9] K Zolna, P B Dao, W Staszewski and T Barszcz 2015 *Nonlinear cointegration approach for condition monitoring of wind turbines*. Math. Prob. Eng. Article ID 978156.
- [10] T Van Gestel, M Espinoza, B Baesens, J K Suykens, C Brasseur and B De Moor 2006 *A Bayesian nonlinear support vector machine error correction model*. J. Forecast. **25** pp.77–100.
- [11] R Perman 1991 *Cointegration: An introduction to the literature* J. Econ. Stud. **18** no 3.
- [12] C Rasmussen 2006 *Gaussian Processes for Machine Learning* (MIT Press)
- [13] B Peeters and G De Roeck 2001 *One-year monitoring of the Z 24-Bridge: environmental effects versus damage events* Earthq. Eng. Struct. Dyn.