Proceedings of the SEM IMAC XXX Conference Jan. 30 – Feb. 2, 2012, Jacksonville, FL USA ©2011 Society for Experimental Mechanics Inc.

Structural Damage Detection Using Multivariate Time Series Analysis

Ruigen Yao[1] and Shamim N. Pakzad[2]

[1] Graduate Student, Department of Civil and Environmental Engineering, Lehigh University, Bethlehem, PA 18015. Email:ruy209@Lehigh.edu

[2]Assistant Professor, Department of Civil and Environmental Engineering, Lehigh University, Bethlehem, PA 18015. Email: pakzad@Lehigh.edu

ABSTRACT: Much research has been focused in the past few decades on data-driven structural health monitoring based on sensor measurements. Modal parameters from system identification are the most widely studied structural state indicators adopted for this purpose; however, recent research has showed that they are not sensitive enough to local damage. In an effort to seek more effective alternatives, univariate autoregressive (AR) modeling on structural response has been investigated in several publications, where model characteristics are used as damage indices. Although these methods are generally successful, they tend to generate false alarms when the environmental conditions are varying because responses from only one location/sensor are considered. To strike a balance between sensitivity and stability, in this paper autoregressive with exogenous input modeling on measurements from several adjacent sensing channels is presented and applied to detect damage in a space truss structure. The damage feature is extracted from the residuals obtained via fitting the baseline model to data from the current structure. Also, damage localization is attempted by examining the estimated mutual information statistic between data from adjacent sensing channels. The damage identification/localization results thus obtained are then compared to those from univariate AR modeling to evaluate their relative pros and cons.

1. Introduction

Damage detection using structural vibration measurements is a field that has been actively investigated by the civil engineering research community in the past few decades[4, 5]. It is hoped that through these investigations automatic structural condition assessment can be realized one day, thereby reducing the cost on infrastructure maintenance as the involvement of human experts becomes less necessary. One of the earliest damage indicators studied is the estimated modal property from system identification [1,9], as it is directly related to structural physics according to classical dynamics theory. However, system identification algorithms, especially those that operate in time domain, are sometimes computationally intensive to implement. Also, the effectiveness of modal property as damage indicator varies depending on the structural layout and damage type. Studies have shown that it can be insensitive to local/minor damage[4] . Its performance could often be improved by feeding the modal property to a finite model updating scheme [10], but that is at the cost of increased algorithm complexity. To circumvent these problems, more efficient alternatives are needed.

Univariate time series analysis on vibration measurements have been successfully adopted in several recent research papers [6, 7, 11, 13] for damage identification purposes. This method bears certain similarities to traditional system identification as both are concerned with numerical modeling, yet the former is more flexible because it uses various damage features that do not necessarily have an explicit physical meaning. Damage features can be functions of either model parameters or model coefficients, and well-known statistical concepts are applied to set the critical damage threshold for the features extracted. Since the method only utilizes responses from a single sensor node at each time, it is relatively sensitive to local damage. But for the same reason, the results produced become less reliable when the environmental condition is not stationary. Moreover, the damage location in general cannot be inferred from the value of the damage indices, confining the application of this method to preliminary damage detection (detecting only the damage existence).

In this paper, multivariate autoregressive (AR) time series analysis [2] will be applied for damage detection. Basically, this method seeks to reach a balance between damage sensitivity and output stability by constructing the model on responses from several adjacent nodes. The damage feature used here is based on the autocorrelation function (ACF) of model residuals, and through examining the amount of change occurred in the feature value obtained for different node combinations damage localization will be attempted. In addition, an index that represents the mutual information [3] between two sensing channels is also used for damage detection here. While being easier to compute than the one based on multivariate AR modeling, this index nonetheless served the purpose for damage detection/localization in the application presented in section 4.

The organization of the paper is as follows; section 2 is a brief review of the univariate AR modeling method for damage identification, and section 3 presents the multivariate AR modeling method and the mutual information method. Section 4 contains the application of the established and new methods to acceleration measurements collected from a space truss structure under ambient loading, and the performances of different algorithms are compared and contrasted.

2. Univariate AR time series analysis for damage identification

Autoregressive models have long been successfully applied to model, validate and predict signals from various types of sources. The definition of a univariate AR model is shown below:

$$x(t)=\sum_{j=1}^{p}\phi_{xj}x(\mathrm{t}-\mathrm{j})+\varepsilon_x(t). \tag{1}$$

where ϕ_{xj} are the AR model coefficients, $\varepsilon_x(t)$ is the model residual, and $x(t)$ is the time series to be analyzed. It is easy to see that this model basically attempts to predict the future output using a linear combination of previous outputs.

Univariate AR model parameters can be very efficiently estimated from a signal using one of the standard algorithms. According to the specific feature extraction process, the damage features from AR model can be classified into two categories: model coefficients based and model residual based. In the remainder of this section, examples from both categories will be presented.

2.1 AR coefficients based damage indicators

It has been proved that if the signal is really an AR process, then any regular coefficients estimator $\{\phi_{xj}\}$ from the signal is asymptotically unbiased and normally distributed with covariance matrix $\sigma_e^2\Gamma_p^{-1}$ [2]. Therefore, Mahalanobis distance, a metric that represents the deviation in the probability space of normal distribution, seems to be a good choice of damage feature.

The estimator of the Mahalanobis distance between a potential outlier vector x_ξ and a baseline sample set can be obtained as

$$D_\xi=(x_\xi-\bar{x})^T\hat{\Sigma}^{-1}(x_\xi-\bar{x}). \tag{2}$$

where $\bar{x}$ is the average of the baseline sample feature vectors, and $\hat{\Sigma}$ the estimated covariance matrix. When applied for damage identification, an assessment of the current structural state is made by comparing the Mahalanobis distances obtained within the baseline/healthy state and those obtained from unknown/current state.

Cosh spectral distance[12] of AR model spectra, another model coefficients based feature, has a formulation similar to that of Mahalanobis distance feature. Corresponding spectrum plot (Fig. 1) can be constructed given an AR model:

$$S_{AR}^{(p)}(\omega)=\frac{\sigma_e^2}{|\phi(e^{j\omega})|^2}=\frac{\sigma_e^2}{|\sum_{k=0}^{p}\phi_k e^{-j\omega k}|^2}. \tag{3}$$

In application σ_e^2 is not calculated and set to 1, since its value is determined by excitation level. And Cosh spectral distance between the mean of baseline spectra $\bar{S}(\omega_j)$ and an unknown state spectrum $S(\omega_j)$ can be computed from Eq. 4:

$$(S,\bar{S})=\frac{1}{2N}\sum_{j=1}^{N}\left[\frac{S(\omega_j)}{\bar{S}(\omega_j)}-\log\frac{S(\omega_j)}{\bar{S}(\omega_j)}+\frac{\bar{S}(\omega_j)}{S(\omega_j)}-\log\frac{\bar{S}(\omega_j)}{S(\omega_j)}-2\right]. \tag{4}$$

When the system is damaged, the Cosh distance value would increase.

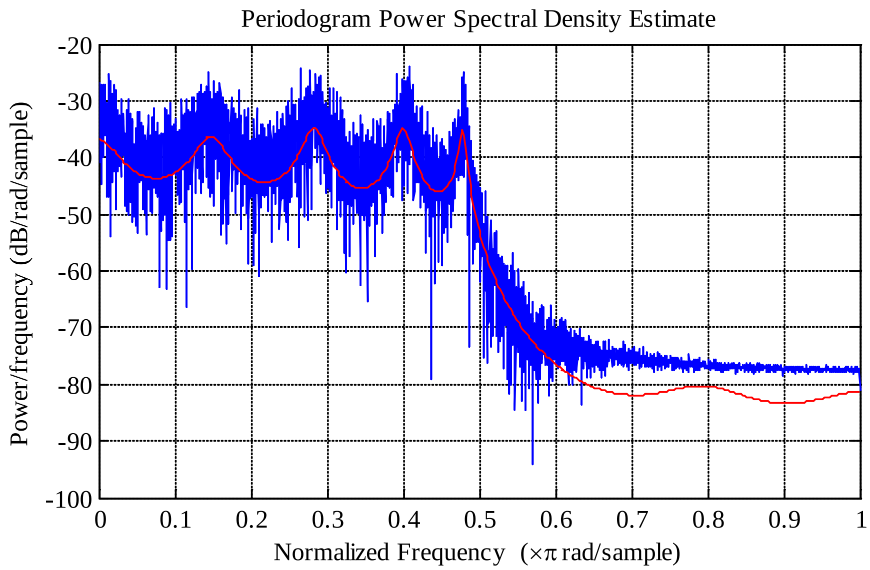

Fig.1 comparison between an AR signal power spectrum density and the estimated AR model spectrum (the smooth red line)

2.2 <u>AR residual based damage indicators</u>

When applying the AR coefficients based algorithms, the modeling process generally has to be repeated for a large number of times to get enough coefficient vectors for structural diagnosis. In the residual based algorithms, however, the model is constructed only once from the baseline signal. Residuals can then be generated by fitting the model to the signals collected.

If the system is damaged, the baseline model will no longer produce a good fit of the new data. In other words, the difference between predicted response from baseline model and real signal will increase. The model residual variance in most cases provides a valid measure of this difference, and has been used to detect damage in several literatures [6, 11]. More sensitive damage features could be obtained by taking into consideration the entire autocorrelation function (ACF) of the residuals, i.e. employing the Ljung-Box statistic [8]:

$$Q = n(n+2)\sum_{j=1}^{h}\frac{\rho_j^2}{n-j}. \tag{5}$$

where n is the sample size, h is the number of lags, and ρ_j is the autocorrelation at the j^{th} lag. This Q-statistic follows a χ^2 distribution under the normality assumption of the signal.

3. Damage identification/localization using multi-channel responses

Time series analysis on measurements from a single sensor node provides an efficient way for damage detection. However, applications have shown that their effectiveness as damage location indicator depends on the specific structural type. Also, due to the information limitation, this family of methods tends to suffer from false alarms when the environmental conditions are varying. Here in this section, damage features constructed from measurements from several adjacent sensors will be introduced in hope of improving algorithm stability and damage localization capability.

3.1 <u>ARX analysis using several adjacent nodal responses</u>

The single-input-single-output (SISO) ARX model[9] is defined as follows:

$$x(t) = \sum_{i=1}^{a} \alpha_i x(t-i) + \sum_{j=1}^{b} \beta_j u(t-j) + e_x(t). \tag{6}$$

where $x(t)$ and $u(t)$ are the output and input signal, respectively. Notice that this is also a linear regression model that associates the current response of a signal with its previous values and an exogenous input. If there are several exogenous inputs (multi-input-single-output case), each input will have its corresponding coefficient vector $\{\beta_j\}$.

Here when this model is applied for damage identification, the output signal will be from a certain sensor node (possibly at a potential damage location) and the input signals will be from its neighboring nodes. Here the damage indicators will be computed from residuals analysis, with the general procedure almost the same as that for univariate AR modeling. Model parameter based analysis will not be attempted here because it will require estimation of the multivariate model for many times, thereby resulting in high computational cost.

3.2 Mutual information between signals collected from adjacent nodes

Mutual information[3] is a statistic defined to measure the mutual dependence/similarity between random variables (Fig. 2). It is first proposed in communication theory to quantify the capacity of data transmission channels. Given the probability distributions of two random variables X and Y, their mutual information can be computed as;

$$I(X;Y) = \int_X \int_Y p(x,y) \log \frac{p(x,y)}{p(x)p(y)} dxdy. \tag{7}$$

$p(\cdot,\cdot)$ and $p(\cdot)$ here denote the joint and marginal probability distribution. It is clear that the value of this metric is always non-negative, and equal to zero only when X and Y are statistically independent. The definition can also be easily extended to the case where X and Y are random vectors.

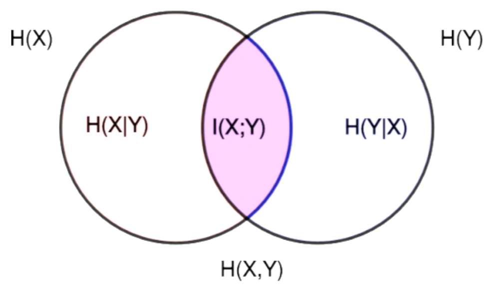

Fig.2 [3] An illustration of the definition of mutual information as the sum of the separate entropies of two random variables X and Y subtracted by their joint entropy H(X,Y). Entropy is essentially a measure of uncertainty for random variables.

If both X and Y are assumed to follow a Gaussian distribution, then the mutual information statistic can be obtained directly from their second-order statistical moments:

$$I(X;Y) = \frac{1}{2} \log \frac{|K\{X\}||K\{Y\}|}{|K\{X,Y\}|}. \tag{8}$$

where $K\{X\}$ stands for the variable covariance matrix and $|\cdot|$ is the matrix determinant. Mutual information can be more efficiently estimated from this formula than from the previous one, as it requires only the variance of the variables, instead of the complete distribution.

In the next section, mutual information between responses from two adjacent sensors will be employed as damage index. If damage (stiffness reduction) occurs between these two nodes, then it is expected that their mutual information value will decrease significantly. Since structural responses at any measured location are always correlated over time, they cannot be treated as single random variable. Rather, a time window of length L will be applied to each response signal and the estimated mutual information $\hat{I}(x_1(t);x_2(t))$ between two signals $x_1(t)$ and $x_2(t)$ becomes:

$$\hat{I}(x_1(t);x_2(t)) = I(x_1(t), x_1(t+1),..., x_1(t+L-1); x_2(t), x_2(t+1),..., x_2(t+L-1)) \tag{9}$$

Here the parameter L needs to be chosen with care so that the autocorrelation of signals after Lth lag will be close to zero.

4. Experimental validation of the statistical algorithms

The statistical algorithms described in the above two sections are applied to acceleration measurements collected from a space truss in the lab. The truss has its four supports fixed to sturdy I-beam sections, and 10 wireless sensors are mounted on the truss nodes in the midspan (Fig.3). Artificial damage is introduced by removing one interior diagonal member between sensor node 1 and 8. For each structural state(undamaged/damaged), two sets of data are collected at a sampling frequency of

280 Hz. As an effort to reduce the noise content, a low-pass filter is applied to all the data, which are subsequently downsampled to 70 Hz. The compressed data is then used as input to the damage detection algorithms, and dataset 1 is the selected baseline signal for all the implementations.

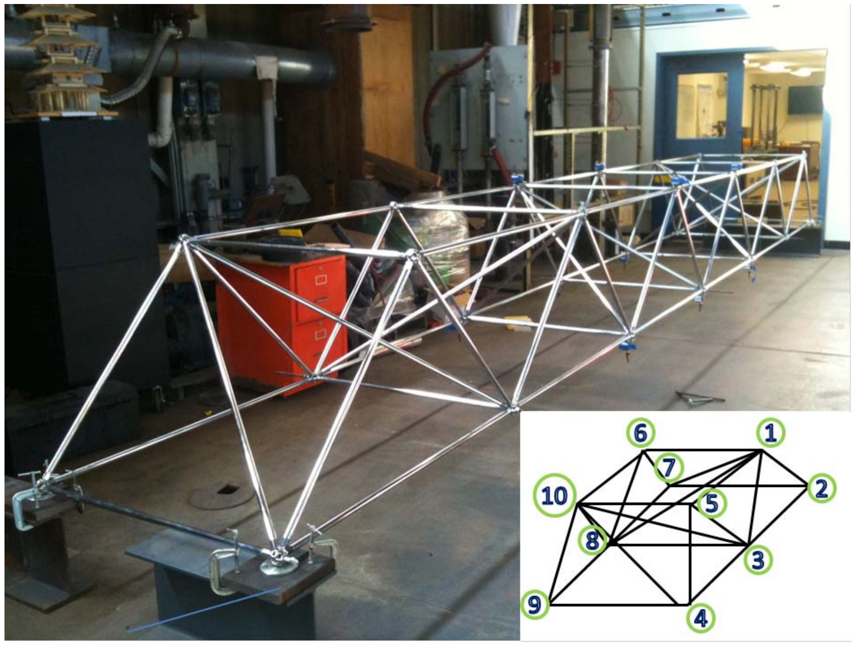

Fig. 3 The space truss model with the sensor node numbers shown in the lower right corner

4.1 Damage identification results from univariate AR coefficients based method

Fig. 4 shows the model coefficient based Mahalanobis distance features and model spectrum based Cosh spectral distance features extracted from different datasets. The first two datasets are from the healthy state of the structure, and the last two datasets are from the damaged state. In application of both methods, the signals are divided into many short segments with large overlap among them so as to produce enough AR coefficient vectors from them for structural state evaluation. It can be seen from the plots that after the damage, a lot of outliers will appear. Also, the Mahalanobis distance method is not as stable as the Cosh spectral distance, i.e. more false alarms are observed in plots of the former. This is because spectral distance generally assigns more weight to the position of system poles than that of system zeros, and thus is less susceptible to noise disruptions. Both damage indices are ineffective damage location indicators in this case, as the magnitude of change in feature values as a result of damage is more or less the same for the two sensing locations, despite one is much closer to the damage than the other.

4.2 Damage identification results from univariate AR residuals based method

As stated in Section 2, damage detection can also be based on the autocovariance/autocorrelation function of the residuals. Fig. 5 contains the ACF plots of the residuals obtained from applying the AR modeling to measurements at sensor 4 and 8. After the structure is damaged, the absolute values of the ACF at non-zero time lags increase significantly. Accordingly, an increase in Ljung-Box statistic of the residual series will also be observed. Due to space limitations, however, its results are not presented here.

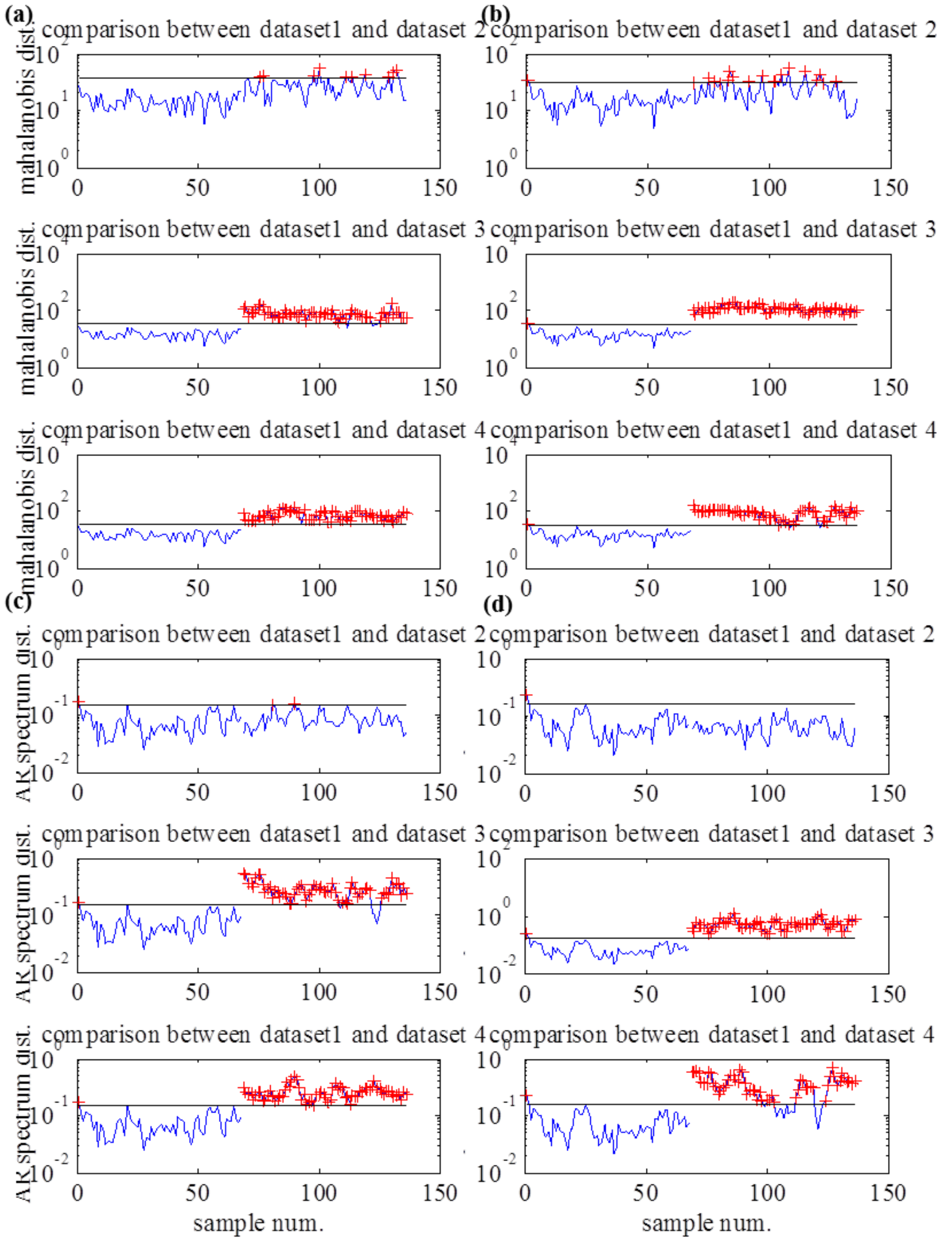

Fig. 4 Damage identification results using AR coefficients based method. (a) and (c) are the Mahalanobis distance and Cosh spectral distance features obtained from measurements at sensor 4, while (b) and (d) are those from sensor 8.

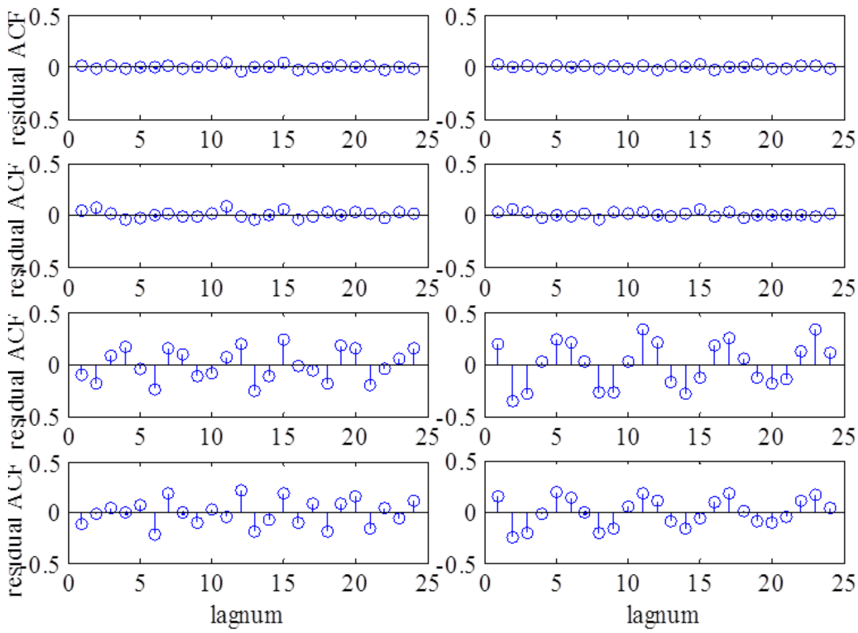

Fig. 5 Damage identification results using AR residual based method. (a) contains the residual autocorrelation function plots obtained from measurements at sensor 4, with subplot 1-4 corresponding to dataset 1-4. (b) contains plots obtained from sensor 8 using the same procedure. Dataset 1 is used as the baseline here.

4.3 Damage identification results from MISO ARX modeling method

Here the acceleration measurements in the vertical direction from sensor node 3 and 8 are used as output in the ARX model, and measurements from their respective neighboring nodes will be used as the model input. Table 1 summarizes the normalized residual standard deviations and the log Ljung-Box statistics computed from signals 1-4. In the first column of the table, the number before the slash represents the output sensor node, where those after the slash are the input nodes. And for the subscripts of these numbers, x stands for measurements in the vertical direction, and y the horizontal direction.

Damage localization can be achieved by comparing results from models constructed with different output nodes (e.g. node 3 vs. 8), and/or models with different combinations of exogenous inputs; a more significant increase in feature values can be observed for models using the sensor 8 as the output node, and for the models that include all neighboring node responses as inputs. Note that the artificial damage occurs between sensor 1 and 8. Because here only 10 truss nodes are instrumented with accelerometers, results from other nodes cannot be presented as the responses of several of their neighbor nodes are unknown.

Table 1 Results from the MISI ARX method. The standard deviations are all normalized with respect to the baseline.

		dataset 1	dataset 2	dataset 3	dataset 4
$3_x/1_x,2_x,4_x,5_x,10_{x,y}$	normalized σ_e	1.000	0.962	1.576	1.460
	log Ljung-Box statistic	5.504	6.070	8.706	7.999
$8_x/1_x,6_x,7_x,9_x,10_{x,y}$	normalized σ_e	1.000	0.990	2.127	1.905
	log Ljung-Box statistic	5.107	5.873	9.596	9.696
$3_x/1_x,2_x,4_x,5_x$	normalized σ_e	1.000	0.937	1.257	1.254
	log Ljung-Box statistic	5.618	6.155	8.153	7.665
$8_x/6_x,7_x,9_x,10_{x,y}$	normalized σ_e	1.000	0.993	1.479	1.264
	log Ljung-Box statistic	5.210	6.148	9.563	8.602

4.4 Damage identification results from mutual information method

The damage index based on mutual information also succeeds to identify the damage existence and location. Fig. 6 is a series of plots containing the estimated mutual information obtained from vertical acceleration measurements from pairs of adjacent nodes. Five estimates are computed from each dataset. In each of these plots, the features are normalized with the mean from the baseline sample group to facilitate cross-comparison. It can be observed that the damage features based on the node pairs located in the vicinity of damage exhibited the most significant change in their values. In addition, the estimates based on nodes that do not lie in the same vertical pane are more sensitive to damage than the rest, indicating that the damage occurs at the interior connections.

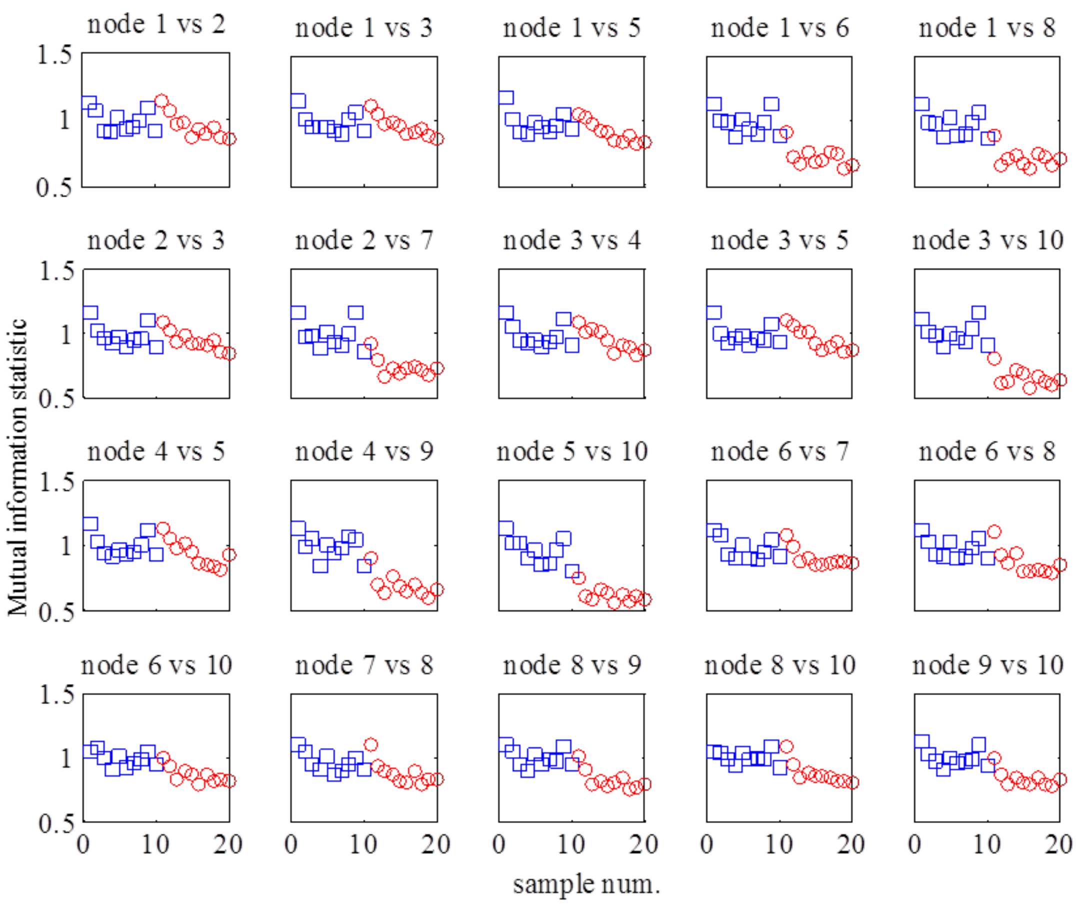

Fig. 6 Damage identification results from mutual information method. The blue squares are features from the undamaged state, and the red circles are from damaged state.

5. Conclusion

Two damage detection methods using time series analysis on responses from several adjacent nodes are described and applied to acceleration measurements on a space truss model here. The first method employs MISO ARX modeling, while the other method utilizes the mutual information concept for feature construction. Their performances are then compared with those of the established damage identification techniques based on AR modeling of responses from a single sensing channel. It is observed that the multivariate time series analysis produces viable damage indices and in the meanwhile is able to predict damage location with greater accuracy.

Univariate AR modeling algorithms has certain advantages when applied for damage identification; they are computationally efficient, suitable for on-sensor-board data processing, and sensitive to small scale damages. As seen in section 4, their application to detect damage existence in the truss model is altogether successful. However, they are not effective at damage localization. Moreover, since this family of algorithms monitors only the statistical properties of the measurements at a single sensor node, it is susceptible to changes in operation conditions that do not concern the structure itself. Such innate false-positive characteristic of these algorithms will make them unreliable for practice.

The MISO ARX modeling algorithm is slightly more sophisticated than the univariate AR method, as more parameters need to be estimated in the former case. But then, the damage localization capability of the algorithm is visibly improved by adding the neighbor nodes responses as external outputs in the model. Recognizing the increased computational cost, the mutual information damage index is introduced and applied to the specimen. This method is of a simpler formulation than the ARX algorithm, yet nonetheless achieves the aim of damage detection/localization as shown in section 3.

In summation, the multivariate time series analysis approach is proposed here as an intermediate approach that seeks to combine the merits of traditional system identification and univariate time series modeling. By including several responses from adjacent nodes, model parameter estimation becomes less affected by excitation variation; yet the model is still constructed on measurements from only a part of the structure, thus retaining the sensitivity to local damage and algorithm computation efficiency. The experimental application used here has confirmed the effectiveness of proposed algorithms, and they will be further examined through implementation on different types of structures in the future.

References:

[1] Alvin, K. F., Robertson, A. N., Reich, G. W. and Park, K. C. Structural system identification: from reality to models, *Computers & Structures*, Volume 81, Issue 12, pp. 1149-1176, 2003.

[2] Brockwell, P. J., Davis, R.A., Time Series: Theory and Methods, 2nd Edition, Springer, NY.

[3] Cover, T. M. and Thomas, J. A., Elements of Information Theory, 2nd Edition, Wiley Series in Telecommunications and Signal Processing, 2006.

[4] Doebling, S. W., Farrar, C. R., Prime, M. B., and Shevitz, D. W., Damage Identification and Health Monitoring of Structural and Mechanical Systems from Changes in Their Vibration Characteristics: A Literature Review. *Distribution*, 1996.

[5] Farrar, C. R., Duffey, T. A., Doebling, S. W., Nix, D. A. A Statistical Pattern Recognition Paradigm for Vibration-based Structural Health Monitoring, *2nd International Workshop on Structural Health Monitoring* ,Standford , CA, 1999.

[6] Fugate, M. L., Sohn, H. and Farrar, C. R., Vibration-based Damage Detection Using Statistical Process Control, *Mechanical Systems and Signal Processing*. Vol. 15, No. 4, pp.707-721, 2001.

[7] Gul, M. and Catbas, F.N., Statistical Pattern Recognition for Structural Health Monitoring Using Time Series Modeling: Theory and Experimental Verifications, *Mechanical Systems and Signal Processing*, Vol. 23, Issue 7, pp2192-2204, 2009.

[8] Ljung, G. M. and Box, G.E.P., On A Measure of Lack of Fit in Time Series Models, *Biometrika*, Vol. 65, No. 2, pp. 297-303, 1978.

[9] Ljung, L, System Identification: Theory for the User, Prentice Hall, Inc., New Jersey, 1987.

[10] Messina, A., Williams, E. J. and Contursi, T. Structural Damage Detection by A Sensitivity and Statistical-Based Method, *Journal of Sound and Vibration*, Vol.216, Issue 5, pp. 791-808, 1998.

[11] Sohn, H. and Farrar, C.R., Hunter, N.F. and Worden, K. "Structural Health Monitoring Using Statistical Pattern Recognition Techniques." *Journal of Dynamic Systems, Measurement and Control*, Vol. 123, Issue 4, pp. 706-711, 2001.

[12] Wei, B., Gibson, J.D., 2000, Comparison on Distance Measures in Discrete Spectral Modeling, *Proceedings of IEEE Digital Signal Processing Workshop*, pp1-4, 2009.

[13] Yao, R. and Pakzad, S. N., Data-Driven Methods for Threshold Determination in Time-Series Based Damage Detection, *Proc. of 20th Annual Structures Congress*, Las Vegas, NV, April 2011.