Data-driven Structural Damage Diagnosis

S.G. Shahidi[1], S.N. Pakzad[2]

1   PhD Candidate, Dept. of Civil and Environmental Engineering, Lehigh University, United States.
    E-mail: sgs310@lehigh.edu
2   Associate Professor, Dept. of Civil and Environmental Engineering, Lehigh University, United States.
    E-mail: pakzad@lehigh.edu

ABSTRACT
Detection of damage in structures and infrastructure systems is the main objective of Structural Health Monitoring projects. Data-driven procedures constitute one category of structural damage diagnosis methods where signal processing or time series techniques are employed on the measured responses to establish damage sensitive features. Then a statistical framework is used to evaluate the significance of potential changes in the extracted features.
This paper presents several data-driven damage identification methodologies. In these algorithms, feature extraction is completed by applying general regression models to data collected through clusters of sensors. For systems with linear topology, substructural regression modeling is also performed on time- and frequency-domain transforms of the measured signals to estimate local stiffness of the structure as damage features. Subsequently, change point analysis is utilized to statistically determine the significance of changes in the extracted features in order to distinguish between changes caused by damage or environmental factors and measurement noise. A toolsuite is developed that facilitates application of such data-driven damage detection algorithms and improve the reliability of the diagnosis results by offering several combinations of regression models, damage features, and statistical tests to process the monitoring data.

KEYWORDS: *structural health monitoring, damage detection, change point analysis, data-driven, damage feature*

1. INTRODUCTION

Detection of early stage damage in the constructed structures and infrastructure systems is one of the central goals in Structural Health Monitoring projects. This goal is accomplish through comparison of structural properties or features from an unknown structural health condition with those from a known and presumably undamaged state of the system. The damage detection algorithms can be categorized into physics-based and data-driven techniques based on the underlying model that are used to extract damage sensitive features for this comparison. In the first category of these methods, Finite Element (FE) calibration techniques are employed to update uncertain parameters of an FE model [1-4]. The calibration is performed with data from different states of structure to investigate whether model parameters would deviate from their baseline. While such damage detection methods can be computationally expensive, they could be advantageous by providing information about the severity of damage in addition to its occurrence and location. In the second category, however, damage identification is performed as outlier detection on features from the structural response extracted through signal processing or time series techniques [5-8]. Then a statistical framework is used to evaluate the significance of potential changes in the extracted damage features. Such data-driven techniques have gained significant attention due to their low computational burden, as well as their capability to signify highly localized damage scenarios in the structure. This paper presents several data-driven damage identification methodologies to analyze univariate and multivariate monitoring datasets. Next section describes the mathematical models used to create damage sensitive features from the measured structural responses.

2. DATA-DRIVEN DAMAGE FEATURES

2.1. General regression models

Application of time series regressive models in damage feature extraction has shown promising results in SHM research [9-11]. One of the time series models used for this purpose is Auto-regressive (AR) model which establishes a relationship between past and present samples of a signal. This model is shown in Eq. 2.1, where p,

α_p and ε denote the order, regression coefficients, and residuals of the (AR) model, respectively. Functions of regression coefficients as well as the residuals have been used as damage sensitive features in previous research [12-13]. These models are beneficial mainly because they can be used on univariate measurements possibly to detect occurrence of a change in structure's behavior. With the use of multivariate sensor networks, performance of these models can be extended to localizing the damage as well.

$$y_i(n) = \sum_{p=1}^{P} \alpha_p y_i(n-p) + \varepsilon(n) \tag{2.1}$$

AR models can be extended to Auto-regressive models with exogenous (ARX) input term as shown in Eq. 2.2. These models establish a relationship between two signals: input (y_i) and output (y_j) of the model. This definition can be employed to relate responses from two locations on the structure to create damage features [14].

$$y_j(n) + \sum_{p=1}^{P} \alpha_{jp} y_j(n-p) = \sum_{q=0}^{Q} \alpha_{iq} y_i(n-q) + \varepsilon(n) \tag{2.2}$$

2.2. Substructural regression models

The band-limited characteristics of mass, stiffness, and classical damping matrices in structural systems with linear topology (i.e. shear building and bridge systems) can be used to estimate the stiffness properties from measured response signals, and therefore detect possible damage in form of local stiffness reduction [15]. In such type of structures, Eq. 2.3 holds for the response measurements at floor i-1, i, and i+1. Therefore, given the mass of middle node (m_i), its excitation, as well as acceleration response, displacement, and velocity of the three neighboring nodes, the stiffness parameters are estimated through regression.

$$m_i \ddot{u}_i = k_{i,i-1}\left(u_{i-1} - u_i\right) + k_{i,i+1}\left(u_{i+1} - u_i\right) - c_i \dot{u}_i - c_{i-1}\dot{u}_{i-1} - c_{i+1}\dot{u}_{i+1} + p_i \tag{2.3}$$

Since excitation is often not available, Eq. 2.3 can be rewritten in form of Eq. 2.4 by taking the correlation of both sides of Eq. 2.3 with middle node acceleration response ($\ddot{u}_i$) to eliminate p_i term. In this equation, τ is a positive time lag.

$$m_i E\left[\ddot{u}_i(t)\ddot{u}_i(t-\tau)\right] = k_{i,i-1}E\left[\left(u_{i-1}(t) - u_i(t)\right)\ddot{u}_i(t-\tau)\right] + k_{i,i+1}E\left[\left(u_{i+1}(t) - u_i(t)\right)\ddot{u}_i(t-\tau)\right]$$
$$- c_i E[\dot{u}_i(t)\ddot{u}_i(t-\tau)] - c_{i-1}E[\dot{u}_{i-1}(t)\ddot{u}_i(t-\tau)] - c_{i+1}E[\dot{u}_{i+1}(t)\ddot{u}_i(t-\tau)]. \tag{2.4}$$

This formulation can be used in time- and frequency-domain to estimate stiffness and damping parameters. The estimation process has lower computational cost in the frequency-domain, since in the time-domain formulation displacement and velocity responses – when unavailable – are reconstructed from acceleration signals, whereas in the frequency-domain formulation they are turned into regression constant [15].

3. CHANGE POINT ANALYSIS

When damage sensitive features are created, statistical frameworks are needed for outlier detection. This section reviews some of the statistics that can be adopted to test the significance of change in damage features and distinguish between changes that are result of damage and common changes due to environmental factors and measurement noise. These change point statistics are mainly used to test any potential change in the mean of the vector of observations; here a vector of damage sensitive features with N elements (X_m where m=1,2,..,N). Eq. 2.5 shows Cumulative Sum (CUSUM) statistics that calculates sum of the differences between each observation and their overall average. This statistics start at zero ($S_0=0$) and eventually goes to zero. An upward slope in CUSUM chart indicates a period when observations tend to be above the average, whereas a downward slope shows that observations are less than the average. Therefore, a distinct change in the slope of CUSUM plot signifies a change in the vector of features [16, 17].

$$S_m = S_{m-1} + (X_m - \frac{\sum_{i=1}^{N} X_i}{N}) \tag{2.5}$$

Another change detection statistics is Exponential Weighted Moving Average (EWMA), shown in Eq. 2.6. This statistic finds a weighted average of all the previous observations, every time a new observation is monitored. The weight is applied by parameter λ, which is typically set between 0.05 and 0.25. Z_0 is an estimate of in-control process mean usually calculated based on historical data. Z_m statistics are compared to an upper and lower control limits (UCL, LCL) based on the mean and variance of in-control process. This statistics is different from the others discussed here in that it can be used for online change detection, since it does not use the information from the entire vector of observations [16, 17].

$$Z_m = \lambda X_m + (1 - \lambda) Z_{m-1} \tag{2.6}$$

Mean Square Error (MSE) is another change statistics shown in Eq. 2.7. MSE_m finds how well two splits of the data made at the mth observation ($1 < m < N-1$) would fit their estimated averages. In effect, the point where MSE shows a V-shaped global minimum indicates a change point in the vector of observations [17].

$$MSE_m = \sum_{i=1}^{m} (X_i - \frac{\sum_{j=1}^{m} X_j}{m}) + \sum_{i=m+1}^{N} (X_i - \frac{\sum_{j=m+1}^{N} X_j}{N-m}) \tag{2.7}$$

Another method to test potential changes in the mean of the damage features is by establishing a test statistics which follows a Student's t-distribution with N-2 degrees of freedom. There are three assumptions for this formulation: (1) samples come from a parent population that is normally distributed, (2) the two sample groups are from populations with equal variances, and (3) sample observations are independent. Eq. 2.8 shows the details of this statistics. In this equation Sp_m denotes the pooled standard deviation of two splits of the vector of observations [14].

$$t_m = \frac{\sum_{j=1}^{m} X_j - \sum_{j=m+1}^{N} X_j}{Sp_m \sqrt{\frac{1}{m} + \frac{1}{N-m}}} \tag{2.8}$$

Finally, a test statistic can be established based on the maximum likelihood function of observations. This statistics which is described in Eq. 2.9 and 2.10, finds the difference between maximum likelihood function of the vector of observations with those from its two segments creating through a split at the mth observation. In Eq. 2.10, $l_{n1:n2}$ denotes the maximum likelihood function of observations from Xn_1 to Xn_2 ($n_2 > n_1$) whose variance is $\hat{\sigma}_{n1n2}^2$. This statistics is first normalized by estimated expected value of in-control statistics to create normalized likelihood test statistics (NLRT), and then is compared with a control limit based on χ^2 distribution [14].

$$lrt_m = -2(l_{1:N} - l_{1:m} - l_{m+1:N}) \tag{2.9}$$

$$l_{n_1:n_2} = -\frac{n_2 - n_1 - 1}{2} \left(\ln(2\pi \hat{\sigma}_{n_1 n_2}^2) + 1 \right) \tag{2.10}$$

4. DIT: DAMAGE IDENTIFICATION TOOLSUITE

As described in the previous sections, several damage features, and change point statistics can be established once structural responses are collected at monitoring intervals. Since in data-driven damage detection, a critical step is to infer about occurrence and location of damage by statistically analyzing the extracted features, it is expected that using combination of different features and tests improve the accuracy of the damage diagnosis procedure. Therefore, a Matlab-based data-driven damage identification toolsuite (DIT) is developed to facilitate comparison of several damage identification techniques described earlier. DIT offers several change point statistics to be created from damage features for which control threshold is constructed using resampling

methods (bootstrapping and permutation) as well as test of significance when applicable [18]. Fig. 4.1 displays this toolsuite which will be described in more details in the following sections through examples of a simulated bridge model, and a scaled two-bay steel frame. DIT can be downloaded from dit.atlss.lehigh.edu.
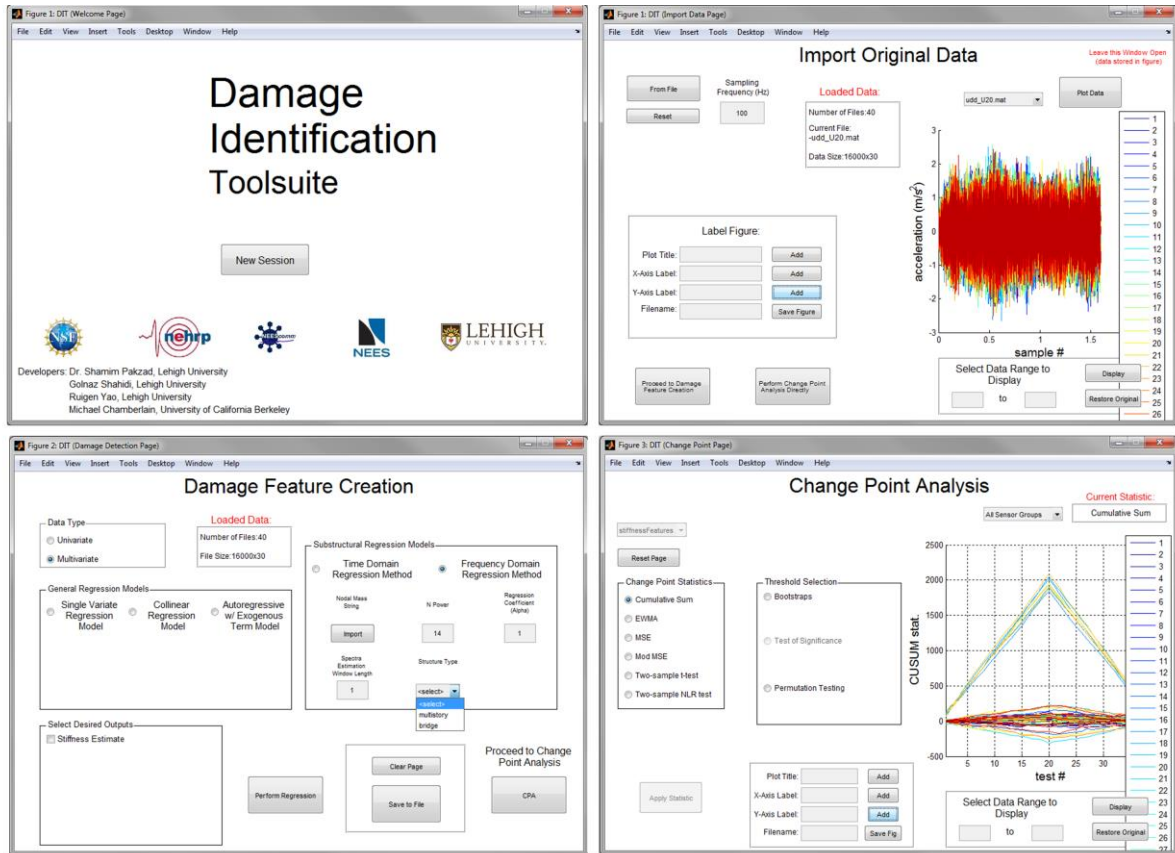

Figure 4.1 DIT main windows

5. DAMAGE DETECTION IN A SIMULATED BRIDGE STRUCTURE

This section presents the damage detection in a shear bridge model with 30 degrees of freedom (dof) and 31 elements. Random excitations were applied on the dofs of the bridge model so they vibrate with 0.05g acceleration standard deviation. In order to consider the instrumental noise, a 5% random white noise signal was added to each acceleration signal. Damage in the model was simulated by 25% reduction in stiffness of 9th, 10th, 21st, and 22nd elements. Twenty sets of simulation were performed on healthy and damage configuration of the bridge model. The simulated datasets along with the nodal mass string were imported to DIT for feature creation and change point analysis. For this example, with assumption of known mass, substructural regression model in frequency-domain was selected to estimate the local stiffness parameters as damage features. Vectors of damage sensitive features were then tested to find any possible change in their means.

The change detection analysis showed clear change points for the statistics related to 9th, 10th, 21st, and 22nd elements. Fig. 5.1 shows EWMA control charts for stiffness parameters of 2nd, 9th, and 26th elements along with their control limits. In calculating this statistics it was assumed that the features from first 10 simulations were coming from a reference (known) state. This assumption was used to estimate the mean and standard deviation of the in-control process. Fig. 5.1 shows that this statistics is successful in identifying the time (i.e. 21st test) and location of damage. The estimated stiffness parameters were also tested through CUSUM control charts. The results are presented in Fig. 5.2 showing a clear change point is detected in the statistics of the damaged elements. Through DIT the tests statistics can also be compared with thresholds based on permutation or resampling of the vector of observations. Fig. 5.2(b) and (c) show the CUSUM statistics of 22nd and 26th stiffness parameters over the statistics from 1000 bootstrapped samples. It is seen that for the 22nd element other thresholding methods confirm the occurrence of damage, whereas for the statistics of 26th element when damage thresholds are set based on bootstrapping methods, there is not enough evidence to infer the change.
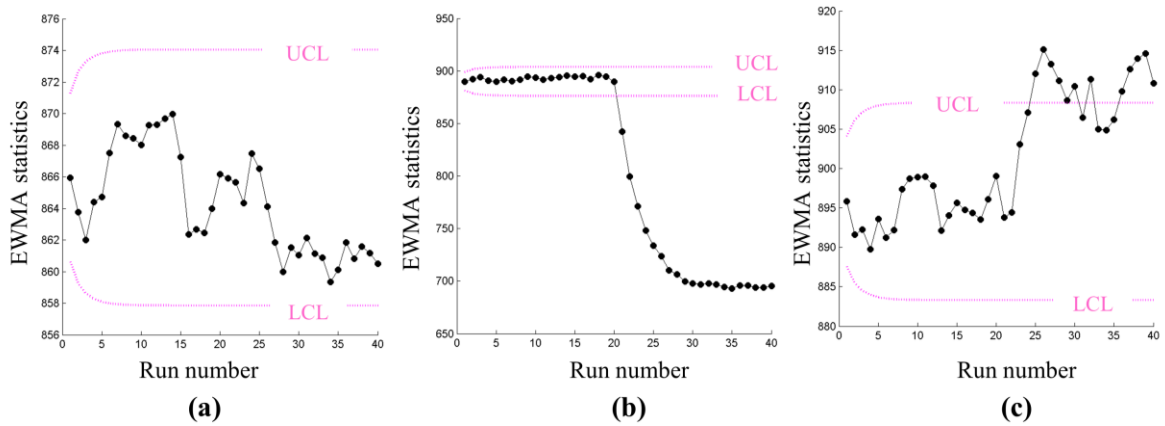
Figure 5.1 EWMA control charts for estimated local stiffness of the bridge model:
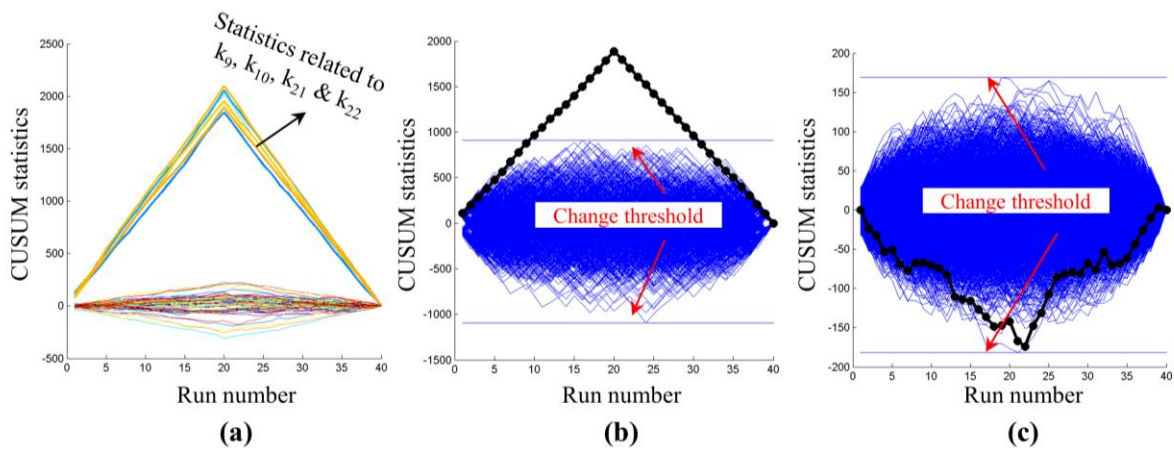(a) 2nd element, (b) 9th element, and (c) 26th element

Figure 5.2 CUSUM control charts for estimated local stiffness of the bridge model:
(a) all the elements, (b) 22nd element, and (c) 26th element

6. DAMAGE DETECTION IN A SCALED STEEL FRAME

This section presents the damage detection process on a two-bay steel tube frame test-bed constructed at the laboratory of Advanced Technology for Large Structural Systems at Lehigh University. This frame (shown in Fig. 6.1) was built as a test-bed for damage detection, mainly to represent typical building frames or bridge girders. It has nine interchangeable sections, 0.2m in length that can be changed throughout the frame in order to simulate damage. These interchangeable sections have different cross-sectional properties than the healthy state which correspond to 20% reduction in member stiffness. The specimen was instrumented with 21 wired accelerometers, labelled in Fig. 6.2 with L, C or R on left, center and right portions of the frame. In order to dynamically excite the frame, impact loading is chosen as the excitation method for this implementation. The impact amplitude was limited to ensure that the linear behavior assumption for the experimental frame holds. Therefore, the acceleration response of frame is recorded while the frame is struck with a hammer on the right column and the frame freely vibrates on its own.

During testing, there were a total of 40 sets of data collected; 20 runs performed on the healthy configuration of the specimen, while 20 run were conducted after the interchangeable section at location of R5 (see Fig. 6.2) was switched with a section with less stiffness. Throughout the damage detection process it was assumed that the first 10 runs from undamaged system are from a known healthy baseline; therefore, the timing of damage should be detected at the 10th runs from unknown state of the system (21st run overall). It should be noted that the data

measured with sensors L1, C3, C5 and C9 were excluded from the damage detection process, since the preliminary inspection of the measured signals revealed their faulty behavior.

ARX and AR models of order 4 were used for damage feature extraction. In the ARX models acceleration from each sensor was paired with the signals from their closest neighbor; therefore, a total number of 15 pairs were tested. The regression coefficients of the ARX and AR models were then compressed to create scalar damage features such as Mahalanobis distance and angle coefficients [14, 17]. These scalar damage features were then tested in the change detection analysis window. Figure 6.3 shows the results of this process with a change threshold corresponding to 99% confidence level. This figure shows that ARX-based damage features are detecting the damage at its correct time at the 10th run; the maximum t-test statistics and NLRT statistics that are above the threshold infer a significant change in the vector of damage features, i.e. damage. It is also seen that the statistics from the right side of the frame are considerably larger, and the highest test statistics are form R5-R6 sensor pair. Fig. 6.3(c) shows the t-test statistics based on AR coefficients which shows that although the statistics imply a significant change at correct timing, damage is not localized to its true location. Since the other AR-based features did not indicate any kinds of change, the damage diagnosis based on AR models is not successful in this experiment.
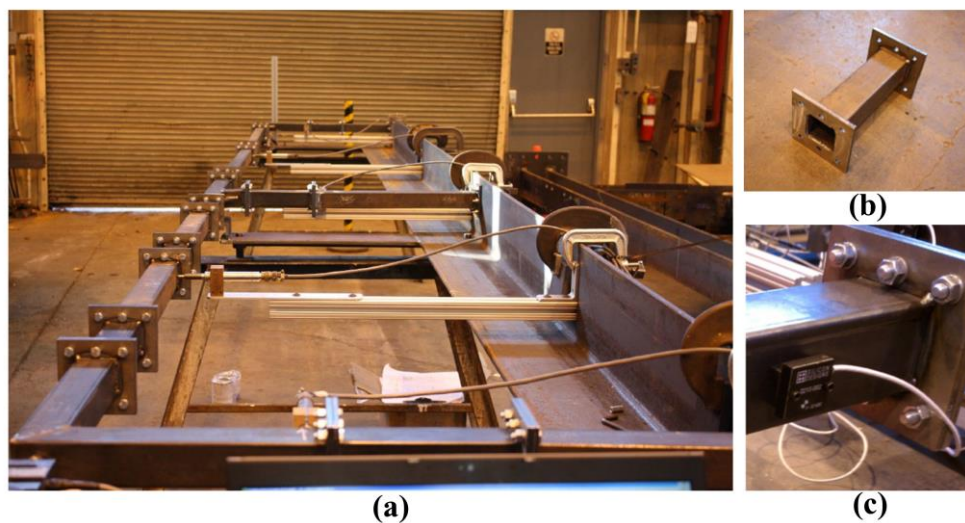

Figure 6.1 Experimental setup: (a) scaled frame, (b) switch-out member, (c) wired accelerometer
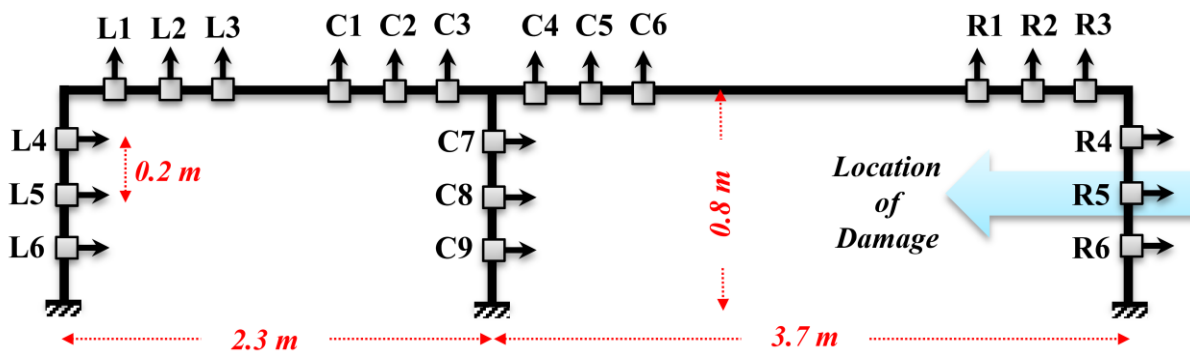

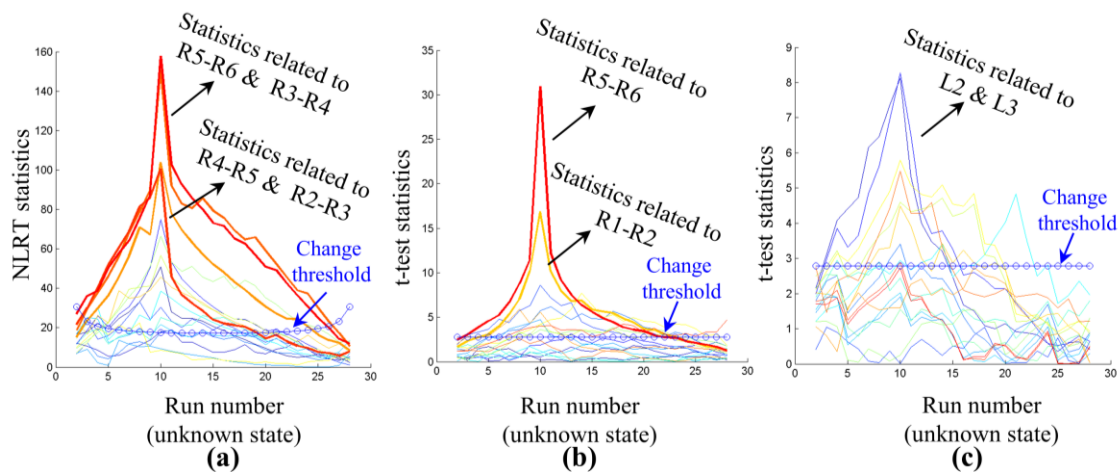
Figure 6.2 Sketch of the specimen and location of damage

Figure 6.3 Control charts of extracted features: (a) NLRT statistics of the Mahalanobis distance of ARX coefficients, (b) t-test statistics of angle features, and (c) t-test statistics of the Mahalanobis distance of AR coefficients

7. CONCLUSIONS

This paper presents data-driven structural damage detection methodologies. In these techniques, the measured structural responses are processed to extract features that are sensitive to damage, yet robust to measurement noise and environmental variability. Regression analysis is performed to extract features from measured signals. Several auto-regressive models are available for this purpose. For structures with linear topology, stiffness parameters of the system can be estimated by using certain substructural regression models. Coefficients of multivariate regression models are condensed into scalar damage features. Vectors of such features that are extracted from different monitoring datasets are then analysed through change point detection in order to statistically test the significant of any potential change. These damage detection methods are implemented in a Matlab-based graphical interface (DIT) in order to facilitate application of several combination of training models, damage features, and change point methods. This paper describes damage detection using DIT through two case studies: (1) a simulated shear bridge structure, and (2) a two-bay scale steel frame tested in the laboratory. It was observed that incorporating multiple mathematical models, damage sensitive features and change detection tests improve the overall performance of these model-free structural damage detection techniques. This shows potential application of such methodologies for automated structural damage diagnosis.

REFERENCES

1. Huth, O., Feltrin, G., Maeck, J., Kilic, N., and Motavalli, M. (2005). Damage Identification Using Modal Data: Experiences on a Prestressed Concrete Bridge. *Journal of Structural Engineering*, **131:12**, 1898–1910,
2. Jaishi, B., and Ren, W. X. (2006). Damage detection by finite element model updating using modal flexibility residual. *Journal of sound and vibration*, **290:1-2**, 369-387.
3. Shahidi, S.G. and Pakzad, S.N. (2014). Generalized Response Surface Model Updating Using Time Domain Data, *ASCE Journal of Structural Engineering*: A4014001. (DOI: 10.1061/(ASCE)ST.1943-541X.0000915)
4. Ebrahimian, H., Astroza, R., and Conte, J. P. (2015). Extended Kalman filter for material parameter estimation in nonlinear structural finite element models using direct differentiation method, *Earthquake Engineering and Structural Dynamics*. DOI: 10.1002/eqe.2532

5. Posenato, D., Kripakaran, P., Inaudi, D. and Smith, I.F.C. (2010). Methodologies for Model-free Data Interpretation of Civil Engineering Structures, *Computers and Structures*, **88:7-8**, 467-482, 2010.

6. Torres-Arredondo, M.A, Tibaduiza, D.A., Mujica, L.E., Rodellar, J., and Fritzen, C. P. (2014) Data-driven Multivariate Algorithms for Damage Detection and Identification: Evaluation and Comparison, *Structural Health Monitoring*, **13(1)**, 19-32.

7. Kumar, R.P., Oshima, T., Mikami, S., Miyamori, Y., & Yamazaki, T. (2012). Damage Identification In a Lightly Reinforced Concrete Beam Based on Changes in The Power Spectral Density, *Structure and Infrastructure Engineering*, **8**, 715–727.

8. Labuz, E.L., Chang, M. and Pakzad, S.N. (2010). Local Damage Detection in Beam-Column Connections Using a Dense Sensor Network, *Proceedings of the ASCE's 42nd Structures Congress*. Orlando, FL. 3143-3154. (DOI: 10.1061/41130(369)282)

9. Cavadas, F., Smith, I.F.C., Figueiras, J., Damage Detection Using Data-Driven Methods Applied to Moving-Load Responses*, Journal of Mechanical Systems and Signal Processing*, **39:1–2**, 409–425, 2013.

10. Dorvash, S., Pakzad, S.N., Labuz, E.L., Chang, M., Li, X., and Cheng, L. (2010) Validation of a Wireless Sensor Network using Local Damage Detection algorithm for Beam-Column Connections, *Proceedings of the SPIE Sensors and Smart Structures Technologies*, San Diego, CA. **7647**, 76719-1.(DOI: 10.1117/12.847581)

11. Dorvash, S., Pakzad, S.N., and Labuz, E.L. (2013). Statistics Based Localized Damage Detection using Vibration Response, *Smart Structures and Systems, An International Journal*, **14:2**.

12. Yao, R., and Pakzad, S.N. (2012). Autoregressive Statistical Pattern Recognition Algorithms for Damage Detection in Civil Structures, *Journal of Mechanical Systems and Signal Processing*, **31**, 355-368.

13. Figueiredo, E., Figueiras, J., Park, G., Farrar, C.R., and Worden, K. (2011). Influence of the Autoregressive Model Order on Damage Detection, *Computer-Aided Civil and Infrastructure Engineering*, **26:3**, 225–238.

14. Shahidi, S.G., Nigro, M.B., Pakzad, S.N., and Pan, Y., Structural Damage Detection and Localization Using Multivariate Regression Models and Two-sample Control Statistics, *Structure and Infrastructure Engineering*, In Press. (DOI: 10.1080/15732479.2014.949277)

15. Yao, R., and Pakzad, S.N. (2014). Time and Frequency Domain Regression Based Stiffness Estimation and Damage Identification, *Structural Control and Health Monitoring*, **21:3**, 356-380

16. Dorvash, S., Pakzad, S.N., Labuz, E.L., Ricles, J.M., and Hodgson, I.C., Localized Damage Detection Algorithm and Implementation on a Large-Scale Steel Beam-to-Column Moment Connection, *(EERI) Earthquake Spectra*, In Press (DOI: 10.1193/031613EQS069M)

17. Nigro, M.B., Pakzad, S.N., and Dorvash, S. (2014) Localized structural damage detection: a change point analysis, *Computer-Aided Civil and Infrastructure Engineering*, **29:6**, 416-432.

18. Pakzad, S.N., Shahidi, S.G., Yao, R., Chamberlain, M.B.W. (2014). Damage identification toolsuite (DIT) user's guide. ATLSS 14-05 (dit.atlss.lehigh.edu)