Data-driven methods for threshold determination in time-series based damage detection

Ruigen Yao, Graduate Student, Department of Civil and Environmental Engineering, Lehigh University, Bethlehem, PA 18015. ruy209@Lehigh.edu
Shamim N. Pakzad, Assistant Professor, Department of Civil and Environmental Engineering, Lehigh University, Bethlehem, PA 18015. pakzad@Lehigh.edu

ABSTRACT

Structural vibration monitoring has received a lot of attention from the research community in the past few years. The objective is to create automatic structural assessment techniques that can be realized through programmed vibration analysis. Till now many vibration-based damage features have been proposed, yet to truly automate the damage identification process, reliable damage threshold construction techniques are also need. In this paper, two data-driven methods based on resampling and nearest neighbor rule are applied for threshold construction for damage features from autoregression (AR) analysis of vibration signals. Both threshold calculation techniques are rooted in empirical feature probability estimation. The proposed thresholds are then tested on features extracted acceleration measurements collected from a 5 DOF test specimen. The resampling method is applied to Mahalanobis distance of AR model coefficients, while the nearest neighbor rule is used on a combination of coefficient distance feature and the residual autocorrelation feature. Both methods perform well in this case study.

KEYWORDS: Structural health monitoring; damage detection; time series analysis; autoregressive modeling; statistical pattern recognition; Mont Carlo method.

INTRODUCTION

For years researchers have been working on automatic damage detection in the hope of reducing a structure's maintenance cost by minimizing human involvement in the process. The rapid development in vibration data acquisition systems for the infrastructure has stimulated a vast interest in vibration-based damage detection and structural assessment, and accordingly a lot of literature are produced [Doebling et al. 1998]. Modal properties are among the earliest damage indicators adapted to this end [Alvin et al. 2003; Doebling 1998], but were found to be insensitive to local damage and computationally expensive. Recently, time series analysis for vibration signals is being investigated in hope of establishing a more sensitive vibration-based damage detection approach [Farrar et al. 1999].

Time series analysis techniques have long been applied to fault detection in mechanical systems [Fassois and Sakellariou, 2007] before they were introduced into the field of civil engineering. This is essentially a data-driven approach that will 'let the data speak for itself'. The general strategy of time- series based damage detection includes five steps: 1) obtain data from baseline/healthy state structure; 2) extract a statistical quantity/damage feature Q from the data via time series analysis; 3) obtain data from current/unknown state structure; 4) extract the Q quantity from the data using the same method; 5) compare Q from the unknown state with that from baseline state using hypothesis testing. If the difference is statistically significant, the system is deemed damaged. The critical quantity Q can be a function of estimated time series model properties or signal spectrum.

Several time-series analysis methods have already been employed to identify damage in civil structures [Sohn and Farrar, 2001; Worden and Manson, 2007; Gul and Catbas, 2009]. The results reported in these papers are encouraging, since structural damage does result in a substantial change in the Q quantity in all cases. The main problem is that the extracted Q quantity can be affected by many other non-relevant factors such as environmental noise and excitation levels, and will suffer some oscillation even when the structural conditions remain the same. Therefore, it is important to set the damage threshold properly such that the number of false alarms and missed cases is minimized.

Statistical hypothesis testing [Koch 1999] is the recognized standard approach for threshold determination. It assumes that the features follow a certain probability distribution, and the threshold is set at a point beyond which the chance for a feature value to occur is small. This approach is theoretically optimal as long as the assumed feature distribution is valid. Hypothesis testing may do well for fault identification in machinery, as the excitation force is well known and the damage types are well-defined. For civil engineering applications, however, there are more uncertainties. When the probability distribution of damage features are too complex to be accurately represented by analytical distribution functions available, threshold constructed using hypothesis testing will yield poor results in damage identification and an alternative approach is needed.

In this paper a data driven threshold determination scheme using cross-validation and resampling techniques [Good, 1999] will be introduced and applied to the Mahalanobis distance [Mahalanobis 1936] feature described in Section 2. This feature is shown to be unstable, and its distribution is unknown. Yet the data-driven method still yields an effective threshold. Also the Mahalanobis distance feature is paired with the autocorrelation feature to improve the damage detection performance of time-series based methods. A nearest-neighbor approach [Duda et al. 2000] is used to set the significance threshold for the feature pair.

The paper is organized as follows: Section 2 is dedicated to a brief review of time-series analysis techniques for damage detection. The damage features can be obtained from either autoregressive (AR) model coefficients or AR model residuals, and hypothesis tests are devised according to their respective theoretical distributions.

In section 3, new data driven approaches for threshold determination are presented. Section 4 contains an application of the Mahalanobis distance feature to detect damage in a lab specimen, and the results using the new threshold construction scheme are presented. Section 5 includes a discussion of feature combination and its application to the same specimen as in section 4.

DAMAGE DETECTION METHODS USING TIME SERIES ANALYSIS

Autoregressive (AR) model is perhaps the most widely adopted time series analysis tool [Brockwell and Davis, 2002]. The definition of a univariate AR model of order p is as below:

$$x(t) = \sum_{j=1}^{p} \phi_{xj} x(t-j) + \epsilon_x(t). \quad (1)$$

In this equation, $x(t)$ is the time series to be analyzed, ϕ_{xj} are the AR model coefficients, and $\epsilon_x(t)$ is the model *residual*. This model basically attempts to express the value of signal at time t as a linear combination of its previous values up to lag p.

According to classical structural dynamics theory, the discretized structural response under random excitation can always be approximated by an AR process of large order. Damage features from AR modeling can be grossly divided into two categories; model coefficients based and model residual based. AR model coefficients can be estimated directly from the data using one of the standard algorithms, and the residual sequence can henceforth be obtained from eqn. (1). In the remainder of this section, two damage features, one from each category, will be presented.

Mahalanobis distance of AR model coefficients

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 $\sigma_e^2 \Gamma_p^{-1}$[Brockwell and Davis, 2009]. Therefore, a metric that represents the deviation in the probability space of normal distribution seems a good choice of damage feature. Mahalanobis distance is such a metric defined from the definition of multivariate normal distribution. The estimator of the Mahalanobis distance between a potential outlier vector x_ξ and baseline sample set can be obtained as

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

where $\bar{x}$ is the average of the baseline sample feature vectors, and $\hat{\Sigma}$ the estimated covariance matrix. When applying this method, the baseline signals are first segmented (often with large overlap) and for each segment an AR coefficient vector are estimated. Signals from current structural state are processed likewise and for each coefficient vector obtained its Mahalanobis distance to the baseline coefficients set will be computed. These Mahalanobis distance features are then compared with the Mahalanobis distances within baseline set. When the structural system is damaged, it is expected that the Mahalanobis distance feature for AR coefficients will increase significantly (Fig 1).

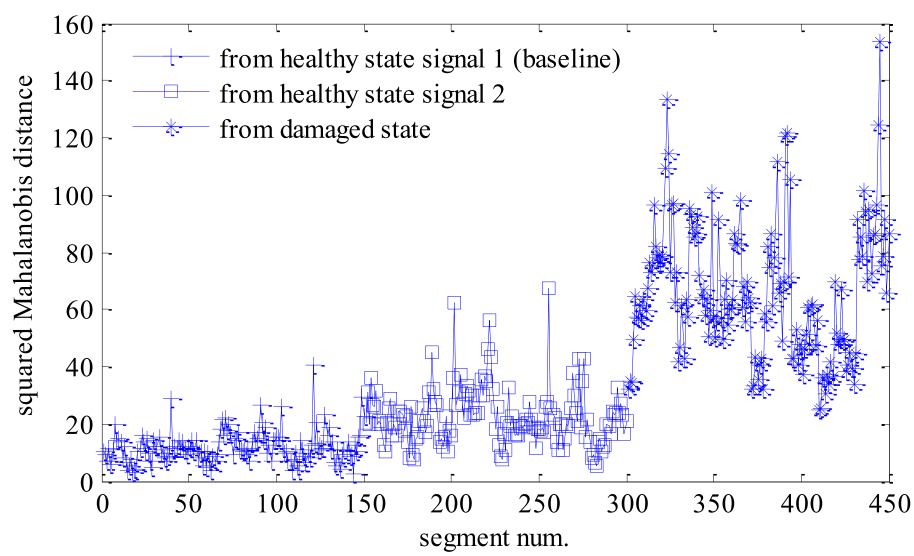

FIG. 1. MAHALANOBIS DISTANCE OF MODEL COEFFICIENTS. SEGMENT SIZE IS 350. THE OVERLAP BETWEEN CONSECUTIVE SEGMENTS IS 300.

Auto-correlation function (ACF) of AR model residuals

It is clear from (1) that if the AR model used to filter the signal is the same as the model from which the signal is generated, the residual series should be a white noise. Otherwise, the residual series will carry certain identifiable patterns that can be captured by its autocorrelation function. ACF can be estimated from residual sequence as:

$$\hat{\rho}(\tau) = \frac{\sum_{i=1}^{i=N-\tau+1}[\epsilon(i+\tau)-\bar{\epsilon}]\,[\epsilon(i)-\bar{\epsilon}]}{\sum_{i=1}^{i=N}[\epsilon(i)-\bar{\epsilon}]^2}.$$

From the baseline signal a baseline AR model can be estimated, which will then be fitted to new signals from unknown structural state. If the structural condition is unchanged, the ACF will resemble a Dirichlet delta function. It can be shown that for large N the sample autocorrelations of a white noise sequence at nonzero lags are approximately identically and independently distributed $N(0,1/n)$ [Brockwell and Davis, 2009], hence the 95% confidence bounds can be drawn at $\pm 1.96/\sqrt{n}$, as 1.96 is the .025 quantile of the standard normal distribution. Accordingly, the system can be identified as damaged when the number of ACF function value outside the bounds become statistically significant (Fig.2).

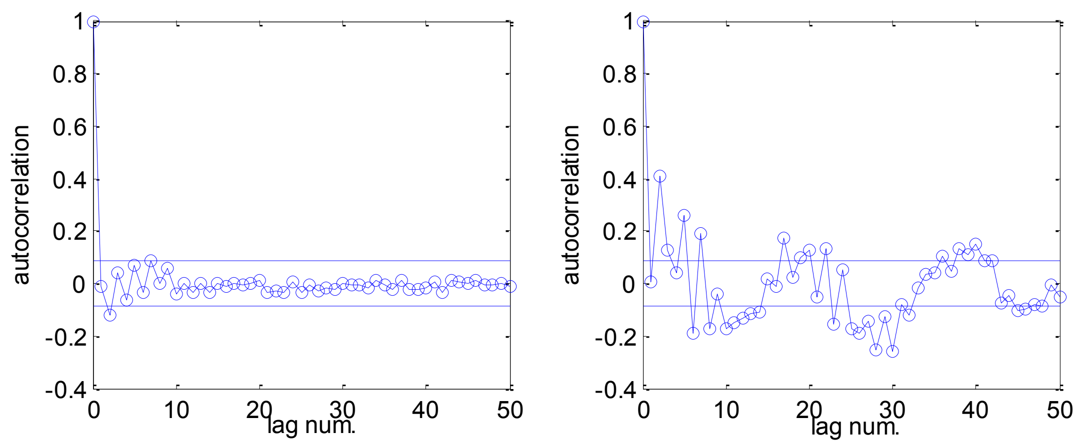

FIG. 2. RESIDUAL AUTOCORRELATION AS DAMAGE INDICATOR; LOTS OF OUTLIERS APPEAR WHEN MODEL NO LONGER FIT THE DATA WELL. THE LENGTH OF RESIDUAL SEQUENCE IS 520.

THRESHOLD CONSTRUCTION SCHEMES FOR DAMAGE FEATURES FROM TIME SERIES ANALYSIS

A desirable threshold is one that strikes up a balance between false alarms and missed cases. Now in most engineering practices this value is determined in an *ad hoc* manner. Statistical hypothesis testing has been tried for automatic threshold construction, but its effectiveness is not guaranteed unless the actual feature distribution is the same as assumed. Here, tow data-driven methods will be introduced for threshold construction.

Threshold calculated from resampling: the 'cross-one-out' method

From Fig. 2 it is clear that the Mahalanobis distance feature suffers from large fluctuations within the baseline sample set. Also, when the baseline samples are extracted from vibration signal segments with overlap, the sample set is in fact not a very good representation of the actual feature probability space. Therefore, hypothesis testing exploiting the multivariate normal distribution will tend to yield a conservative threshold.

The first problem is perhaps resulted from environmental variations and statistical modeling and estimation deficiencies and cannot be helped; one solution to the second problem is to use vibration signals without overlap for feature extraction. However, when the data available is limited, this method is not practicable.

To address this problem here a 'cross-one-out' resampling technique is adopted. First a segment is cut from the baseline signal at a random time point and reserved for testing, and sample segments of the same size are cut with a preset overlap from the remaining signal. The Mahalanobis distance between the AR model coefficients of left-out segment and those of the sample set is then computed and stored. This process is repeated for a large number of times and the value beyond which 5% of the tests occur is used as threshold in subsequent analysis. This is essentially an estimation of the feature distribution by recomputing the statistic for many a time by leaving out a certain portion of observation, and can be viewed as a combination of jackknife and cross-validation technique [Shao and Tu, 1995].

Other methods have been proposed for threshold determination of Mahalabnobis distance feature; a Monte Carlo method has been used [Sohn, 2001] to produce the desired threshold by calculating the 5% quantile of empirical feature distribution of simulated coefficient vectors whose components are drawn independently from standard normal distribution. However, as noted in previous text, the AR coefficients are not mutually independent and may not have a unit variance. The assumption of this simulation is not well-grounded, thereby this approach is not used here though reported successful in a couple of literatures.

Empirical density estimation for multiple features; the nearest neighbor rule

Threshold determination becomes complicated when there is more than one feature in the algorithm. Fortunately, the nearest neighbor rule (Fig. 3) provides an intuitive way for empirical density estimation, from which damage threshold can be derived.

Suppose a dataset $\{x_i\}$ is generated from a certain probability distribution $p(x)$, then an empirical estimate of the probability density at x_i can be obtained exploiting the condition below

$$\hat{p}(x_i) \propto 1/\min d^n(x_i, x_j) \ (i \neq j).$$

Here n stands for the feature vector dimension, and (x_i, x_j) is the distance between point x_i and x_j. It is clear that the larger the distance from a point to its nearest neighbor, the smaller is its probability to occur.

A decision strategy for multiple features can hence be established as follows; for each feature vector from unknown state, search for its nearest neighbor in the baseline feature set and record the squared Euclidean distance. When the number of feature vectors above threshold exceeds a certain amount, the system is identified as damaged.

Therefore, the only thing remains to be determined for applying this strategy is the threshold value and its corresponding significance level. In the application presented in Section 5, a 5% significance level threshold is used. The threshold value for the test (unknown state) sets is four times the distance value D^* that has 5% of within-baseline nearest neighbor distances above it. This decision is based on following considerations; assume that both baseline and test features follows a same unimodal distribution that decays exponentially with respect to the distance to feature mean, then around 2.5% of the total features will be expected to have squared nearest neighbor distance beyond $2D^*$ given that the baseline and test set are of equal size. Since in the proposed decision strategy only baseline features are used for nearest neighbor search, this will increase the threshold to $4D^*$ because only half of the samples remain. To simply data processing, here the threshold is tested against 5% of the test features instead of 2.5% of the total features. It is but a slight relaxation

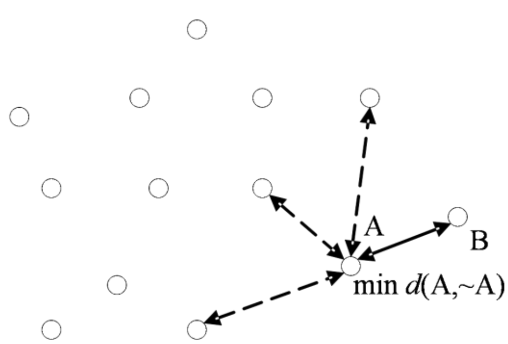

FIG. 3 ILLUSTRATION OF THE NEAREST NEIGHBOR RULE

of conditions, as few of the baseline features will go beyond this critical value.

APPLICATION OF THE RESAMPLING THRESHOLD CONSTRUCTION SCHEME

The damage detection algorithms are applied to acceleration measurements collected from a 5 degree-of-freedom structure (Fig 4) subjected to base excitation. Wired accelerometers are mounted to the shaking table and each floor. Damage is simulated by adding weight to the 4th floor. For each structural scenario, two sets of acceleration signals are recorded from two random excitation experiments. The sampling frequency is 100 Hz for all datasets.

FIG. 4 THE PLEXIGLASS 5 DOF LAB SPECIMEN

Fig 5 displays some of the results from Mahalanobis distance method. (a) is the result from 2nd floor response using threshold from proposed technique; (b) is the result from 4th floor response using threshold from proposed technique; (c) is the result from 4th floor response using threshold from hypothesis testing based on multivariate normal assumption. The significance level in all cases is 5%.

The AR model order adopted is 5, and 350-point segments are cut from acceleration signals with 300 overlap. It is obvious from Fig. 5 (a) and (b) that the change in feature value becomes more prominent at sensing locations more close to structural damage. Also, a comparison between Fig. 5 (b) and (c) shows that the threshold determination technique proposed here does yield a superior performance to that of frequentist hypothesis testing. In all the plots in Fig. 5 logarithmic scale is used for the y axis in order to decrease the oscillation of feature values.

Because in this case a controllable artificial excitation source is used in all the experiments, the overall statistical characteristic of Mahalanobis distance feature does not vary much over time. When this method is applied for damage detection from ambient vibration responses, this feature can become even more unstable. However, this resampling-based technique is still proved reliable in a couple of experiments on a space truss under ambient load, on the condition that there is enough data available.

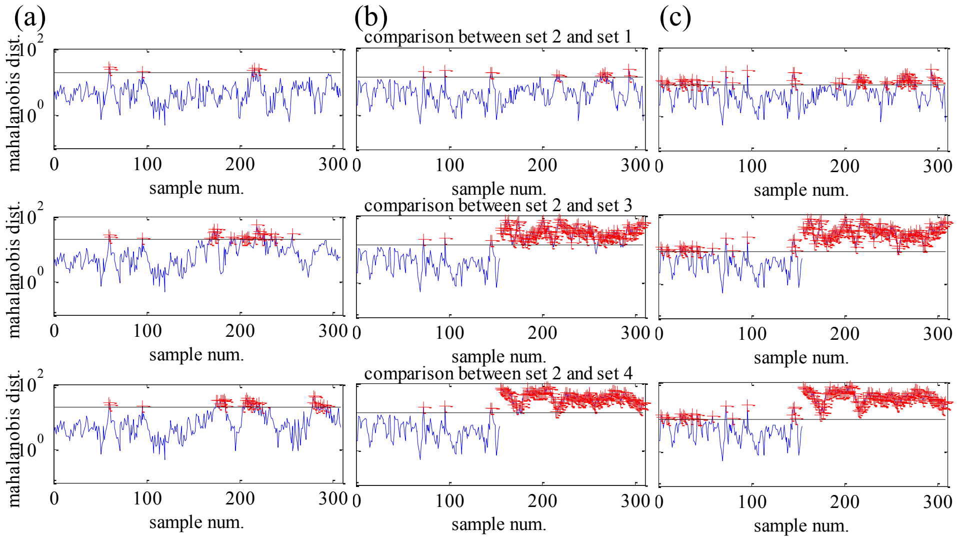

FIG. 5 DAMAGE DETECTION USING THE MAHALANOBIS DISTANCE OF AR MODEL COEFFICIENTS.

FEATURE COMBINATION

Most of the current research on damage detection by time series analysis employs only a single feature, or a set of features from either residual analysis or model parameter analysis. More damage sensitive methods can probably be established by taking features from both categories into consideration, as they in some sense reflect different aspects of the vibration signals. Thus here the two features mentioned in section 2 will be combined for damage identification.

To begin with, two feature pair sets of equal size are extracted from baseline and unknown signal, respectively. The extraction of Mahalanobis distance feature is still the same as before, i.e. for baseline signal feature values are generated using the 'cross-one-out' method introduced in Section 3, and for unknown state signal they are obtained using the routinely procedure described in Section 2. Note here that the logarithm of the distance is used to reduce the oscillation of feature values. Correspondingly, the residual sequence of each signal segment is obtained by filtering each data segment with the AR model estimated from the whole baseline signal. The autocorrelation feature value is then computed from the formula below;

$$Q_2 = \sum_{i=-10}^{10} |\rho(i)|$$

Where $\rho(i)$ is the ACF of the residual sequence. The segment size and overlap length adopted are still 350 and 300.

Fig 6 contains two scatter plots of baseline and test feature clusters extracted from acceleration responses at 4th floor of the lab specimen afore mentioned. Test features in the left plot come from healthy state, while those in the right one are from damaged state. It can be seen that when the structural condition remains same, the two clusters largely overlap each other; when damage has occurred, the test cluster will drift away from the baseline.

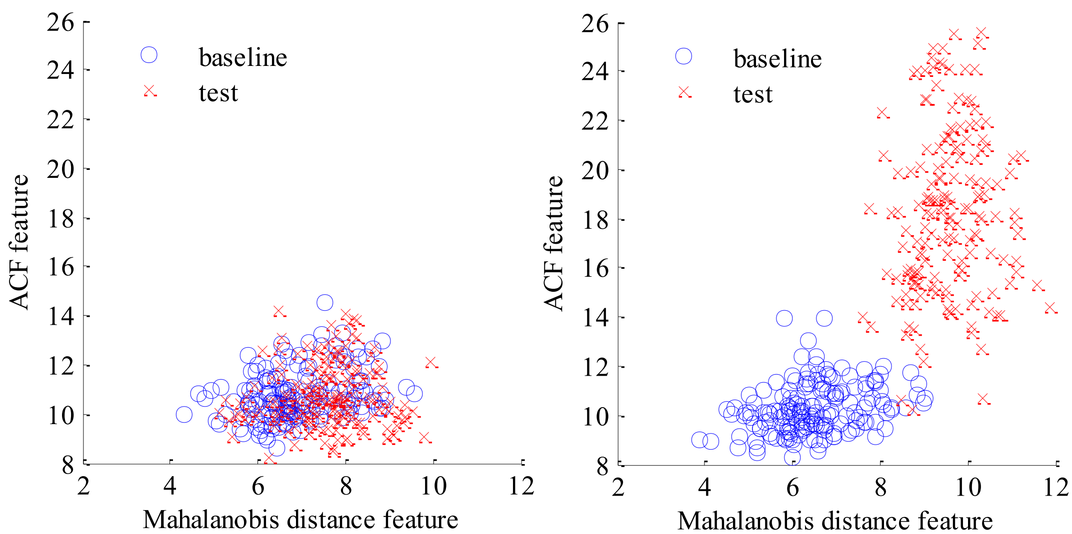

FIG.6 SCATTER PLOTS OF DAMAGE FEATURE PAIRS.

The nearest neighbor scheme is applied here to determine the damage threshold, and the results are summarized in table 1. The value before the slash is the number of outliers, and after the slash is the decision on structural state. The notation in statistical hypothesis testing is used; H_0 denotes healthy state, while H_1 denotes damaged state. Feature variables are all normalized beforehand with respect to the corresponding baseline standard deviation so that contribution of the two features to damage identification is equal. In each comparison the size of both baseline and test set is 166, and the threshold value for the number of outliers is accordingly set at 9. For each sensor location different AR model order is used to better fit the data. The performance of this algorithm is excellent in this case, with only one false alarm for the healthy state and perfect recognition for damaged case. It seems that the responses at 1st and 4th floor are most affected by the simulated damage.

TABLE 1 DAMAGE IDENTIFICATION RESULT USING NEAREST NEIGHBOR SCHEME.

Sensor location / Dataset num.	` 1^{st} floor	2^{nd} floor	3^{rd} floor	4^{th} floor	5^{th} floor
1 (healthy state; baseline)					
2 (healthy state; validation)	$7/H_0$	$7/H_0$	$7/H_0$	$3/H_0$	$14/H_1$
3 (damaged state)	$126/H_1$	$78/H_1$	$32/H_1$	$161/H_1$	$88/H_1$
4 (damaged state)	$112/H_1$	$37/H_1$	$27/H_1$	$159/H_1$	$30/H_1$

CONCLUSION

For this case studies presented in this paper, thresholds based on data driven techniques are proved successful. They tend to generate fewer false alarms than frequentist hypothesis testing, yet still correctly report damage for most tests. All of these techniques employ some sort of empirical density estimation. Although the procedure is somewhat computationally complex, it provides a relatively reliable way to construct thresholds for features with large variation or of unknown distribution and thus automate the process of damage detection.

However, it must be stated that neither improved schemes for threshold construction, nor combination of features, can replace the quest for features that are more stable and damage-sensitive. Features are always the most important topic in damage detection. Adjusting threshold determination methods and combining features can help well-chosen features to perform better, but they cannot save a bad feature from yielding bad performance. Such features based on autoregression on single node response are innately ‘fragile’ because the information is limited and can be affected by other factors. When the loading condition or operation environment of the structure monitored is subjected to change, features based on longer data or data from several sensing locations will probably be more reliable. This will be a good subject for future research.

ACKNOWLEDGMENTS

The research described in this paper is partially supported by the National Science Foundation through Grant No. CMMI-0926898 by Sensors and Sensing Systems program, and by a grant from the Commonwealth of Pennsylvania, Department of Community and Economic Development, through the Pennsylvania Infrastructure Technology Alliance (PITA).

REFERENCES:

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

[2] Brockwell, P. J., Davis, R.A. (2002). *Introduction to Time Series and Forcasting*, 2nd ed. Springer, NY

[3] Brockwell, P. J., Davis, R.A. (2009). *Time Series: Theory and Methods*, 2nd Edition, Springer, NY.

[4] Doebling, S. W., Farrar, C.R., and Prime, M. (1998). "A Summary Review of Vibration-Based Damage Identification Methods." *The Shock and Vibration Digest*, Vol. 30, No. 2, pp.91-105.

[5] Duda, R. O., Hart, P. E., and Stork, D. G. (2000). *Pattern recognition*, Wiley-Interscience, 2nd Edition.

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

[7] Fassois, S. D. and Sakellariou, (2007). J. S, "Time-series methods for fault detection and identification in vibrating structures." *Phil. Trans. R. Soc. A*, Vol. 365, pp. 411-448.

[8] Good, P. I. (1999). *Resampling Methods: A Practical Guide to Data Analysis*, Birkhauser.

[9] Gul, M. and Catbas, F.N. (2009). "Statistical Pattern Recognition for Structural Health Monitoring Using Time Series Modeling: Theory and Experimental Verifications." *Mechanical Systems and Signal Processing*, Vol. 23, Issue 7, pp. 2192-2204.

[10] Koch, K., (1999). Parameter Estimation and Hypothesis Testing in Linear Models, Springer; 2nd Edition.

[11] Mahalanobis, P.C. (1936). "On The Generalized Distance in Statistics." *Proceedings National Institute of Science*, India, Vol. 2, No. 1., pp. 49-55.

[12] Shao, J. and Tu, D. (1995). *The Jackknife and Bootstrap*, New York: Springer-Verlag.

[13] Sohn, H., Farrar, C. R. (2001). "Damage Diagnosis Using Time Series Analysis of Vibration Signals." *Smart Materials and Structures*, Vol.10, No 3, pp. 446–452.

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

[15] Worden, K. and Manson, G., (2007). "the Application of Machine Learning to Structural Health Monitoring." *Phil. Trans. R. Soc*, A 365, pp. 515–537.