DATA MINING AND VISUALISATION FOR ANOMALY DETECTION AND DIAGNOSIS IN CIVIL STRUCTURES

CONFERENCE PAPER · JANUARY 2005

DOI: 10.13140/RG.2.1.4367.8243

Piotr Omenzetter
University of Aberdeen
139 PUBLICATIONS **395** CITATIONS

Pilate Moyo
University of Cape Town
50 PUBLICATIONS **403** CITATIONS

DATA MINING AND VISUALISATION FOR ANOMALY DETECTION AND DIAGNOSIS IN CIVIL STRUCTURES

By JMW Brownjohn, P Omenzetter, P Moyo
[1]School of Engineering, University of Plymouth, Drake Circus, Plymouth PL4 8AA, UK

ABSTRACT

Management of data generated by SHM systems is a major issue to be addressed in future developments. Even with data compression and embedded systems to convert large quantities of data to more manageable amounts of information, there remains the need for procedures to manage the data and in particular to present it to various levels of user. Experience with a number of SHM systems has shown the need to condense data, to develop simple interfaces for quick visual inspection, to provide second and third levels of inspection via statistical analysis tools to identify performance anomalies, more sophisticated parametric modelling and data mining techniques to characterise the anomalies and links to validated structural models for diagnosis. The paper presents experiences with a combination of dynamics-based structural assessment and continuous remote monitoring of static and dynamic effects and response and some of the tools that have been developed to manage and interpret the data.

INTRODUCTION

Data management is an major issue for structural health and performance monitoring systems. The landmark WASHMS system implemented for the Lantau Fixed Crossing [1] employs over 800 sensors and acquires over 2GBytes of data per hour, 120MB deriving from conventional sensors, the lion's share from video recordings and GPS data. The monitoring system installed at Republic Plaza [2] acquires 920kB of data from 16 analog channels and 2MB from a dual-rover GPS system per hour. These data derive from dynamic and static performance of the structures, accounting for both the loading and the system itself, hence they contain information that can be used to characterize the structural system and the input. Structural system identification from such data, which is not limited to modal representations, is still a major research area with new procedures being developed to obtain more accurate or representative characterizations.

The challenge of real-time on-line system identification for permanently monitored structures is even greater than for post-processed data and the aims of civil infrastructure SHM may shift towards motives other than damage detection. In fact continuous structural performance monitoring may have several motivations:

1 Checking as-built performance of a novel structural system
2 Supporting the move towards performance-based design and calibrating loading codes
3 Monitoring and alerting on limit states of structures under construction or affected by external works
4 Assessing the effect of structural upgrades
5 Assessing post-disaster structural integrity
6 Tracking long term structural performance to identify degradation of materials or structural damage
7 Aiding maintenance strategies for growing stock of aging infrastructure
8 Imposition by legislation, insurance or other contractual requirements

SHM developments are heavily driven by academic research and tend to focus on items 1 and 6 yet future developments and real world applications are more likely to be driven by the other requirements, and SHM system design needs to account for these needs at least as much as 'damage detection'. Damage is just one form of anomaly that may be recognizable from monitored signals, with significant research effort being directed towards 'vibration based damage detection' (VBDD), which usually attempts to recognize and interpret subtle changes in modal properties. These procedures are fallible, and it is still not proven that for civil infrastructure it is possible to detect reliably (let alone locate and diagnose) structural damage using modal parameter changes. A recognized difficulty in this area is the normalization of data to filter out ambient effects and the level of accuracy required in modal measurements.

Because of the known difficulties in traditional VBDD [3], alternative strategies have been developed that include non-modal analysis and data normalization to provide reliable indication that something has changed in the structural system. Such an indication can then be used to trigger (automatically if possible) further detailed investigation, aided by historical performance information provided by the SHM system.

For civil infrastructure a major part of the problem with detecting a change in the system is determining the baseline system. Unlike aerospace or automotive industries, each structure is a prototype with unique performance characteristics in the as-built conditions, and a major part of the SHM program has to be devoted to establishing this baseline. In fact for landmark civil structures where high-visibility SHM systems are installed, the first two motivations in the list dominate, providing a wealth of valuable information about structural performance and unusual loading while providing the opportunity to develop SHM technology.

Two SHM systems are studied in this paper. The first, at Republic Plaza was installed due to motives 1 and 2 and has required development of procedures for identifying the loading and response mechanisms and for scanning data to retrieving interesting events. The second, at Tuas Link in Singapore was originally due to motive 8 but has provided more value in respect of motive 6 where it has been used to develop algorithms for performance anomaly detection.

TALL BUILDING PERFORMANCE MONITORING: VISUALISATION

In 1993 Shimizu Corporation began constructing the superstructure of a 280m building (Figure 1) in Singapore using a structural system likely to be employed in Japan. The structural system comprises a concrete (shear) core for resisting lateral loads and a perimeter ring of concrete filled steel tubes for resisting vertical loads.

Figure 1 Republic Plaza

Shimizu installed a set of static stress and strain gauges in the core, columns and beams in a segment of the building at a lower level and arranged for these to be read manually at intervals during the construction. At the same time, starting with a crude vibration recorder, natural frequencies were tracked as construction progressed, and over a decade this evolved into a permanent monitoring installation including biaxial accelerometers at basement and roof, UVW anemometers to capture strong wind characteristics and a GPS system to identify absolute deflections.

The effectiveness of the structural system was proven via a combination of static monitoring, ambient vibration survey and finite element model updating [4] and the monitoring program now has the following objectives:

1. Capture records of ground and building motion due to earthquakes originating outside Singapore
2. Identify the characteristics of the different types of wind and the static and dynamic response of the building
3. Identify any changes in the structural system over the long term of ten years of monitoring

In fact the deep structural knowledge gained over the years and the mechanisms and correlations between load and response has allowed the building to be used as a super-sensor for wind and seismic loads.

Dynamic response recording system

The monitoring system comprises a master analog recording system continuously collecting data from 16 analog signal channels at 64Hz. Four acceleration, six wind speed, two temperature and four RTK GPS (displacement) signals are recorded into a 16k data buffer. These 16-second records are decimated 8-fold using a steep low-pass filter and concatenated into an array holding 512 seconds of 8Hz response data. For each 512-second record, statistics of mean and variance, as well as (for acceleration data) narrow and modal RMS are calculated and saved in a statistics file.

An event triggering strategy has been developed so that if modal dynamic response levels briefly but reliably exceed background thresholds, the system saves the 512-second event. Hence instead of saving 1MB of dynamic data per hour, 1MB of statistics file represents 18 days or monitoring, and only 2-4% of raw time series are saved. Fig. 2 shows how the system identifies and save strong wind and seismic events.

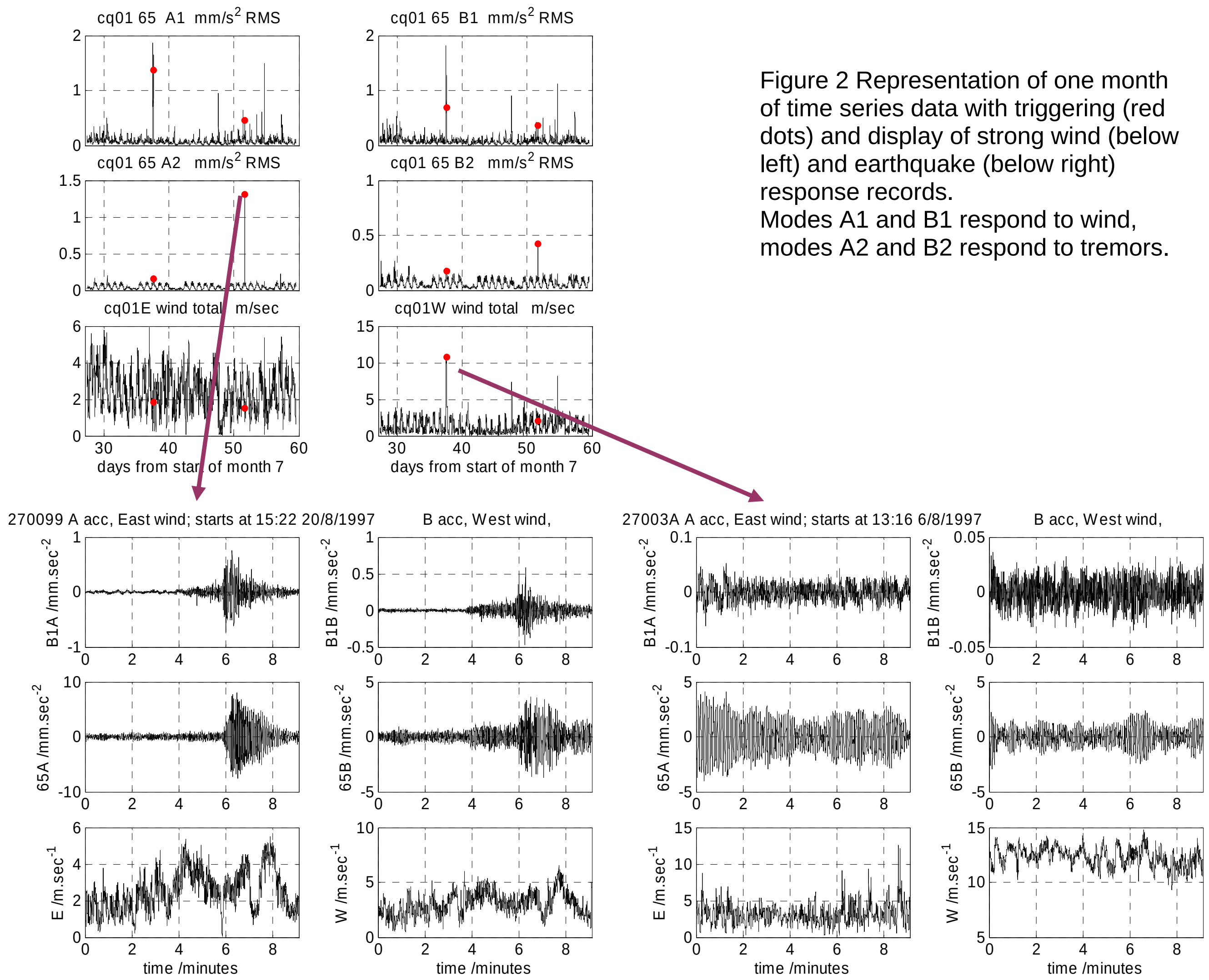
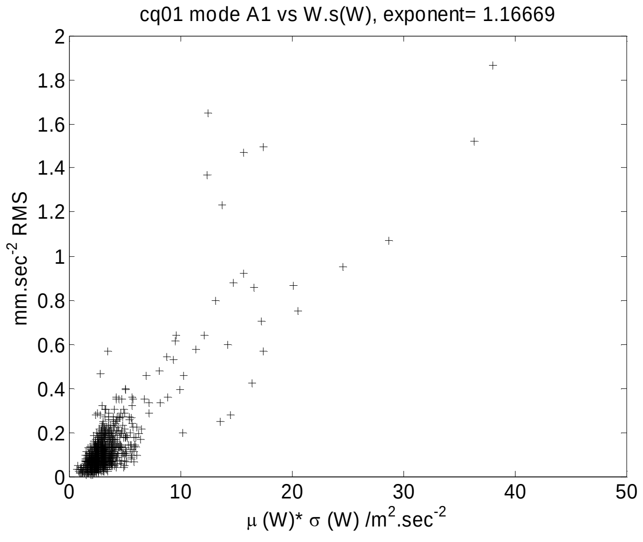


Figure 2 Representation of one month of time series data with triggering (red dots) and display of strong wind (below left) and earthquake (below right) response records.
Modes A1 and B1 respond to wind, modes A2 and B2 respond to tremors.



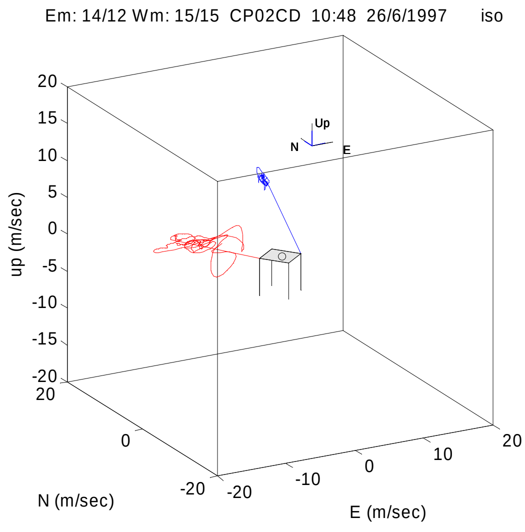


Figure 3 Mode 1 RMS response vs wind strength

Figure 4 Visualising wind vectors for storm

The effect of wind on dynamic response is identified by correlation of modal strength and the product of wind speed and its standard deviation (Fig. 3). Having established this type of correlation allows for dynamic response of the building to be used as an indicator of wind strength. This inverse method is necessary since, as shown in Fig. 4 the wind signals that derive from anemometers on the corner of the building are affected by the building itself. Wind that is forced to blow over the building appears to have higher turbulence as well as an upwards trajectory.

While modal parameters may not be reliable indicators of damage, they may be dependent on modal amplitude and ambient conditions. Attempting to recover modal estimates from output-only data is tedious even for individual short records with significant reseach in improved techniques[5]. For Republic Plaza data, an automated procedure has been applied to samples of continuously record low frequency response. The eigensystem realisation algorithm is used, dividing the 17-day record into 10 minute frames from which modal paramters are estimated. The bar heights and colour show the mode strength, and it is clear that the mode frequency has diurnal variation, as does the total response RMS (the blue line).

The procedure can, with some degree of uncertainty, also visualise variations in damping ratio, known to be a parameter sensitive to structural degradation [6].

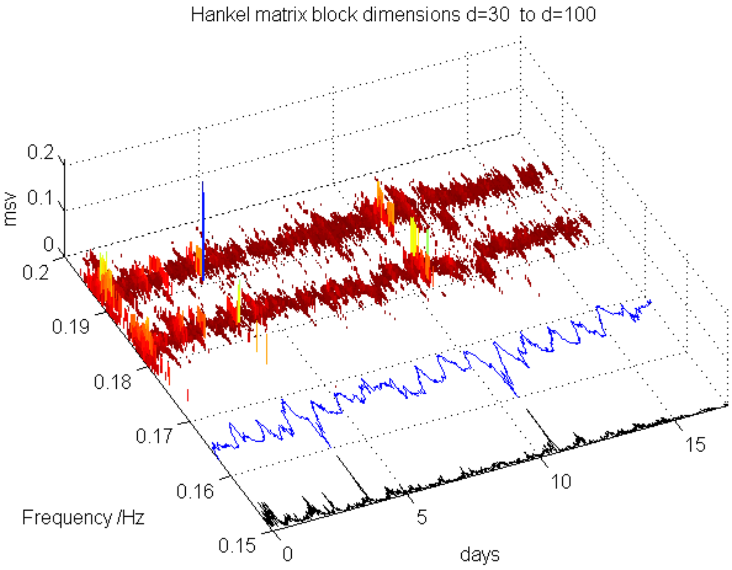


Figure 5 ‘ERAgram’ showing building biaxial fundamental mode frequency and amplitude variation over 17 days.

TUAS SECOND LINK: ANOMALY DETECTION

The recent trend in SHM is to compress data to information at the sensor, i.e. to develop ‘embedded systems’. In a sense this is being done at Republic Plaza but in the acquisition system rather than at the individual sensor. The condensed form of data, possibly called information, would characterize the signal in a more advanced way than developed for Republic Plaza, and certainly it should be possible to transmit reliable estimates of modal parameters, or other characterizations of the signal (e.g. polynomials, auto-regressive coefficients). The sample rate for these parameters would depend on the size of block or frame used to recover them (for Republic Plaza it is 512 seconds) so the data reduce to slowly sampled time series. This is just another form of data and the task remains to investigate by normalising, looking for patterns, correlations and more important, changes in patterns and correlations. The authors believe that this will be the major growth area in SHM in the coming years as more efficient and cost-effective sensors are increasingly used with faster communication rates to deliver ever increasing quantities of data.

There are many tools that can be applied in this area for reduction of data and detecting anomalies. For example autonomous neural networks can be used to detect changes in patterns established by a training phase, while discrete wavelet transforms can be used to filter the signals into frequency bands and generate time-frequency information for signals as well as correlations between signals [7]. Principal component analysis (PCA) can be used to compress multi-channel data into smaller numbers of channels and filter uncorrelated information into separate time series, and techniques of statistical process control (SPC) [8] can be used to identifier outliers from established patterns. Analogous to extracting modal parameters form dynamic response, various forms of auto-regressive models for moving average, exogenous inputs and varying coefficients [9] can be fitted to single or multi-channel data. These techniques are used by themselves or in series (for example applying SPC and time series modeling to wavelet coefficients) to identify outliers (one-off deviations) or system changes from time-sampled data. These data may be either static response parameters such as stress, strain and temperature or data derived from dynamic response, such as frequency and mode shape estimates.

Hence there is a formidable array of tools to reach the first level of SHM, which is to identify the existence of a 'defect'. Since the authors have been dealing with newly built high profile structures, the terms 'defect' and 'damage' are inappropriate; 'structural performance anomaly' is a more honest term that implies an altered structural state or a change in the loading, either temporary or permanent. This leads to a higher level of SHM which is an attempt to characterize the anomaly. In the terms of traditional VBDD the next level would be location and quantification of damage. The authors have used a different approach with static time series data from a bridge monitoring program in Singapore and have used, characterizing performance events during construction and attempting to identify analogs during continuous operation.

Fig. 6 shows the bridge under construction and Fig. 7 a sample of the strain data from gauges embedded in top and bottom corners of the box. A sequence of construction events such as shifting formwork, casting, post-tensioning and span closure (continuity stitch) are highlit as step rapid, slow permanent or transient changes in the data. Some of the more subtle changes are not visible to the naked eye, hence wavelets are used to enhance these and are then processed for significance.

Figure 6 (above) Tuas Second Link: monitored span under construction

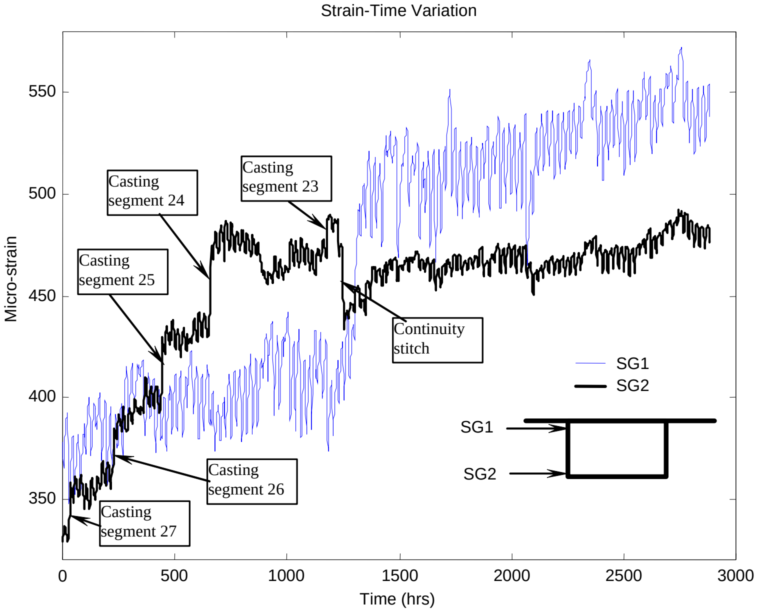


Figure 7 (right) strain gauge readings in box deck segment (#31) close to pier as span construction progresses and extends the cantilever.

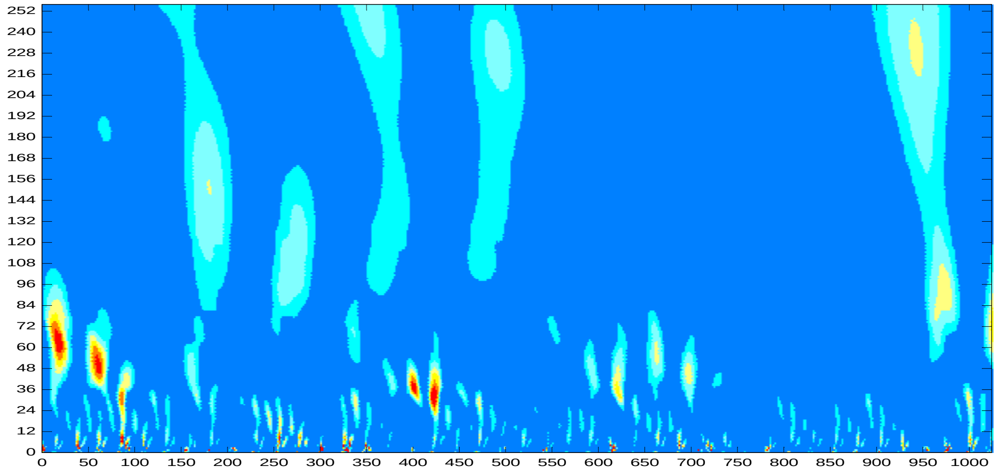


Figure 8 (right) Wavelet cross spectrum of strains at top corners of box segment. Power in scales 6-36 is related to daily temperature cycles.
Power in scales 36-84 represent short-lived changes, such as concreting.
Bands of high power above scale 84 depict ends of segment construction.

Wavelets provide a powerful and visual means to identify events buried in signals that show features related to different sources. When the structure is in service, events that resemble construction events may occur, for example loss of tension in a cable or pier settlement. Such events may be gradual changes, in which case changes in the system model may be detected [9], while more sudden events could be distinguished and categorized by modeling wavelet coefficients as ARMA time series and identifying outliers.

The operation of such a procedure [10] is illustrated in Fig. 9 for data from all four strain gauges in a segment. Outliers are detected as being significant according to Mahalanobis distance, a statistical measure of deviation from the best fit model which differs from the Euclidian distance in that it accounts for the relative dispersions and correlations among vector elements.

Given an outlier event has been detected, the contributions of the sensors to the even can be determined and intervention analysis undertaken on the original strain data [11] to categorise the event as a permanent or transient shift in strain.

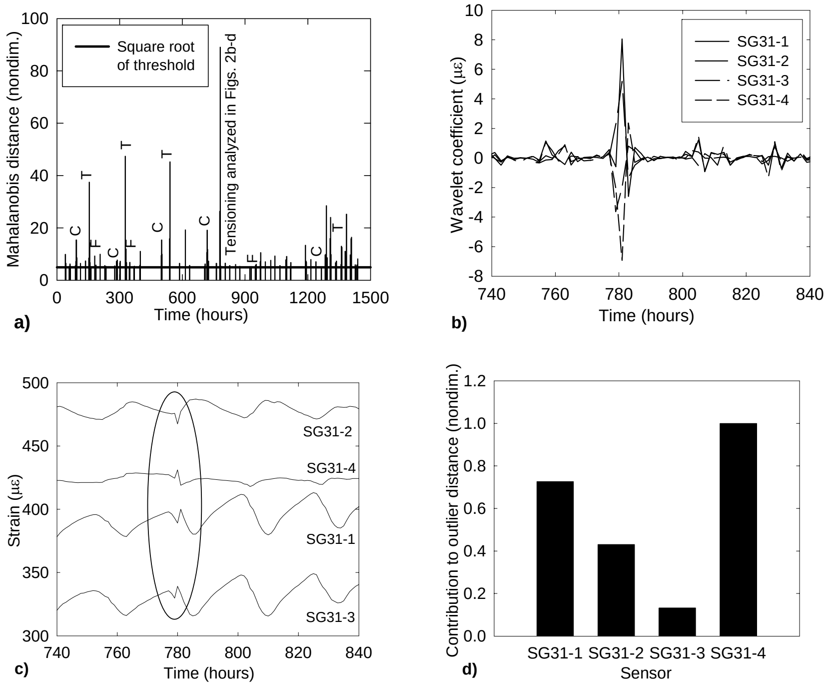


Figure 9. Monitoring during construction: a) Mahalanobis distances for identified events, b) wavelet coefficients for analyzed tensioning event, c) strains for analyzed tensioning event, d) sensors' contribution to outlier distance for analyzed tensioning event.

REFERENCES

1 Wong KY, Instrumentation and health monitoring of cable-supported bridges. Structural Control and Health Monitoring 11(2), 91-124, 2003.

2 Brownjohn JMW, Long-term performance monitoring of a tall building under wind and seismic excitation The Second International Conference on Structural Engineering, Mechanics and Computation, Cape Town, Balkema, p659-664, 5-7 July 2004.

3 Doebling SW, Farrar, Prime M, Schevitz DW, Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a review. Los Alamos National Labs Report LA-13070-MS UC-900, 1996

4 Brownjohn JMW, Pan TC, Deng X, Correlating dynamic characteristics from field measurements and numerical analysis of a high rise building, Earthquake Engineering and Structural Dynamics **29**, p523-543, 2000.

5 Peeters B, Vecchio A, Van der Auweraer H, PolyMAX modal parameter estimation from operational data. ISMA 2004 International Conference on Vibration Engineering, Leuven Belgium, September 2004.

6 Jeary AP, Chiu GC, Wong JCK, Wholistic structural appraisal. Proc. 8th Int. conf. and structural safety and reliability: ICOSSAR 2001 .

7 Sohn H, Farrar CR, Hunter N, Worden K, Applying the LANL statistical pattern recognition paradigm for structural health monitoring to data from a surface-effect fast patrol boat. Los Alamos National Report LA-13761-MS, January 2001.

8 Moyo P, Brownjohn JMW, Detection of anomalous structural behaviour using wavelet analysis. Mechanical Systems and Signal Processing **16**(2-3), p429-445, 2002.

9 Fugate ML, Sohn H, Farrar CR, Vibration-based damage detection using statistical process control. Mechanical systems and Signal Processing **15**(4) p701-721 2001.

10 Omenzetter P, Brownjohn JMW, Moyo P, Tracing the behavior of a Bridge Using Time Series Modeling. 4th International workshop on SHM, Palo Alto, p491-498. September 2003

11 P Omenzetter P, Brownjohn JMW, Moyo P, Identification of unusual events in multi-channel bridge monitoring data using wavelet transform and outlier analysis. Proceedings SPIE **5057**, p157-168, 2003