Benchmark Data for Structural Health Monitoring

Jyrki Kullaa

7th European Workshop on Structural Health Monitoring
July 8-11, 2014. La Cité, Nantes, France

BENCHMARK DATA FOR STRUCTURAL HEALTH MONITORING

Jyrki Kullaa

*Helsinki Metropolia University of Applied Sciences,
P.O. Box 4021, 00079 Metropolia, Finland*

jyrki.kullaa@metropolia.fi

ABSTRACT

Data analysis is a key function in structural health monitoring (SHM). To develop algorithms for SHM, one needs realistic data. A library of SHM data is introduced with simulations of vibration measurements resulting in three challenging SHM cases: (1) a beam structure with environmental and operational influences, (2) a beam with a non-linear breathing crack, and (3) moving loads on a beam, modelling traffic on a bridge. In all these cases, the excitation is varying and unknown. Also, the environmental or operational variables are random and unknown. A different modelling strategy is used in each case. Also the damage scenarios are different. Data analysis examples for damage detection and localization are presented utilizing redundancy of the sensor network. There are endless possibilities to utilize the benchmark data.

KEYWORDS : *structural health monitoring, damage detection, environmental or operational effects, moving load, breathing crack.*

INTRODUCTION

Data analysis is a key function in structural health monitoring (SHM). The objective is usually damage detection and localization. A statistical deviation of the dynamic behaviour of the structure is an indication of damage. Training data are first acquired from the undamaged structure under normal environmental or operational conditions. Once damage has been detected, its location can be estimated relative to the sensor positions in the sensor network.

In vibration-based SHM, damage identification is performed from vibration signals measured simultaneously at different locations of the structure. Damage detection can be performed in the time domain from the raw sensor data or in the feature domain, in which damage-sensitive features are first extracted from the time series.

To develop algorithms for SHM, one needs realistic data. In many cases, data analysts do not have resources to acquire data or perform realistic simulations. Therefore, it would be useful to have a library of SHM data available for testing and validating new algorithms.

Obviously, the most realistic data would be vibration measurements from a full-scale structure. The data would origin from a large sensor network with different types of sensors (e.g. accelerometers and strain gauges) and include measurement noise, occasional outliers, normal environmental or operational variability, unknown ambient excitation (with varying amplitude, frequency content, or spatial distribution), transient or steady state vibrations, missing samples of data, damage, sensor faults, and non-linearities. From the aforementioned list it can be concluded that most simulations may not be very realistic, because great effort is needed to build a model. Moreover, it would be very tempting to generate data to fit one's own algorithms. Therefore, an independent source of SHM data would be valuable to assess the capabilities and restrictions of different algorithms.

SHM data can be divided into two categories, features and time series, which can be univariate or multivariate. A univariate feature represents a vibration characteristic extracted from the measurement data, for example the first natural frequency. It can be a Gaussian or a non-Gaussian variable, and damage can be either a change in the mean or change in the variance. A multivariate feature represents a vector of vibration characteristics, for example the lowest natural frequencies and mode shapes. The number of features may be very high resulting in the curse of dimensionality. Also environmental or operational variability can have an effect on the features.

Direct simulation of features is often relatively easy for a data analyst. Simulation of time series from a vibrating structure is more demanding, but also results in more realistic data. Finite element method (FEM) is often used for simulation. In this study, a few benchmark cases are designed for SHM with different challenges: (1) A beam structure with environmental and operational changes, (2) A beam with a non-linear breathing crack, and (3) moving loads on a beam, modelling traffic on a bridge. In all these cases, the excitation is varying and unknown. Also, the environmental or operational variables are random and unknown. Each case introduces a different modelling approach and a different damage scenario.

The paper is organized as follows. A damage detection algorithm is described, which utilizes the redundancy of the sensor network and the generalized likelihood ratio test (GLRT) in the time domain. Three different benchmark cases are introduced followed by examples of data analyses for damage detection and localization. A short conclusion is given in the end.

1 DAMAGE DETECTION AND LOCALIZATION

The sensor network is assumed to form a redundant system; for spatial correlation, the number of sensors should be greater than the number of active modes. Let $\mathbf{x}$ be the measured variable, typically a simultaneously acquired sample of accelerations or strains. Working in the time domain avoids the curse of dimensionality, because the number of samples is often much larger than the data dimensionality.

To identify a model for a single sensor in the network, the minimum mean square error (MMSE) estimation is used [1], in which the signal of each sensor is estimated in turn using the other sensors in the network. The sensors are divided into observed sensors $\mathbf{v}$ and missing sensors $\mathbf{u}$:

$$\mathbf{x} = \begin{Bmatrix} \mathbf{u} \\ \mathbf{v} \end{Bmatrix} \tag{1}$$

with a partitioned data covariance matrix $\mathbf{\Sigma}$ of the training data

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{uu} & \mathbf{\Sigma}_{uv} \\ \mathbf{\Sigma}_{vu} & \mathbf{\Sigma}_{vv} \end{bmatrix} = \begin{bmatrix} \mathbf{\Gamma}_{uu} & \mathbf{\Gamma}_{uv} \\ \mathbf{\Gamma}_{vu} & \mathbf{\Gamma}_{vv} \end{bmatrix}^{-1} = \mathbf{\Gamma}^{-1} \tag{2}$$

where the precision matrix $\mathbf{\Gamma}$ is defined as the inverse of the covariance matrix $\mathbf{\Sigma}$ and is also written in the partitioned form. A linear MMSE estimate is [2]:

$$\hat{\mathbf{u}} = \boldsymbol{\mu}_u - \mathbf{\Gamma}_{uu}^{-1}\mathbf{\Gamma}_{uv}(\mathbf{v} - \boldsymbol{\mu}_v) \tag{3}$$

where $\boldsymbol{\mu}_u$ and $\boldsymbol{\mu}_v$ are the mean of $\mathbf{u}$ and $\mathbf{v}$, respectively. The error covariance matrix is

$$\mathbf{\Phi} = \mathrm{cov}(\mathbf{u} \mid \mathbf{v}) = \mathbf{\Gamma}_{uu}^{-1} \tag{4}$$

Assuming Gaussian distribution, the conditional probability density function of $\mathbf{u}$ becomes:

$$p(\mathbf{u} \mid \mathbf{v}) = \left|2\pi\mathbf{\Phi}\right|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{u} - \hat{\mathbf{u}})^T \mathbf{\Phi}^{-1}(\mathbf{u} - \hat{\mathbf{u}})\right] \tag{5}$$

Damage detection is done using the hypothesis test for the MMSE model parameters applying the generalized likelihood ratio test (GLRT) [3]. The test statistic is the log-likelihood ratio $l(\mathbf{u} \mid \mathbf{v})$ for each sample:

$$l(\mathbf{u} \mid \mathbf{v}) = \ln \frac{p(\mathbf{u} \mid \mathbf{v}; H_1)}{p(\mathbf{u} \mid \mathbf{v}; H_0)} \tag{6}$$

where $p(\mathbf{u} \mid \mathbf{v}; H_i)$ is the probability according to the hypothesis H_i, $i = 0, 1$. The hypothesis H_0 is that the model parameters are the same as those of the training data (normal), and the hypothesis H_1 is that the parameters are different to those of the training data (anomaly). The distributions $p(\mathbf{u} \mid \mathbf{v}; H_0)$ and $p(\mathbf{u} \mid \mathbf{v}; H_1)$ are obtained by identifying the parameters from the training data and the current measurement, respectively. The hypothesis test is based on the Neyman-Pearson (NP) lemma [3]. To maximize the probability of detection P_D for a given probability of false alarm $P_{FA} = \alpha$, decide H_1 if $\sum l(\mathbf{u} \mid \mathbf{v}) > \gamma$. The threshold γ is found from the false alarm constraint $P_{FA} = \alpha$. Choosing the threshold for detection is discussed in the following.

Each sensor yields one test statistic, resulting in a total number of variables equal to the number of sensors in the network. The dimensionality can be reduced by using principal component analysis (PCA) [4] to the log-likelihood ratios l. The distributions of l and the PCA scores are unknown. Therefore, the statistics used for novelty detection are the maxima and minima of the first principal component scores, and using the theory of extreme value statistics (EVS) [4, 5], the thresholds are designed so that the probability of false alarms is 0.001. The extreme values are computed from 100 subsequent variables. Finally, the statistics are plotted on a control chart [7].

Damage localization is performed by computing the average log-likelihood ratio for each sensor in the measurement. The largest value is assumed to reveal the sensor closest to the damage location.

The algorithm described is used in the following three benchmark studies.

2 A BEAM STRUCTURE WITH ENVIRONMENTAL AND OPERATIONAL CHANGES

This is a modified case of that presented in [8]. Consider a simply-supported beam with length 1.4 m, a uniform rectangular cross-section of 50 mm × 5 mm (Figure 1). The beam is also supported with a spring 612.5 mm from the support, with the spring constant k depending non-linearly on temperature:

$$k = k_0 + aT^3 \tag{7}$$

where k_0 = 100 kN/m, a = –0.8 (with compatible units), and T is temperature with a uniform random distribution between –20 and +40°C. Notice that seasonal variation of temperature would probably have been more realistic.

The beam is divided into three sections of equal length. The Young's modulus E_i in the ith section has linear relationship with a corresponding independent and dimensionless environmental variable z_i:

$$E_i = E_0 + \sigma_i z_i, \quad i = 1,2,3 \tag{8}$$

where E_0 = 207 GPa, z_i are standardized Gaussian variables: $z_i \sim N(0,1)$, and the standard deviations σ_i of different sections are: σ_1 = 5 GPa, σ_2 = 3 GPa, and σ_3 = 7 GPa.

The structure is modelled with 144 simple beam elements and a single spring element. Independent random excitations with different amplitudes in each measurement excite the structure at three points (Figure 1). Modal superposition with a static correction procedure [9] is used in the response analysis. Transverse acceleration is measured at 47 equidistant points along the beam. Gaussian noise with standard deviation σ = 0.01 m/s^2 is added to each sensor. In the average the noise level is approximately 1% of the signal. The sampling frequency is 571 Hz and the number of samples is 2859 in each measurement. The first 50 measurements are from the undamaged structure. A slow environmental variability is assumed, justifying a constant environment during each measurement.

As a conclusion, the environmental or operational variability thus introduced originate from (1) a variable spring with a non-linear relationship between temperature and the spring constant, (2) three regions with independently varying Young's moduli, and (3) random load distribution at three points.

Damage is a decrease in the beam depth at the spring support in two elements along a total length of 19.4 mm. This type of damage could represent local corrosion around the spring joint. Sensor 21 is located in the middle of the damaged region (Figure 1). The height of the damaged beam varies in five different levels: 4.5, 4, 3.5, 3, and 2.5 mm. Each damage level is monitored with 10 measurements at variable unknown environmental conditions.

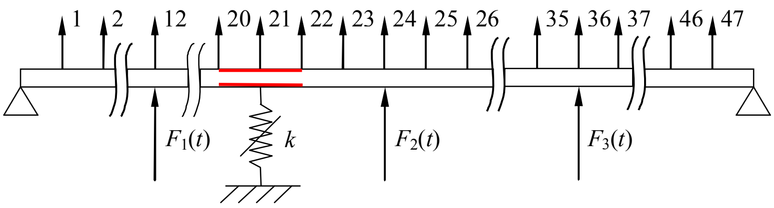

Figure 1: Simply-supported beam with a variable spring, three independent concentrated loads and structural damage. The response is measured at 47 DOF shown with numbered arrows [8].

2.1 Data analysis for damage detection and localization

The training data are the first 25 measurements. These are also used to design the control charts. The first principal component of the generalized likelihood ratio is used for damage detection. The extreme value statistics of subgroup size 100 are plotted on the control chart. The test data include 75 measurements, 25 measurements (26–50) from the undamaged structure and 50 measurements (51–100) from the damaged structure. Figure 2 shows the EVS control chart for damage detection and the likelihood ratio for each sensor for damage localization (right). Damage is well detected without too many false alarms. Damage is also correctly localized to sensor 21.

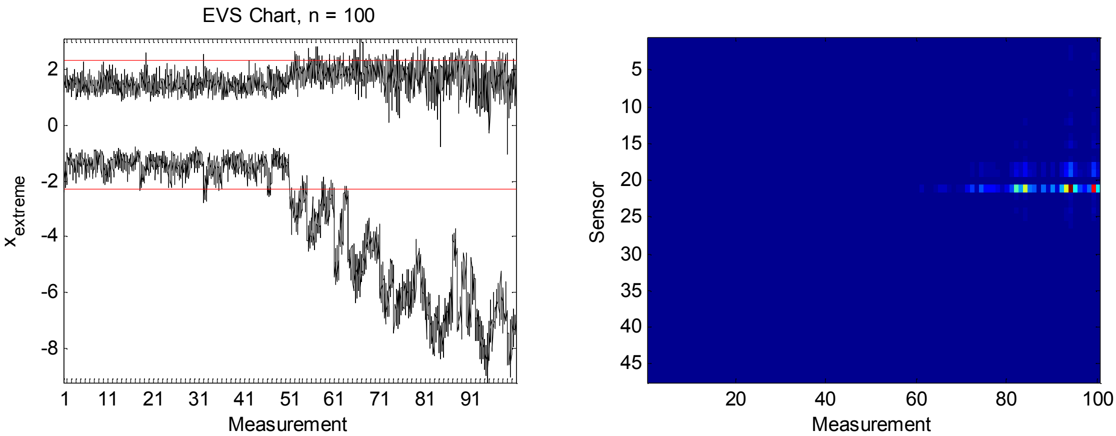

Figure 2. EVS control chart for damage detection (left) and likelihood ratio for damage localization (right).

3 A BEAM WITH A NON-LINEAR BREATHING CRACK [10]

Cracking is a common failure mechanism in machines, vehicles, and structures. Therefore, it is important to detect cracks at an early stage before catastrophic failure. Detection of cracks may be approached using wave-propagation or vibration based methods. Finite element simulations were performed to study the smallest crack length that can be detected using vibration measurements [10].

Crack models can be divided into two main categories: (1) open crack and (2) breathing crack. An open crack is often assumed, in which the crack edges are never in contact. There are two main reasons to use this assumption: (1) the structure behaves linearly, and linear models are easy to use both in simulation and in system identification; and (2) preparing test objects with open cracks is relatively easy. However, a breathing crack is more realistic, in which the crack opens and closes due to external loading. As the crack closes it transmits compressive stress, while in the opening stage, no normal stress is present. This leads to a sudden change in the dynamic properties and a non-linear behaviour, which is more difficult to study both mathematically and experimentally.

3.1 Modelling

The structure is a simply supported beam with the following dimensions: length 5 m, height 0.5 m, and width 0.01 m. It was modelled with 4-node linear 2D elements with reduced integration, and Abaqus Explicit finite element code was used for the simulations. The beam ends were supported on the neutral axis of the beam. Furthermore, the beam end nodes were forced to follow the Euler Bernoulli beam theory by assuming that the planes at the beam ends remain planes. Rayleigh damping [9] was applied, because a full damping matrix was needed. The parameters were chosen to produce low damping in the frequency range of interest. The contact of the crack surfaces was assumed to be frictionless in the tangential direction. In the normal direction the Abaqus surface-to-surface option was used. Figure 3 shows a cracked beam model in a displaced configuration.

A uniform transverse random load history, different in each case, was applied to the top surface of the beam. The load histories were low-pass filtered below 1000 Hz, resulting in five active dynamic modes of the structure. The measurement period was two seconds including 4001 samples.

In the undamaged case, no crack was present. Damage was a single vertical crack at the bottom of the midspan with different crack lengths of 10, 20, 30, 50, 100, and 150 mm.

The total number of sensors was 30. They measured accelerations in the transverse (vertical) direction at the top or bottom edge of the beam. The top sensors are depicted in Figure 3. The bottom sensors were located at the same longitudinal positions on the opposite side of the beam. The bottom sensor in the midspan was located at the crack edge, which may result in too an idealistic case.

Noise was added to the acceleration records obtained from the finite element analyses. The signal-to-noise ratio (SNR) was 30 dB, which is a typical value in vibration measurement systems. The SNR value greatly affected the minimum crack size that can be detected [10].

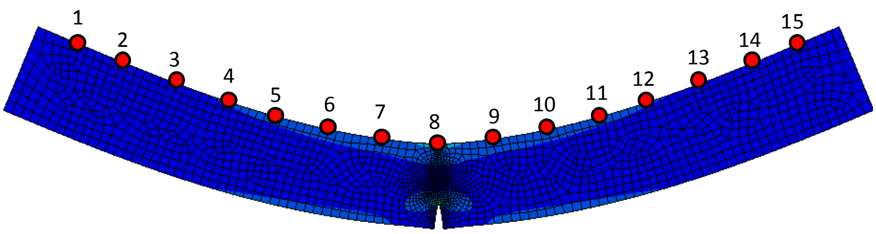

Figure 3. Monitoring of a cracked beam with 15 accelerometers.

3.2 Data analysis for damage detection and localization

The training data for model identification and the in-control data for the control chart design were measurements 1–6. The first principal component of the generalized likelihood ratio was used for damage detection. The extreme value statistics of subgroup size 100 were plotted on the control chart. The test data included six measurements, four measurements (7–10) from the undamaged structure and two measurements (11–12) from the damaged structure.

Figure 4 shows the extreme value statistic (EVS) control chart for damage detection and the log-likelihood ratio for crack localization using the 15 top sensors with crack size of 50 mm. This size of crack (10% of beam height) was successfully detected with the given SNR value. The crack was correctly localized to the closest sensor.

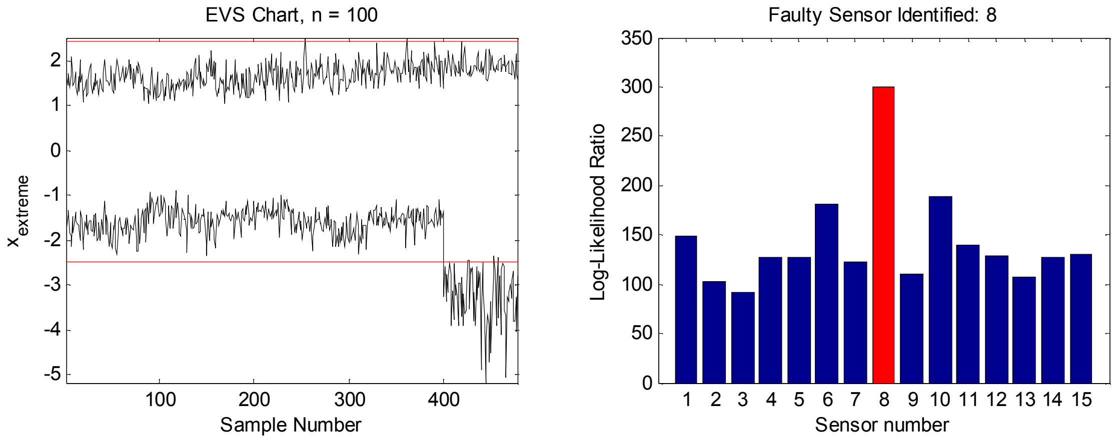

Figure 4. EVS control chart for damage detection (left) and likelihood ratio for damage localization (right).

4 MOVING LOADS ON A BEAM, MODELLING TRAFFIC ON A BRIDGE

This case is a modification of Tedesco [11]. A simply supported uniform bridge girder is subjected to a moving load having constant velocity v (Figure 5). This simulates a vehicle crossing the bridge. The bridge is analyzed as a continuous system. Assuming zero initial conditions, transverse acceleration of the beam $\ddot{y}(x,t)$ is computed. The increased mass or the vehicle dynamics are not taken into account.

The natural angular frequency of mode n is

$$\omega_n = n^2\pi^2\sqrt{\frac{EI}{\rho AL^4}} \tag{9}$$

The displacement response is

$$y(x,t) = \sum_{n=1}^{\infty}\frac{F_0}{M_n\omega_n}\frac{1}{\left(\frac{n\pi v}{L}\right)^2-\omega_n^2}\left(\omega_n\sin\frac{n\pi vt}{L}-\frac{n\pi v}{L}\sin\omega_n t\right)\sin\frac{n\pi x}{L} \tag{10}$$

where $M_n = \frac{1}{2}\rho AL$ is the modal mass of mode n. Equation 10 is valid when the vehicle is on the bridge, or for the time interval $0 \le t \le t_1$, where $t_1 = \frac{L}{v}$. For $t > t_1$, the girder is in free vibration with modal initial conditions $q_n(t_1)$ and $\dot{q}_n(t_1)$, and the displacement response is

$$y(x,t) = \sum_{n=1}^{\infty}\left[\frac{\dot{q}_n(t_1)}{\omega_n}\sin\omega_n(t-t_1)+q_n(t_1)\cos\omega_n(t-t_1)\right]\sin\frac{n\pi x}{L} \tag{11}$$

The acceleration response is derived by integrating (10) and (11) twice with respect to time t. The response consists of a series with infinite number of terms. It is assumed that the first d modes respond dynamically, while the higher modes respond quasi-statically. For the first d modes, Equations (10) and (11) are used, and a static correction term is added to approximate the contribution of the higher modes [9].

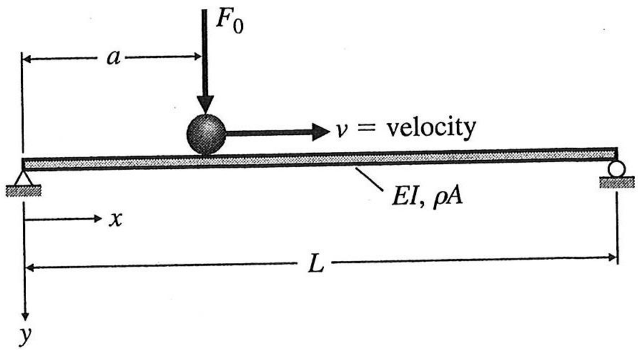

Figure 5. A moving load on a beam modelling a vehicle on a bridge [11].

Several vehicles are allowed to cross the one-way bridge during the measurement period. The vehicle mass, velocity, and the entering time are random variables. Linear superposition is used to compute the bridge response. The vehicle mass varies uniformly between 500 and 20 000 kg, the velocity varies uniformly between 50 and 120 km/h, and the entering time varies according to a Poisson process with a mean of 5 s between events.

The measurement period is 30 s with a sampling frequency of 100 Hz. Accelerations are measured at 98 equidistant points along the bridge. The number of modes in the analysis is 10. The measurement system is assumed to include a low-pass filter that removes the contribution of the higher modes.

The following material and structural parameters are used: span length L = 100 m, flexural rigidity EI = 754110 kNm^2, and mass/length ρA = 262 kg/m. Damping is zero. Damage is deterioration of the support, resulting in an increase of span length by 50, 100, and 200 mm. The first 20 measurements are from the undamaged structure, and two measurements are acquired from each damage level. Gaussian noise with standard deviation σ = 0.1 m/s^2 is added to each sensor.

4.1 Data analysis for damage detection and localization

The training data for model identification and the in-control data for the control chart design are measurements 1–10. The first principal component of the generalized likelihood ratio is used for damage detection. The extreme value statistics of subgroup size 100 are plotted on the control chart. The test data include 16 measurements, 10 measurements (11–20) from the undamaged structure and 6 measurements (21–26) from the damaged structure.

The EVS control chart is plotted in Figure 6 left. Damage is well detected without too many false alarms. Damage is localized to the sensor closest to the deteriorated support (Figure 6 right).

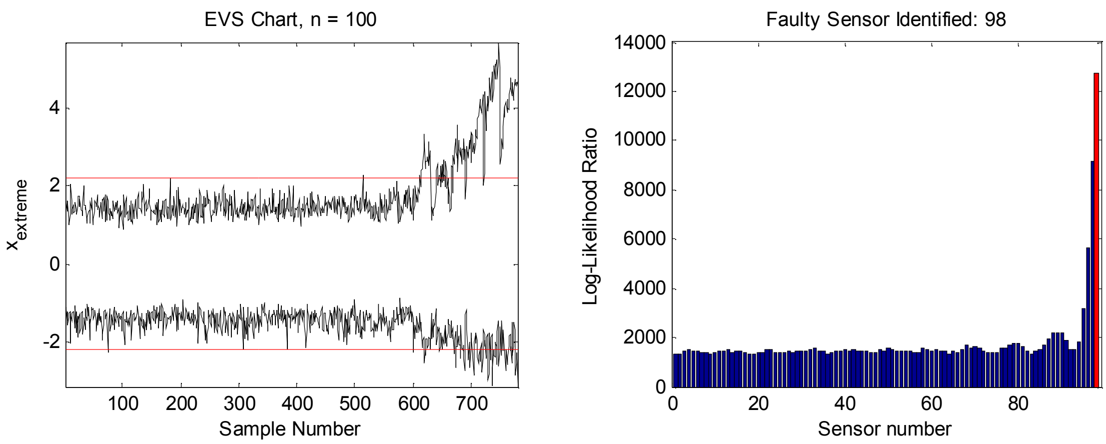

Figure 6. EVS control chart for damage detection (left) and likelihood ratio for damage localization (right).

CONCLUSION

Three benchmark data sets for vibration-based SHM were created with different challenges, including environmental variability, non-linearity (a breathing crack), and moving loads. Also different modelling strategies were used: finite element method with modal superposition, finite element method with explicit time integration, and a distributed (continuous) system. The damage scenario was different in each case.

The benchmark data can be used for different purposes. For example, time series data could be used for system identification before damage detection. Also, with a sensor network, one can select a subset of sensors only or create a sensor fault in one or more channels. More noise can be added to simulate low-cost sensors, or time shifts can be generated to represent inaccuracy in time-synchronization in a wireless sensor network (WSN). Missing data or outliers can be created, a lower sampling rate can be applied, training data can be ignored in a baseline-free analysis, a subset of damage levels can be selected, and so on.

All benchmark data are documented: the model, environmental or operational influences, sensors, noise, damage, and measurements from the undamaged and damaged structure. The challenges of the data analysis for each benchmark are described. Examples of data analyses are presented for sample benchmark data. The data sets can be downloaded at http://users.metropolia.fi/~kullj The database is not complete, and would be supplemented with more data. Suggestions for new data are welcome.

REFERENCES

[1] Sorenson, H.W. 1980. *Parameter estimation. Principles and problems.* New York. Marcel Dekker.

[2] Kullaa, J. 2010. Sensor validation using minimum mean square error estimation, *Mechanical Systems and Signal Processing*, 24, 1444–1457.

[3] Kay, S.M. 1998. *Fundamentals of statistical signal processing. Detection theory*, Upper Saddle River, NJ, Prentice-Hall.

[4] Sharma, S. 1996. *Applied multivariate techniques*. New York, John Wiley & Sons.

[5] Castillo, E. 1988. *Extreme value theory in engineering*, San Diego, CA, Academic Press.

[6] Worden, K., Allen, D., Sohn, H., Farrar, C.R. 2002. Damage detection in mechanical structures using extreme value statistics, in: *SPIE Proceedings*, Vol. 4693, 9th Annual International Symposium on Smart Structures and Materials, San Diego, CA, 289–299.

[7] Montgomery, D.C. 1997. *Introduction to statistical quality control*. 3rd ed. New York. John Wiley & Sons.

[8] Kullaa, J. 2011. Distinguishing between sensor fault, structural damage, and environmental or operational effects in structural health monitoring. *Mechanical Systems and Signal Processing*, 25, 2976–2989.

[9] Clough, R.W., Penzien, J. 1993. *Dynamics of structures*. 2nd edition, New York, McGraw-Hill.

[10] Kullaa, J., Santaoja, K., Eymery, A. 2013. Vibration-based structural health monitoring of a simulated beam with a breathing crack. *Key Engineering Materials* Vols. 569–570, Damage Assessment of Structures X. Trans Tech Publications, Switzerland, 1093–1100.

[11] Tedesco, J.W., MacDougal, W.G. and Ross, C.A. 1999. *Structural Dynamics: Theory and Applications*. Menlo Park, CA., Addison-Wesley.