Detection of structural damage using novelty detection algorithm under variational environmental and operational conditions

M. El Mountassir, S. Yaacoubi and F. Dahmene
Institut de Soudure, R&D department, 4 Bvd Henri Becquerel, 57970 Yutz, France

E-mail: m.elmountassir@institutdesoudure.com

Abstract. Novelty detection is a widely used algorithm in different fields of study due to its capabilities to recognize any kind of abnormalities in a specific process in order to ensure better working in normal conditions. In the context of Structural Health Monitoring (SHM), this method is utilized as damage detection technique because the presence of defects can be considered as abnormal to the structure. Nevertheless, the performance of such a method could be jeopardized if the structure is operating in harsh environmental and operational conditions (EOCs). In this paper, novelty detection statistical technique is used to investigate the detection of damages under various EOCs. Experiments were conducted with different scenarios: damage sizes and shapes. EOCs effects were simulated by adding stochastic noise to the collected experimental data. Different levels of noise were studied to determine the accuracy and the performance of the proposed method.

1. Introduction

SHM of structures using mainly ultrasonic piezoelectric sensors is a promising technique that has been developed in the recent years. The main advantage of this technique over conventional non-destructive evaluation is that the sensors are permanently mounted on the structure to be monitored [1]. Therefore, damage detection is an automatically process which does not need human intervention. It is based on several algorithms and signal processing methods that are capable of recognizing early stage structural damage. Nevertheless, sensors measurements, which should be basically sensitive to the presence of damage, are also sensitive to some environmental and operational conditions (EOCs) [2]. In other words, the effects of these EOCs could be the same as those produced by the damage. This would result in false warnings. Different methods of compensation were developed to overcome this problem, but they have been only designed to perform with an analytical damage detection method, namely subtraction method [3]. On the contrary, statistical methods are less sensitive to the EOCs and can be easily implemented to the monitoring system [4]. They fall into two categories: supervised and unsupervised. When it is a priori possible to retrieve data from damage state, damage identification is done by a supervised learning algorithm such as Artificial Neural Network (ANN) [4]. However, it is generally difficult to provide damaged data information; in this case the damage detection is achieved by an unsupervised learning algorithm based on novelty detection also known as outlier analysis.
This paper is organised as follows. First, a theoritical background wich includes outlier analysis and the Discrete Wavelt Transform (DWT) will be presented. Afterward, the experimental setup that has

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

been used to simulate different damage scenarios, will be discussed in detail as well as the results. Finally, we conclude this paper and indicate the future work.

2. Theoretical background

2.1. Novelty detection

The problem of novelty detection is to identify if the measured signal has deviated from the normal conditions while in-situ monitoring of engineering structures [5]. Obviously, two aspects can be derived from this problem, firstly how can we define the normal conditions and what is the efficient mean of calculating the deviance from these normal conditions. Since, the structure is exposed to different EOCs; the normal conditions should include all factors that can affect somehow the guided wave propagation (e.g. temperature, humidity, load vibration, etc.). In this study, to account for some variation in these EOCs, a digital stochastic noise was added to ultrasonic measured signals in form of signal-to-noise ratio. The second aspect has been inspired from statistics where the novelty detection is generally based on outlier analysis which is a well-studied field that has been recently utilized for damage detection purposes. In the data mining and statistic literature, outliers are also referred to as anomalies, abnormalities, deviants. It follows that an outlier is defined as a data which is significantly different form a baseline dataset [6]. Depending on this baseline data form, the outlier analysis can be calculated through two methods: univariate or multivariate.

2.1.1. Univariate analysis

Given a set of variables representing data which can be considered as an outlier candidate, the univariate analysis explores each variable separately. The discordancy of this outlier can be measured with different tests. The most common test is the deviation statistics, which can be defined as:

$$\mathbf{z} = \frac{|\mathbf{x}-\bar{\mathbf{x}}|}{\boldsymbol{\sigma}} \tag{1}$$

where x is the potential outlier candidate, $\bar{x}$ and σ, the main and the standard deviation of the baseline data respectively. This discordancy value is then compared to a threshold which will be discussed in detail later to determine whether the data is an outlier or not.

2.1.2. Multivariate analysis

In the multivariate analysis, all variables must be combined in a multi-dimensional vector. In this case, the deviation statistics used in the univariate analysis for discordancy test is replaced by the Mahalanobis Square Distance (MSD) given by [7]:

$$\mathbf{D} = (\mathbf{x}-\bar{\mathbf{x}})^{\mathbf{t}}.\mathbf{K}^{-\mathbf{1}}.(\mathbf{x}-\bar{\mathbf{x}}) \tag{2}$$

where x is the potentiel outlier damage index vector, $\bar{x}$, the mean vector of the baseline, K, the covariance matrix of the baseline and t, matrix transpose index. As in the univariate analysis, the MSD value must be compared to a threshold.

2.1.3. Threshold computation

The threshold value is necessary to determine whether a candidate is an outlier or inlier. Its computation depends on the baseline data distribution. When the data are supposed to be non-parametric, the threshold is calculated using the extreme value theorem (based on Monte Carlo simulation) [8]. Otherwise, following a parametric approach, more particularly, if the data is assumed to represent a Gaussian distribution, the threshold value is taken as 99.73% of the Gaussian limit confidence. Therefore, if a candidate is classified as an outlier, there is only 0.27% chance of false classification. Notice that this threshold is a choice, and hence the lower the threshold, the higher the false calls probability.

2.2. Discrete Wavelet transform (DWT)

The DWT is a discrete implementation of the wavelet transform which is a time frequency representation, well designed for the analysis of non-stationary signals (e.g. ultrasonic guided waves) [9]. The DWT decomposes the original signal by computing its correlation with a short-duration wave called the mother wavelet into approximation and detail coefficients as it is shown in Figure 1. Two major applications of this signal processing method can be exploited in this study: denoising and compression [10]. This can be achieved if only a few wavelet coefficients containing the essential information of interest are retained and the remaining coefficients, related to unnecessary information such as noise, are eliminated.

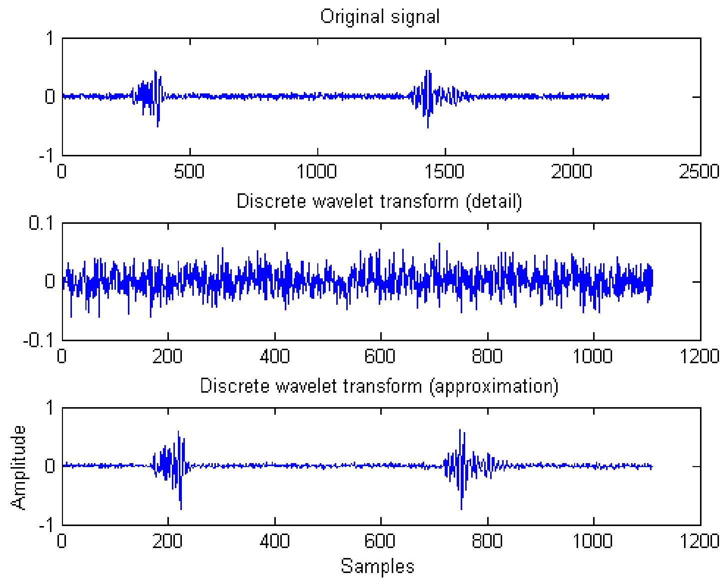

Figure 1. Discrete wavelet transform of an experimental ultrasonic signal (top) showing the detail (medium) and the approximation coefficients (down)

3. Case of study

This section aims to apply the theory of novelty detection to detect and characterise damages severity (if possible) in pipeline structure during in-situ monitoring via ultrasonic guided waves (UGW). Both the univariate and the multivariate techniques will be investigated to determine which one is the most efficient for damage detection.

3.1. Experimental setup

Data acquisition (emission and reception of UGW) was performed using MsS System, designed for pipeline non-destructive testing via ultrasonic guided waves. In order to test the performance of defect detection method, damages were created artificially by adding a magnets cluster on the surface of the structure as it is illustrated in Figure 2.

Figure 2. Artificially simulated damages in a pipeline structure

Journal of Physics: Conference Series **628** (2015) 012126 doi:10.1088/1742-6596/628/1/012126

A single measurement of each damage system states (baseline, 1 defect, 2 defects, 3 defects, 4 defects) was taken. Afterward, a digital random noise was added in order to create the statistical population (database) and to simulate some variation of the signal-to-noise ratio which can be induced by several factors namely, EOCs effect, sensors ageing etc. The noise was created using randn function predefined in MATLAB. This function generates an arrays of random numbers normally distributed with zero mean and a standard deviation equal to 1. It was multiplied by a factor determining the noise level (0.01 low level noise and 0.05 for a high level noise). For each level noise, a total of 200 samples were created for all damages states. Once the database is created (Figure 3), UGW signals must first be gated to remove all redundant information as well as significantly reduce the data size. The critical information of the pipe's condition was stored in the time frame between the excitation of the tone burst pulse and the arrival of the end of the pipe (EOP). The gating technique first detects the excitation pulse and the arrival time of the EOP, which can be easily determined by knowing the signal velocity and the length of the pipe. After gating, the DWT was applied to compress and to denoise the data. The wavelets coefficients were used to calculate the damage-sensitive features. RMS, variance, peak to peak and maximum amplitude were chosen as damage index (DI).

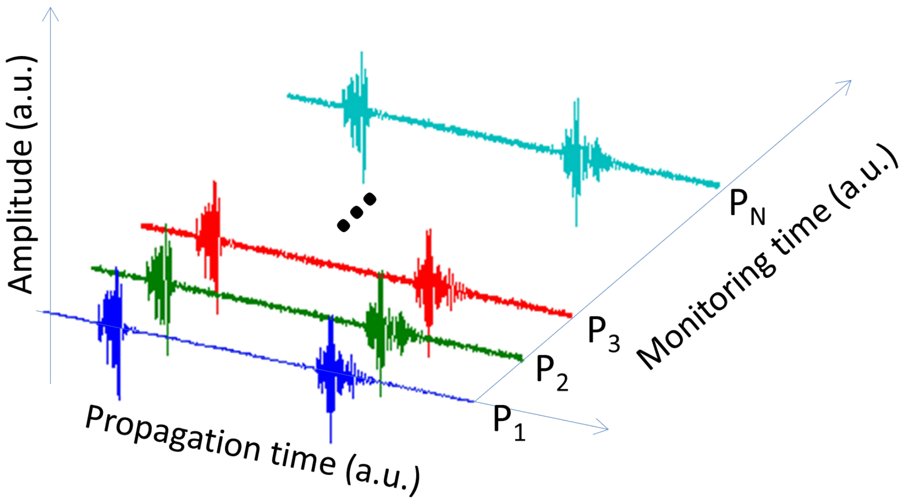

Figure 3. Building the database for the baseline population

3.2. Results and discussion

3.2.1. Univariate analysis (low noise level)

As it was described in section 2.1, in the univariate analysis all damage sensitives features (RMS, Variance, Peak to peak amplitude, maximum amplitude) were considered separately. Figure 4 illustrates the results of discordancy test for all damages types including the baseline and for all damage-sensitives features in the case of a low level of noise. Each value was obtained using equation 1 where the mean and the standard deviation were calculated from the baseline data. Since the baseline data are normally distributed, the threshold was representing 99.73% Gaussian confidence limit. Each damage type has been properly detected even in the case of one defect. Besides, the differentiation between all damage types has been clearly noticed by the increasing steps at 201, 401, 601 and 801. Interestingly, for all damage-sensitives features, there is no indication of false alarm, because all samples have been correctly classified. Despite these satisfactory results, the use of univariate analysis doesn't guarantee that all the damages-sensitive features will behave in same manner. In other words, some features could for example indicate the presence of a defect and others not. This fact reduces the applicability of such a method.

3.2.2. Univariate analysis (high noise level)
When using a high level of noise, results shown in Figure 5 and Figure 6 demonstrate that the threshold increases greatly, therefore, the sensitivity of detection decreases. In this case, all damages type's samples are correctly classified except for the type of one defect where almost all samples are below the threshold which indicates negative false alarms.

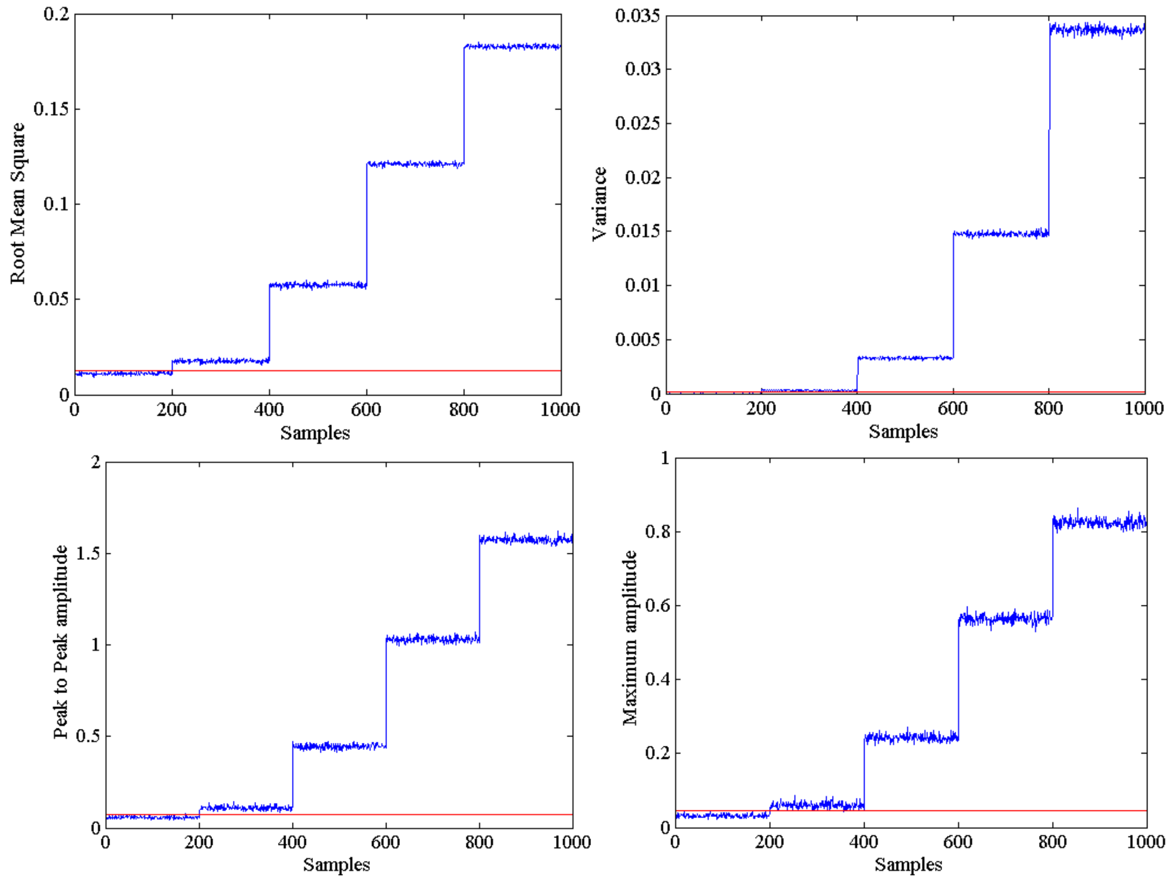

Figure 4. Univariate analysis for a low noise level and for four damage sensitive features (RMS, Variance, Peak to peak amplitude and Maximum amplitude). The red line corresponds to the threshold

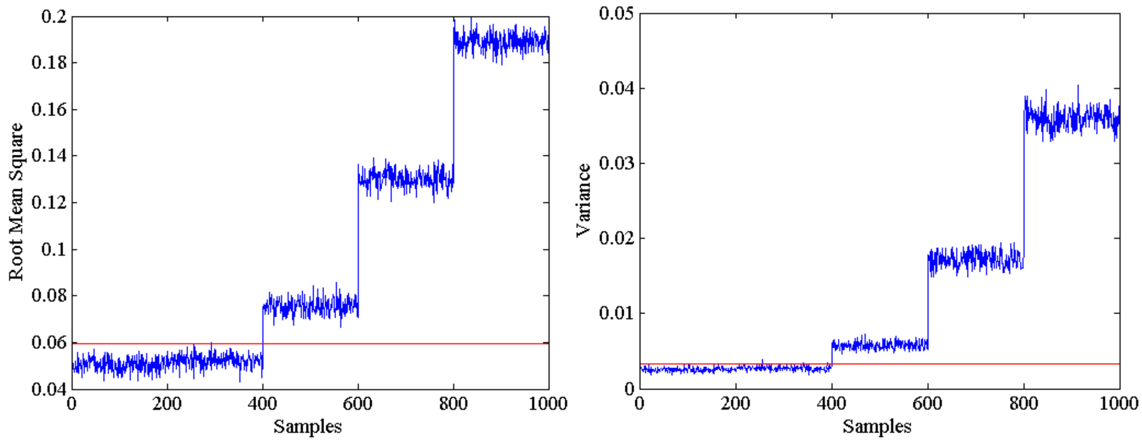

Figure 5. Univariate analysis for high level noise (left: RMS, right: Variance). The red line corresponds to the threshold

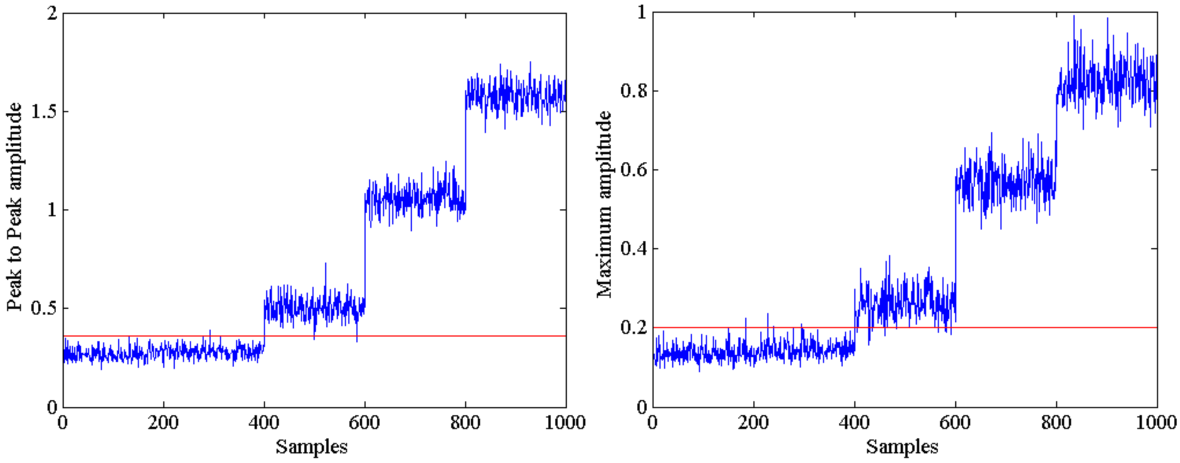

Figure 6. Univariate analysis for high level noise (left: Peak to peak amplitude, right: maximum amplitude). The red line corresponds to the threshold

3.2.3. Multivariate analysis

As it was described in section 2.1.2, in the multivariate analysis the four damages sensitive features are used simultaneously to construct a four dimensional vector. The MSD calculated from the baseline all damages conditions for data corrupted by a low noise level is illustrated in Figure 7 (left). It can be clearly noticed that the multivariate performs well than the univariate analysis. Firstly, the values of the MSDs are much greater the corresponding values of the univariate analysis. As a consequence, the sensitivity to damage is enormously improved. Secondly, there is a good discrimination between all damage types. For high noise level, results shown in Figure 7 (right), are excepted because the sensitivity to damage was decreased and hence one type defect has not been identified.

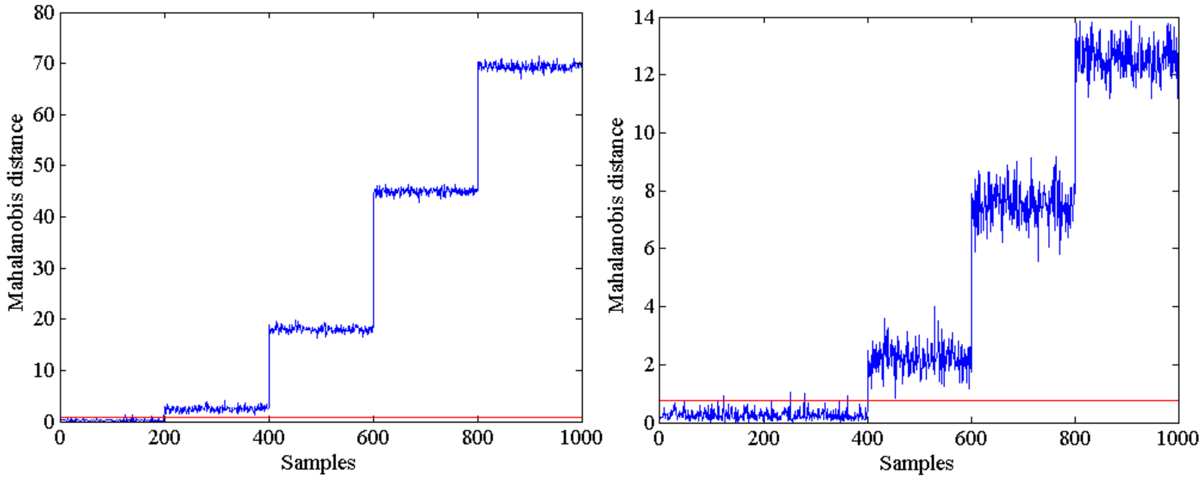

Figure 7. Multivariate analysis for low noise level (left) and for high noise level (right)

4. Conclusion

In this paper, the application of novelty detection as damage identification technique was investigated under varying EOCs. Features extraction were performed using the coefficients of DWT. The use of DWT is essential for noise suppression, and signal compression if there is need. Results have shown that the multivariate analysis is much better than the univariate analysis in terms of sensitivity to damage and discrimination between all damages types. It has been also shown that the level of noise compromises greatly the sensitivity of detection either in univariate or multivariate analysis. In the present work, RMS, Variance, Peak to peak amplitude and Maximum amplitude are chosen to be investigated. Some other features could be used and might be more sensitive that the current tried ones. Further studies concerning this item are to be performed in the near future.

5. References

[1] Farrar CR, Worden K. An introduction to structural health monitoring. Philos Trans A Math Phys Eng Sci 2007;365:303–15. doi:10.1098/rsta.2006.1928.

[2] Sohn H. Effects of environmental and operational variability on structural health monitoring. Philos Trans A Math Phys Eng Sci 2007;365:539–60. doi:10.1098/rsta.2006.1935.

[3] Croxford AJ, Moll J, Wilcox PD, Michaels JE. Efficient temperature compensation strategies for guided wave structural health monitoring. Ultrasonics 2010;50:517–28. doi:10.1016/j.ultras.2009.11.002.

[4] Sohn H, Farrar CR, Hunter NF, Worden K. Structural Health Monitoring Using Statistical Pattern Recognition Techniques. J Dyn Syst Meas Control 2001;123:706. doi:10.1115/1.1410933.

[5] Worden K, Manson G, Fieller NRJ. Damage Detection Using Outlier Analysis. J Sound Vib 2000;229:647–67. doi:10.1006/jsvi.1999.2514.

[6] Pimentel M a. F, Clifton D a., Clifton L, Tarassenko L. A review of novelty detection. Signal Processing 2014;99:215–49. doi:10.1016/j.sigpro.2013.12.026.

[7] Rizzo P, Sorrivi E, Lanza di Scalea F, Viola E. Wavelet-based outlier analysis for guided wave structural monitoring: Application to multi-wire strands. J Sound Vib 2007;307:52–68. doi:10.1016/j.jsv.2007.06.058.

[8] Odening M, Hinrichs J. Using Extreme Value Theory to Estimate Value-at-Risk. Agricultural Finance Review, 2003.

[9] Adeli H, Zhou Z, Dadmehr N. Analysis of EEG records in an epileptic patient using wavelet transform. Journal of Neuroscience Methods Volume 123, Issue 1, 15 February 2003, Pages 69–87

[10] Brussel VU. Noise suppression and signal compression using the wavelet packet transform. Chemometrics and Intelligent Laboratory Systems. Volume 36, Issue 2, April 1997, Pages 81–94