

Research Article

Statistical Pattern-Based Assessment of Structural Health Monitoring Data

Mohammad S. Islam^{1,2} and Ashutosh Bagchi²

¹ University of Nevada, Las Vegas, NV, USA

² Concordia University, Montreal, QC, Canada H3G 1M8

Correspondence should be addressed to Ashutosh Bagchi; ashutosh.bagchi@concordia.ca

Received 28 April 2014; Revised 16 July 2014; Accepted 21 July 2014; Published 17 September 2014

Academic Editor: Qingsong Xu

Copyright © 2014 M. S. Islam and A. Bagchi. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In structural health monitoring (SHM), various sensors are installed at critical locations of a structure. The signals from sensors are either continuously or periodically analyzed to determine the state and performance of the structure. An objective comparison of the sensor data at different time ranges is essential for assessing the structural condition or excessive load experienced by the structure which leads to potential damage in the structure. The objectives of the current study are to establish a relationship between the data from various sensors to estimate the reliability of the data and potential damage using the statistical pattern matching techniques. In order to achieve these goals, new methodologies based on statistical pattern recognition techniques have been developed. The proposed methodologies have been developed and validated using sensor data obtained from an instrumented bridge and road test data from heavy vehicles. The application of statistical pattern matching techniques are relatively new in SHM data interpretation and current research demonstrates that it has high potential in assessing structural conditions, especially when the data are noisy and susceptible to environmental disturbances.

1. Introduction

In a study by Mirza and Haider [1], it was found that more than 40% of the bridges in service in Canada are over 30 years old. Currently, the percentage of older bridges is possibly higher as very few new bridges have been constructed in the past decade. Some of these aging bridges are in urgent need of diagnosis, rehabilitation, or even partial reconstruction in order to make them adequately safe for traffic and prevent sudden collapse and down time. In a similar study conducted by Chase and Washer [2] on the bridge infrastructure in the United States found that about 187,000 bridges representing more than 25% of all bridges were deficient at that time, and about 5,000 bridges were becoming deficient every year. A later study by Research and Innovative Technology Administration [3] puts the number of deficient bridges to about 12% of the total US National and State bridges. The reduction from 25% in 1997 to 12% 2007 is perhaps due to the continued reconstruction and rehabilitation efforts made during that period.

To maintain the overall highway operational safety with current funding limitations, it is beneficial to adopt structural health monitoring (SHM) systems that are continuous, automatic, and low cost [4]. An SHM system includes data acquisition, storage, and broadcasting; it is often used to continuously evaluate the status of an entire structure or structural component. Hence, monitoring a structure would augment the findings from visual inspection and assist in assessing structural conditions and detecting damage in structures.

Sensors measuring strains and vibration of a structure produce signals that always respond to the change in environmental and operational conditions Sohn et al. [5]. Each group of signals can be considered a pattern that has some relation to the structural and ambient conditions. These researchers proposed that if the effect of ambient conditions to the patterns is normalized, the strain and vibration measurements should be identical or close to one another for similar vibration effect as long as the structural vibration property remains the same. Changes in physical

properties, mainly stiffness, should reflect in the processed signal blocks or patterns. Noman et al. [6] and Islam et al. [7] developed statistical pattern based techniques based on the above concepts to interpret the sensor data from structural monitoring systems.

In a continuous or periodic monitoring paradigm, it is often necessary to identify the differences in a signal from a sensor obtained at different times. The objective of this paper is to utilize statistical methods to develop a set of techniques to compare a pair of time history signal blocks corresponding to a sensor. Such a pair of signal blocks could be either obtained from a sensor at different times or one of the signals may represent the real signal and the others simulated using a mathematical model of the system being monitored. The degree of similarity or difference in a pair of such signals can then be interpreted in the context of the condition of the structural system. Initially, acceleration data from roads test on heavy vehicles and corresponding data from simulation have been used to test the developed techniques. Further, the developed techniques have been applied to a case study with strain data from an instrumented bridge pier.

2. Damage Identification Approach by Pattern Comparison

The basic concept of this approach was first proposed by Sohn et al. [8]. It is logical to assume that the patterns in data corresponding to a certain state, either steady or agitated, taken at various points of time of the structure will not vary significantly if the structure does not change significantly. Conversely if the structure has undergone a significant change, it should reflect in the pattern of data in a given state. In order to observe the variation of structure by studying the pattern of signals or data blocks, it is necessary to nominate certain block as reference data block with which patterns of the other data series or blocks are compared. Usually, the reference data blocks for particular conditions are taken from the earlier time of the observation of the structure and other data blocks are called test blocks. The time series model particularly developed for a reference block is defined as a reference model. As the structure undergoes change due to degradation of materials or damage, the pattern of the monitored data block changes. Therefore the pattern of the concerned data block will not match closely that of the reference block.

The degree of similarity between the signal blocks or data series from a sensor could be used to assess the potential damage and reliability of sensors' data. Time series data pattern of the same sensor at different times and time series data pattern of two different sensors at the same time should always follow a similar pattern. The reason of their dissimilarity may either be defective sensor or damage in structure assuming other conditions, that is, load condition and environmental condition remain the same. To assess such dissimilarity various statistical metrics have been used in this research which are described in the following steps. Figure 1 shows a block diagram of the analysis scheme.

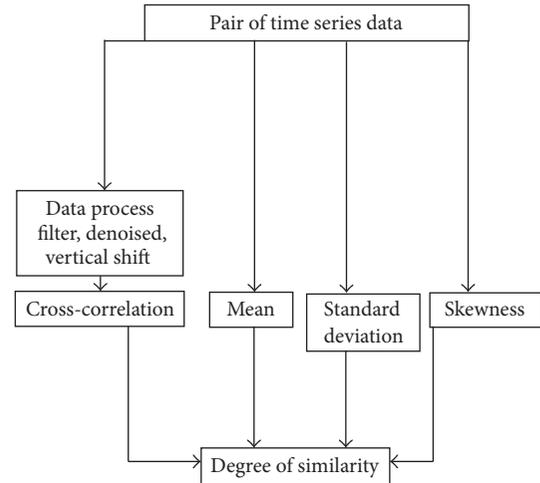


FIGURE 1: Schematic diagram for assessing degree of similarity among sensor data.

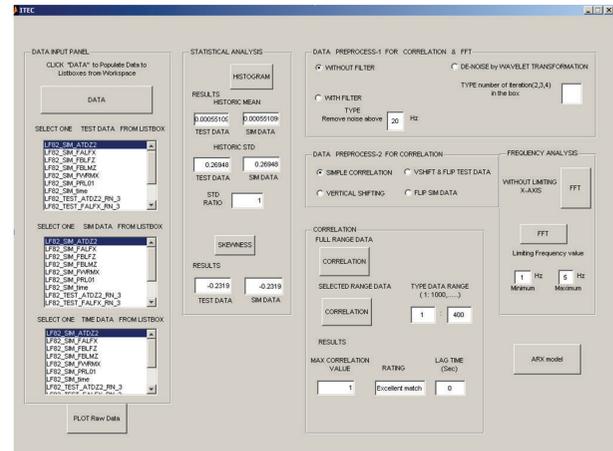


FIGURE 2: Graphical user interface of the tool.

Step 1. Two kinds of data sets have been taken: in the first case, data from the same strain gauges but at different times are considered, while in the second case, signals from two different strain gauges taken at the same time window are considered.

Step 2. Mean, Standard deviation, Skewness parameters are calculated without preprocessing the data and cross correlation is applied with preprocessed data.

Step 3. Finally all results are taken together to find out the relationship between the pair signals under consideration.

A software tool has been developed in the MATLAB environment to implement statistical pattern recognition techniques. A snapshot of the user interface of the tool developed for the present study is shown in Figure 2. As the statistical pattern recognition method does not require any modeling of the bridge structure, this tool can be efficiently used by selecting a time history data of acceleration or strain gauge sensors of any type of structure or system. The data

series' can be preprocessed by Wavelet-based denoising and Filtering in frequency domain. Here, data denoising has been performed using Haar wavelet transform as it is a simple and powerful technique which allows for the rapid evaluation of similarity between time series in large databases. A filter is usually in need to perform frequency dependent alteration of a data sequence. For example, a low-pass filter could be applied to remove noise above 30 Hz from a data sequence sampled at 100 Hz. A more rigorous specification might call for a specific amount of passband ripple, stopband attenuation, or transition width. For bridge vibration a low-pass filter with 30 Hz cut-off frequency is usually sufficient as the usable natural frequencies of a bridge are generally lower than 15 Hz and the fundamental frequency is lower than 5 Hz. After denoising and filtering, a pair of time history signals is compared by computing their mean, standard deviation, skewness, and also the correlation coefficient between them. It may be necessary to shift one signal against another vertically and/or horizontally to improve the closeness or the degree of similarity between them. While computing the correlation coefficient between then signals, they are also shifted horizontally against each other to obtain the best correlation. The amount of shift needed for that is referred to here as the lag time.

It should be noted that the effectiveness of the method in damage detection or structural condition assessment is affected by the environmental factors (e.g., temperature change) and operational loadings which are usually time-varying. Isolating the effects of environmental factors from the effect of operational loadings is a challenging task in structural condition assessment. For that reason, data segments corresponding to the steady state and similar environmental conditions are selected here. Also the time windows considered are small, so that the effect of temperature does not interfere with the similarity metrics. Shifting of the strain signals horizontally/vertically with respect to each other also minimizes the effect of temperature and time varying effect.

3. Interpretation of the Correlation Coefficient for Assessing Degree of Similarity

Several authors have offered guidelines for the interpretation of a correlation coefficient. Cohen [10], for example, suggested the following interpretations for correlation coefficient in psychological research: a correlation coefficient of 0.1 to 0.3 indicates small correlation, 0.3 to 0.5 medium level of correlation, and 0.5 to 1.0 stronger level of correlation. However, all such criteria are in some ways arbitrary and should not be observed too strictly. This is because the interpretation of a correlation coefficient depends on the context and purposes. A correlation of 0.9 may be very low if one is verifying a physical law using high-quality instruments but may be regarded as very high in cases where there may be a greater contribution from many complicating factors. Accordingly, it is important to remember that "large" and "small" correlations should not be taken as synonyms for "good" and "bad" correlations in terms of determining that a correlation is of a certain size. For example, a correlation of

1.0 or -1.0 indicates that the two variables analyzed are equivalent modulo scaling. Scientifically, this more frequently indicates a trivial result than an important one. But in the context of SHM data, the following interpretation is proposed by Islam and Bagchi [11]: the correlation coefficient exceeding 0.9 is interpreted as an excellent match, between 0.7 and 0.89 indicates a good match, the correlation between 0.5 to 0.69 indicates a fair match, and lower than 0.5 is regarded as a poor match.

4. Details of the Data Used in the Present Study

Data from two different sources have been used in the present study. The first source of data corresponds to road tests on heavy vehicles where the vibration data were measured using accelerometer installed at different points of interest on the vehicle (e.g., truck, bus). The acceleration data corresponding to the road tests were also simulated using the vehicle-road interaction model using commercial software. The test and simulation were conducted by a manufacturer of such vehicles. These data are used here in Test 1 and presented later in this paper.

The second source of data belongs to the structural health monitoring systems of the Portage Creek Bridge located in Victoria, British Columbia (Figure 3(a)). The bridge has three spans and is 407 ft (124 m) long with a reinforced concrete deck supported on two reinforced concrete piers; it has abutments on H piles. The deck has a roadway width of 53 ft (16.2 m) with two 6 ft 6 inch (1.98 m) sidewalks and aluminum railings. On each column, eight bidirectional rosette type electrical strain gages are located. Moreover, on Column 2, a temperature sensor is located, as described in Figure 3(b). Each bidirectional strain gage has two channels, one for horizontal strain and another for vertical strain, producing a total of 16 sets of strain data for each column. Then strain data from this bridge acquired during 2003–2006 period are used here in Tests 2 and 3 as presented later in this paper. The data were originally sampled at 32 Hz as the fundamental frequency of the bridge is less than 2 Hz. The data was resampled at 1 s and 1 hr intervals for further analysis. The sensor data from the instrumented bridge was obtained from the Intelligent Sensing for Innovative Structure (ISIS) Canada Research Network.

4.1. Test 1: Acceleration Data from Road Tests on Heavy Vehicles. The data set in Test 1 contains a pair of time series data, one corresponding to vibration response of a heavy vehicle obtained during a road load Test 2, while the other corresponds to similar data produced using simulation. It is required to determine the level of similarity between the experimental and simulation data to test reliability of the simulation data. The results of the similarity test between the signals are outlined below.

- (1) Raw data of acceleration time history from the test and simulation are plotted to get a preliminary idea of the relation between them visually (Figure 4). Relevant statistical parameters such as mean, standard

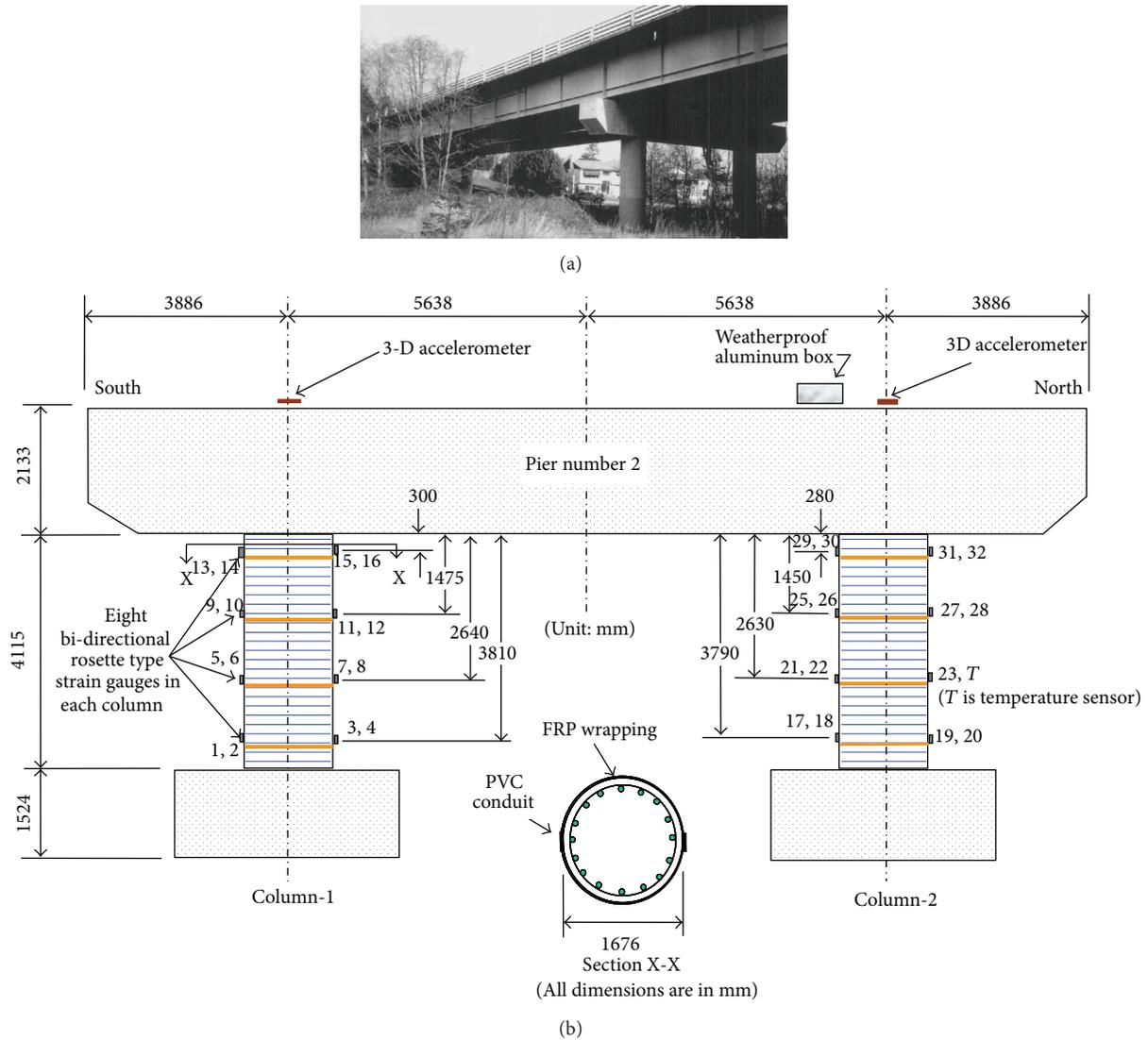


FIGURE 3: Portage Creek Bridge, Victoria, British Columbia, Canada: (a) photo; (b) sensor locations on columns of Pier 2 of the Portage Creek Bridge [9].

deviation and skewness of the test and simulated data are calculated and shown in Table 1.

- (2) The data are then processed in a number of ways, such as removing the mean for each series to eliminate the effect of bias from both signals and filtering the test data using Fourier Transform or denoising using Wavelet Transform. The correlation on the full range data is performed, which also calculates the lag between the signals to achieve maximum correlation as shown in Table 2.
- (3) In order to establish the relation between the two series, at different time windows, the series are also divided into multiple segments in time and each pair of them is analyzed, separately. The results of correlation on such segmented series' are shown in Table 3. The variation of the correlation coefficient is shown in Figure 5.

TABLE 1: Various statistical parameters on the time series data.

	Test data	Simulated data
MEAN	-532.64	-606.34
STD	283.47	170.52
SKEWNESS	0.0788	-0.024

- (4) Filtering is performed by specifying the cut-off frequency.

From the results presented here, data sets from test and simulation are found to be similar. The correlation coefficient of 0.89 on the raw data or 0.91 on filtered or denoised data indicates a high degree of similarity. The correlation improves when smaller time windows are used on data sets. The test data filtered in frequency domain using a low-pass filter with cut-off frequency of 30 Hz in one case and denoised with

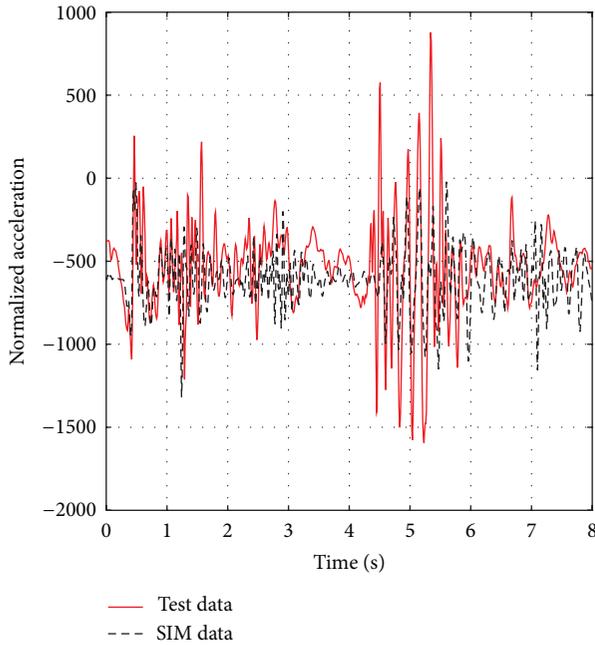


FIGURE 4: Plots of the test and simulated data in their raw forms (SIM data indicate SIMULATED data).

TABLE 2: Effect filtering and denoising of time series data.

Data preprocess	Maximum correlation	Lag time (sec)
Without filter	0.89	0
With filter (30 Hz)	0.89	0
Denoise (4)	0.91	0

wavelet transform, in another case. The statistical comparison metrics such as correlation coefficient are not affected too much for these signals. One of the main reasons why a high degree of similarity between the test and corresponding simulated data is achieved is the controlled environment where the test was conducted and sophistication of the simulation process. It will be shown later that in case of structural health monitoring, data are usually noisy and data obtained from the same sensor at different times, but under similar conditions, may not produce a very high degree of similarity.

4.2. Test 2: Strain Monitoring Data from Portage Creek Bridge Piers. The data in Test 2 contain a pair of time series representing strain data in two different time windows from a strain gauge (Channel Strain1_1_c1, which is at the bottom of Column 1 measuring vertical strain) installed on the bridge pier of Portage Creek Bridge. These statistical patterns of these two data series' are compared to determine whether the response of the structure at the corresponding time windows is similar (Figure 6). In this case the data from one of the strain sensors, Strain1_1_c1 data have been used. Two series of the strain data, one acquired on 25th of December 2003, starting at 11.00.00 and the other on 25th of August 2006, starting at 11.00.00 are considered for the comparison. Each

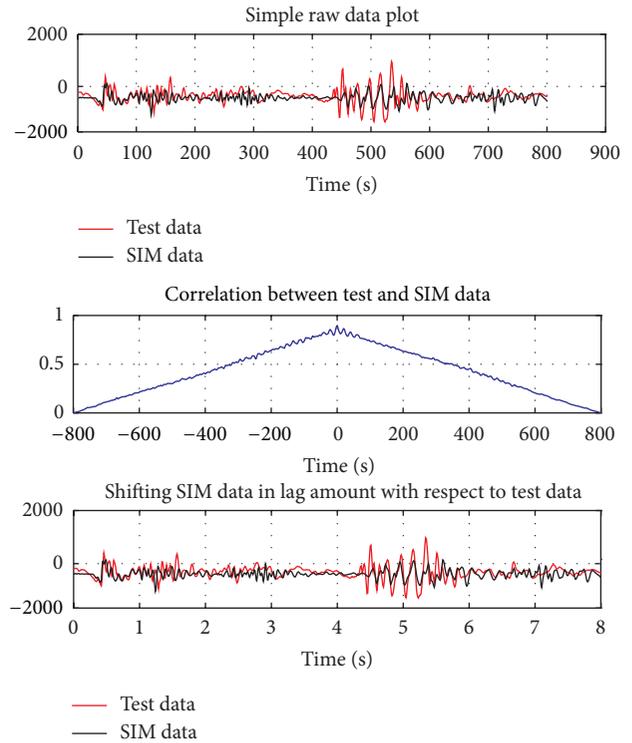


FIGURE 5: Correlation between TEST DATA and SIMULATED DATA (SIM data indicate SIMULATED data).

of these data sets has a length of 302 s. The mean values of the two data sets are different (Table 4) and correlation coefficient is found to be 0.74 which indicates that they have a good match in terms of their patterns, but one signal is stronger than the other. The data sets are produced by the same sensor and the time difference between them is two years and eight months. So it can be said that there might be some minor changes that may have occurred in the structure or the magnitude of the excitation. Standard deviations are almost the same and they skew at the same direction that shows they follow same pattern which indicate the data sets are from the same sensor and the changes in the structural properties are less likely. Frequency analysis of the pairs of the time series of the strain as considered here does not produce a meaningful information (Figure 8) as the data represent the steady state response of the structure when the vibration response is not appreciable. This result is consistent with an earlier work by Noman et al. [6] on the bridge using the same sets of data, but with a different statistical pattern recognition algorithm. In their study, first AR process was applied to extract the coefficients which were then statistically modeled for damage classification by X-bars. From the X-bars of strain and vibration readings, the percentage of outliers was not found to be so high to indicate any damage in the structure or structural degradation. Secondly, pattern comparison based on fitting of the reference models to test blocks was performed. Computed *R*-values that represent the goodness of fit also did not show any trend or consistent discrepancies to indicate any damage in the structure.

TABLE 3: Analysis of subsets of time series data.

Data range	With raw data		With filtering		With denoising	
	Maximum correlation	Lag time (sec)	Maximum correlation	Lag time (sec)	Maximum correlation	Lag time (sec)
1: 200	0.89	0.06	0.896	0	0.93	0
200: 400	0.95	0	0.94	0	0.95	0
400: 600	0.85	0	0.85	0.09	0.90	0
600: 800	0.95	0	0.94	0.12	0.95	0

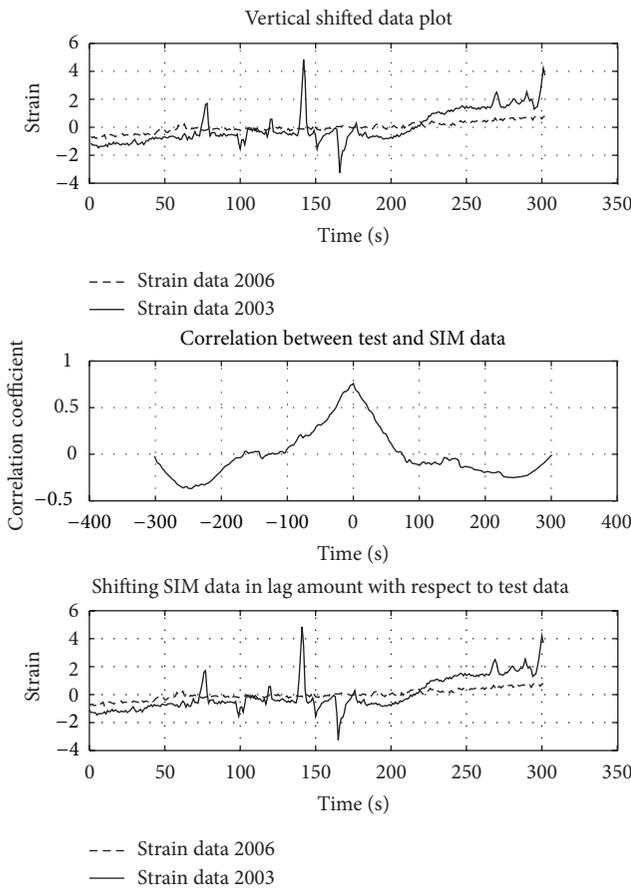


FIGURE 6: Correlation of two data sets, starting time of query 2003/12/25,11.00.00 and starting time of query 2006/08/25,11.00.00 (SIM data indicate SIMULATED data).

4.3. Test 3: Strain Monitoring Data from Portage Creek Bridge Piers. The data in Test 3 contains a pair of time series' representing the strain data in a given time window from two different strain gauges (Channels Strain1.1.c1 and Strain4.1.c2) installed on the bridge pier. The objective in this test is to identify the relationship between the data generated at the same time, but from different sensors. These data sets each having a length of 220 s were acquired on 1st of April 2006, starting at 10.00.00. The mean values of the two data sets are different and correlation values indicate they are only 31% similar (Table 4). The data sets are produced by two different sensors at the same time window (Figure 7). They are skewed in the same direction but their skewness values are quite different which indicates that the data sets are not similar.

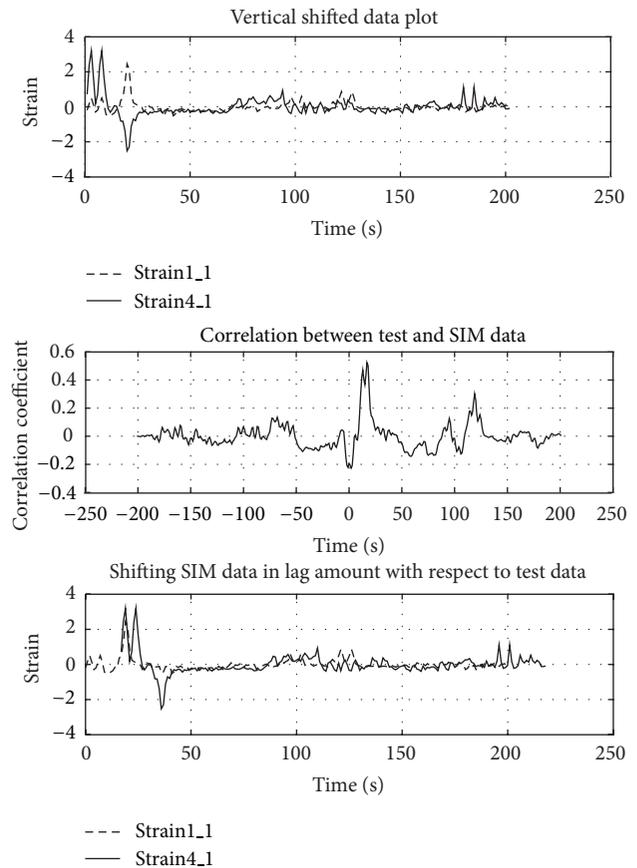


FIGURE 7: Correlation of Strain1.1.c1 and Strain4.1.c2 data (SIM data indicate SIMULATED data).

Frequency analysis of the above pairs of the time series of the strain does not produce meaningful information (Figure 9) as in the earlier case. However, a significant difference in Fourier amplitude is noted between the two series' indicating that while the data patterns may be similar, the origin of each data series is different.

5. Discussion and Conclusions

Structural health monitoring (SHM) has been widely studied during the past two decades and significant progress has been achieved through the development of new sensors and system that are capable of monitoring the performance of a structure. It is found from various studies that structural damage affects the dynamic properties of a structure, causing a change in

TABLE 4: The statistical parameters on different pairs of time history of strain.

	Strain1.1.c1 data at different times		Data from different sensors at a given time in 2006	
	2003	2006	Strain1.1.c1	Strain4.1.c2
Mean	-505.52	165.49	-261.335	42.9626
STD	1.34	0.96	0.47	0.46
Skewness	0.31	0.51	1.58	6.1
Correlation	0.74		0.31	
Lag time (sec)	0		43	

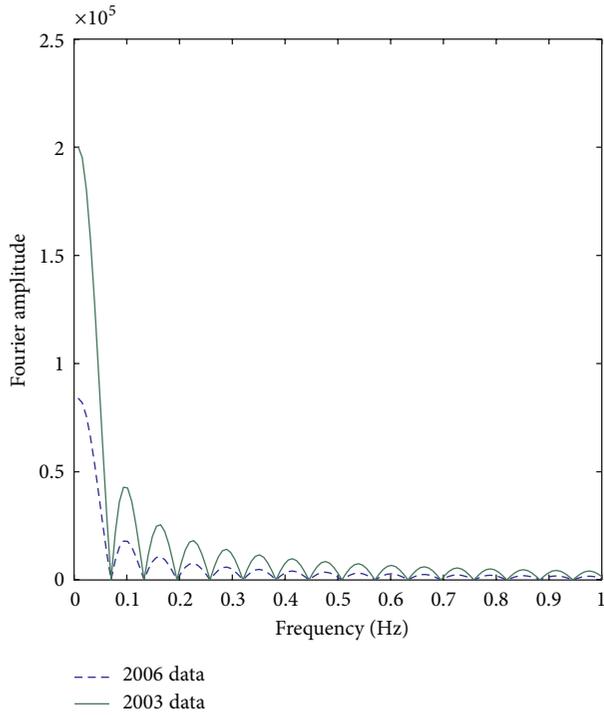


FIGURE 8: FFT of Strain1.1.c1 data from 2003 and 2006.

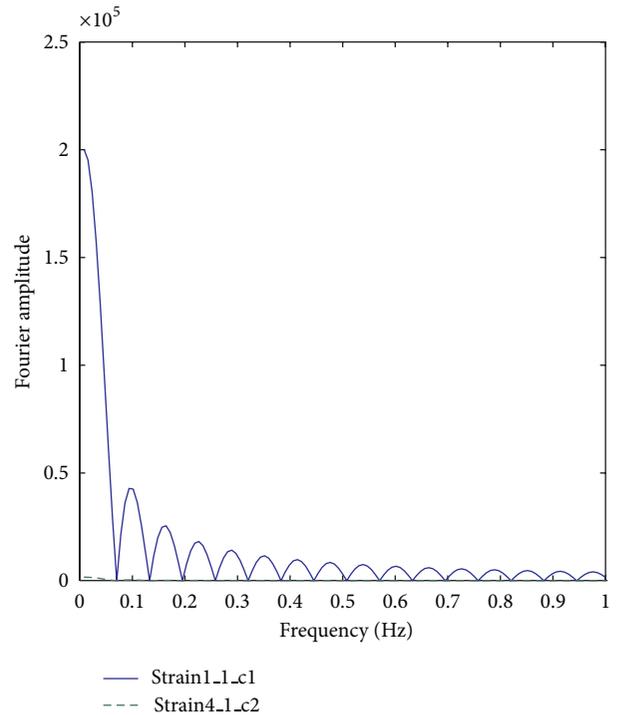


FIGURE 9: FFT of Strain1.1.c1 and Strain4.1.c2 data from 2006.

the vibration signals, that is, strain and acceleration time histories. Damage detection can be performed using time series analysis of vibration signals measured from a structure before and after damage. In this study the method proposed is degree of similarity which is used to assess the structural condition and sensor reliability. The time series data of the same sensor at different times should have similar patterns if the loading or ambient condition remains unchanged (e.g., steady state or excited state). The time series data pattern of two different sensors at the same time should also have a consistent pattern irrespective of the time window considered. The reason for their dissimilarity may be due to either defective sensor or damage in structure assuming that other conditions such as load and environmental conditions are unchanged. Numerous tests have been done to assess the structural condition and reliability of sensors' data by quantifying the degree of similarity between the pairs of sensor data. In this paper three test results have been presented. In Test 1, mean, standard deviation and skewness values of the two data

sets are found to be close, which indicate that the simulation data are reliable. Thus, expensive and time consuming tests can be optimized based on the results of the simulation if they are proved to have high fidelity with the test results. In Test 2, while the mean values of the two data sets from the same sensor, but taken at different time windows, are found to be different, the standard deviation and skewness values are close and correlation is 74%. Thus, the degree of similarity of strain data in two different time windows can be considered high. It can be implied that structural conditions are not degraded from the initial condition. In Test 3, data from two different sensors at the same time window have been considered, where the mean, standard deviation values of the two data sets are not found to be close, correlation values indicate they are only 31% similar, and while the data sets are skewed in the same direction, their skewness values are quite different. It can be implied that sensors are measuring different quantities, but the data are related. When the data from all different pairs of sensors are thus tested, if

the similarity pattern is consistent among all the pairs, the sensors can be said to be working reasonably for the time windows considered. Apart from Test 3 presented here, other pairs of data series' from different sensors are compared in this study to ascertain that they follow consistent patterns. The results are found to be consistent with the study reported in Noman et al. [6] on Portage Creek Bridge using a different approach as discussed earlier.

Based on the results of the present study, the following conclusions/observations are made.

- (i) To determine the reliability of the sensors' data and to assess the structural condition, a set of statistical parameters, mean, standard deviation, skewness, and correlation can be used in a holistic manner as demonstrated in the present study.
- (ii) The information about the degree of similarity among various sensors data in a structure and detection of defective sensors using the proposed methods can be extended to detect damage in structure.
- (iii) Low degree of similarity among sensors data or multiple sensors detected as the responsible ones for the change of the statistical patterns of data indicates either the presence of damage/degradation in structure or change in load/environmental conditions.

However, the observations from the present study also indicated some limitations of the application of the developed techniques which are explained below.

- (i) In the present study, it is assumed that data dissimilarity occurs due to the presence of damage in structure. But data dissimilarity may also occur due to the presence of individual or combined effects of sensor malfunction or excessive load or presence of damage in the structure. Further work should consider these effects in isolating the individual factors.
- (ii) For correlation analysis, there is no concrete benchmark to rating the similarity. The coefficient of determination (i.e., R^2 measure) or other similarity metrics might be used to define the rating.
- (iii) In the case of the bridge, the data from steady state conditions, when no vehicle passes over the bridge, have been considered. However, there is always a low frequency noise in the data due to ambient disturbance. This may be responsible for a lower level of correlation as compared to the data sets obtained from the tests on heavy vehicles where the tests were conducted in a controlled environment.

It should also be noted that the volume and type of data available for the present study are limited. Although the statistical parameters used here are straight forward, their application in the context of the bridge monitoring data and the road test data needs a careful integration of these parameters so that the interpretation of the results is coherent. The methodology should be further studied with a richer source of data and a case study structure with known damage conditions.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The support of the Quebec research grant agency FQRNT is gratefully acknowledged. The authors also thank ISIS Canada Research Network and International Truck and Engine Corporation for providing sample data required in this research.

References

- [1] S. M. Mirza and M. Haider, *The State of Infrastructure in Canada: Implications for Infrastructure Planning and Policy*, Infrastructure Canada, Ottawa, Canada, 2003.
- [2] S. B. Chase and G. Washer, "Non-destructive evaluation for bridge management in the next century," *Public Roads*, vol. 61, no. 1, 1997.
- [3] RITA, *Research and Innovative Technology Administration*, U.S. Department of Transportation (US DOT), Washington, DC, USA, 2010.
- [4] L. Ping, *A statistical based damage detection approach for highway bridge structural Health Monitoring [Ph.D. dissertation]*, Iowa State University, 2008.
- [5] H. Sohn, J. A. Czarnecki, and C. R. Farrar, "Structural health monitoring using statistical process control," *Journal of Structural Engineering*, vol. 126, no. 11, pp. 1356–1363, 2000.
- [6] A. S. Noman, F. Deeba, and A. Bagchi, "Health monitoring of structures using statistical pattern recognition techniques," *Journal of Performance of Constructed Facilities*, vol. 27, no. 5, pp. 575–584, 2012.
- [7] M. S. Islam, A. Bagchi, and A. Said, "Developing a pattern-based method for detecting defective sensors in an instrumented bridge," *Journal of Civil Structural Health Monitoring*, vol. 3, no. 3, pp. 195–206, 2013.
- [8] H. Sohn, C. R. Farrar, N. F. Hunter, and K. Worden, "Structural health monitoring using statistical pattern recognition techniques," *Journal of Dynamic Systems, Measurement and Control*, vol. 123, no. 4, pp. 706–711, 2001.
- [9] S. Huffman, A. Bagchi, A. A. Mufti, K. Neale, D. Sargent, and E. Rivera, "GFRP seismic strengthening and structural health monitoring of portage creek bridge concrete columns," *The Arabian Journal For Science and Engineering*, vol. 31, no. 1, pp. 25–42, 2006.
- [10] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, Erlbaum, Hillsdale, NJ, USA, 2nd edition, 1988.
- [11] M. S. Islam and A. Bagchi, "Application of Statistical Pattern Recognition Techniques in Structural Health Monitoring," in *Proceedings of the 11th CANSIMART International Workshop on Smart Materials and Structures*, Montreal, Canada, 2008.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

