

Clustering Techniques and Artificial Neural Network for Acoustic Emission Data Analysis

U. B. Attanayake^{1*}, H. M. Aktan², J. Mejia³, and R. Hay⁴

¹Western Michigan University, Kalamazoo, United States

²Western Michigan University, Kalamazoo, United States

³TISEC Inc., Morin Heights, QC J0R 1H0, Canada

⁴TISEC Inc., Morin Heights, QC J0R 1H0, Canada

E-Mail: upul.attanayake@wmich.edu, TP: +1 269 276 3217

Abstract: Acoustic emission (AE) sensor technology is commonly used for real-time monitoring of fatigue sensitive details. This is mainly due to its ability to detect fatigue events (crack initiation and opening) by mounting sensors in the vicinity of potential crack location. Also, AE data can be used for damage location detection. Even though AE provides many capabilities with regard to fatigue monitoring, many implementation challenges exist. A majority of the challenges is associated with noise elimination, AE signal analysis, and interpretation of the results. This article describes AE implementation for monitoring a fatigue-sensitive detail and use of data analysis techniques such as cluster analysis, non-linear mapping (NLM), and three-class classifiers to identify the relationship of each cluster to the characteristics of crack opening signals, background noise, and structural resonance.

Keywords: Acoustic Emission, Artificial Neural Network, Cluster Analysis, Data Analysis, Fatigue Monitoring.

1. Introduction

Fatigue is one of the most critical problems for steel bridges as well as for any steel structures that needs to be considered during design and operation. Irrespective of the causes of cracking, fatigue events (i.e., crack initiation or crack growth) need to be identified to and monitored to assure safety. An acoustic emission (AE) monitoring system with strain gages is one of the most effective technologies for fatigue event detection. AE has been successfully implemented in the field and evaluated for continuous monitoring of fatigue-sensitive details. At this time, AE is the only technology that is capable of real-time monitoring of fatigue events and providing data for damage location detection. The reasons for widespread use of the technology are;

- AE can be used as a local as well as a global crack growth monitoring tool [1-3].
- AE is capable of locating the source of failure [2, 4].
- AE is capable of detecting and locating defects in areas obscured from view or in areas that are difficult to inspect (e.g. weld defects, material imperfections, etc.) [2, 5].

- The data from an AE monitoring system can be used to track the history of crack growth activity [1].
- Parametric data (strain, displacement, temperature, etc.) can be used to correlate AE events in order to improve the accuracy of data analysis results [2].
- Very minimal surface preparation is required to mount AE sensors [1].
- Frequent access to a detail is not required once the sensors are installed [1, 2].
- Technology has been used for decades in many disciplines and the experience is well documented.

With any technology there are advantages as well as implementation challenges. This is no different for AE technology implementation and data interpretation. The following is an abbreviated list of challenges associated with AE monitoring, data analysis, and results interpretation [1, 2, 4-7]:

- AE monitoring requires extensive expertise to plan, set up the sensors, test, and interpret results.
- The service environment contributes extraneous noise to the signals.

- Challenging to implement standard noise reduction methods because the signals are transient and random in time.
- A large amount of data is recorded during monitoring; hence, effective data analysis and management are necessary, especially for long-term monitoring.
- Complicated geometries and low strength signals make tasks more difficult.
- Unable to detect dormant cracks using AE monitoring.
- Unable to quantify the extent of damage using AE data.
- Unable to repeat AE measurements.

A majority of the challenges is associated with noise elimination, AE signal analysis, and interpretation of results. This article describes AE implementation for monitoring a fatigue sensitive detail (local monitoring) and use of data analysis techniques such as cluster analysis, non-linear mapping (NLM), and three-class classifiers to identify the relationship of each cluster to the characteristics of crack opening signals, background noise, and structural resonance; thus, eliminating a majority of challenges associated with noise elimination, AE signal analysis, and interpretation of the results.

2. AE Monitoring System Implementation

2.1 System Configuration

The monitoring system selected for this project has a low-power computer. The operating system and essential software are installed in a 2 GB hard drive. The supplemental software and sensor data are stored in a 110 GB drive. The monitoring system components are shown in Figure 1. The monitoring system consists of only one AE board (PCI/DSP-4) with four-channels. The system capability can be extended to accommodate 4 AE boards with a total of 16 AE sensors. The monitoring system included 4 PK30I narrow band sensors (Figure 2a). Each AE sensor has an integral, ultra-low noise, low power preamplifier with a 26 dB voltage gain. The frequency range for the sensor is 200 – 450 kHz with a resonance frequency of 300 kHz. The data acquisition settings included a 40 dB preamplifier voltage gain, 45 dB threshold, 1 kHz to 1 MHz analog filter range, and waveform settings of 1MSPS sample rate, 256 μ s pre-trigger, and 1k waveform length. Magnetic holders were used to mount the AE sensors (Figure 2b and c).

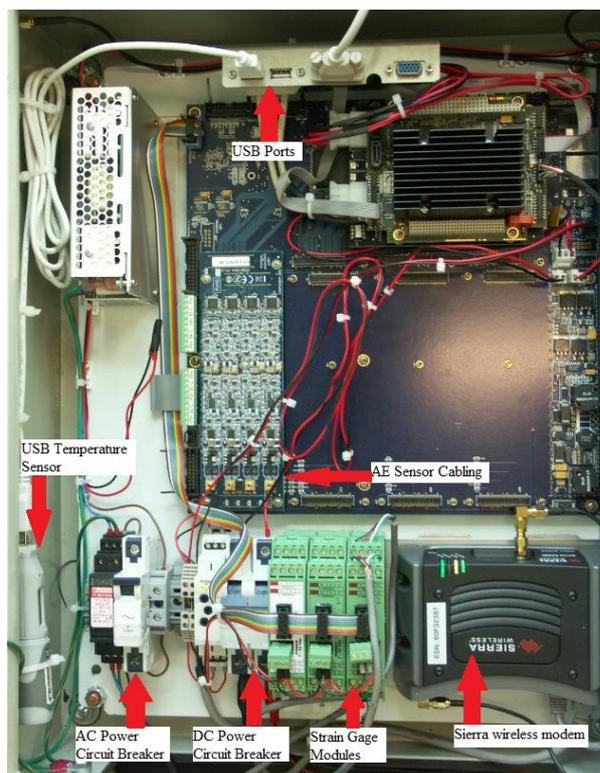


Figure 1: Monitoring system components in the enclosure



(a) An acoustic emission sensor (b) A spring-loaded magnetic holder



(c) AE sensors mounted on a steel girder

Figure 2: (a) An acoustic emission sensor, (b) a spring-loaded magnetic holder, and (c) AE sensors mounted on a steel girder

2.2 System Implementation

The bridge (S16 of 11015) is located in Stevensville, Michigan, and carries I-94 over Puetz Road. After reviewing biennial inspection reports and conducting a field visit to document the bridge superstructure and substructure condition, the I-94 EB bridge was selected. The longest span of the EB Bridge is 56 ft - 6 in. and has a 54.5° skew. The

span is supported on an integral abutment and a pier with expansion bearings. The superstructure consists of 12 steel I-girders (10-W30×108 and 2-W30×99) and a 9 in. thick cast-in-place concrete deck. The girders are connected transversely using intermediate and end diaphragms. The partial depth diaphragm connection detail is classified as a category C' fatigue-sensitive detail [8].

Once the AE sensors were mounted and the data acquisition started, AE source locations appeared on a source location page. Pencil lead break (PLB) signals are used to demarcate the area of interest as well as to fine-tune the data acquisition settings: preamplifier voltage gain and the signal threshold. Pencil lead breaks were performed and the waveforms were recorded. Figure 3 shows the source locations calculated based on the PLB signal arrival time. These source locations mark the boundaries of the area of interest for continuous monitoring.

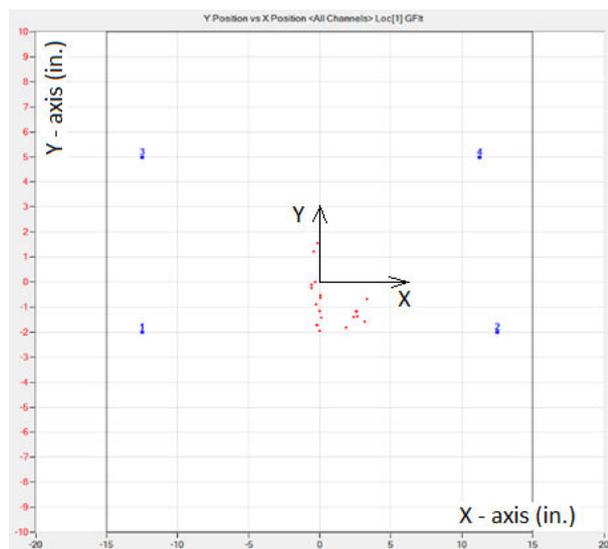


Figure 3: AE source locations generated through pencil lead breaks

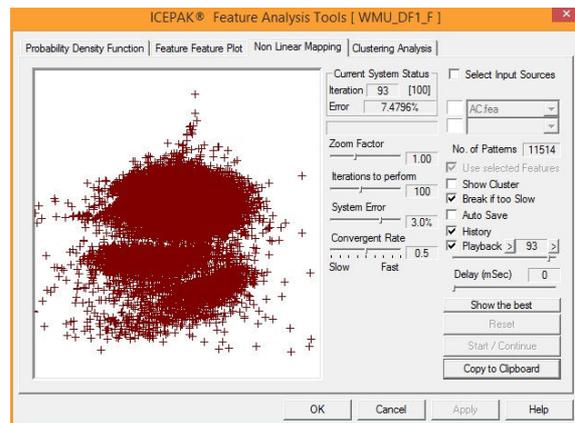
3. AE Data Analysis and Results Interpretation

The ICEPAK™ software developed by TISEC Inc. was used to classify data via pre-trained classifiers designed by the software package [9].

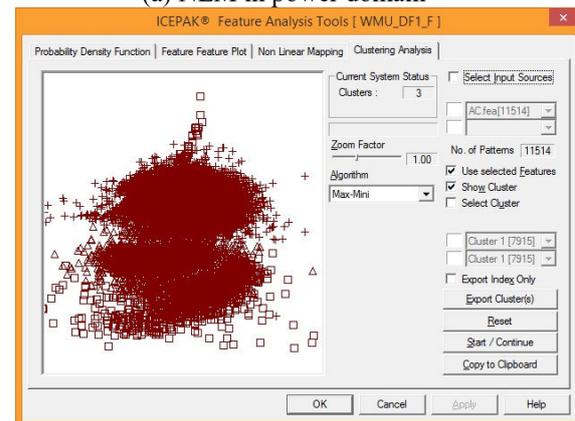
3.1 Unsupervised Learning via Clustering

The data collected from the bridge was examined directly to identify any significant similar AE activity formations using non-linear mapping (NLM) and clustering analysis available in ICEPAK™. NLM can be performed in time, power, phase, cepstral, and auto-correlation domains. The data set used for NLM and clustering analysis included little more than 11,000 AE signals that were above the set threshold of 45dB. NLM

was performed using one feature domain at a time to visually detect significant naturally forming concentrations. Out of the 5 domains, the spectral power domain produced three significant concentrations as shown in Figure 4a. Clustering was performed using the same spectral power domain features, and produced three significant concentrations as presented in Figure 4b. The clusters are aligned with the visual presentation of the NLM result.



(a) NLM in power domain



(b) Clustering in power domain

Figure 4: NLM and clustering in power domain

The individual data clusters were exported and labelled as [c1], [c2], and [c3]. Then, each cluster was used to train a three-class classifier. Four statistical classifiers (i.e., linear discriminant, K-nearest neighbour, empirical Bayesian, and minimum distance classifiers) and a neural network classifier were tested. The design of the classifiers included optimizing the feature sets. The design procedure included separating available data into two groups; one was used to train the classifiers and the other to test the performance of the classifiers. As an example, results of the linear discriminant three-class classifier is shown in Figure 5. Cluster [c1] had a total of 7,915 data points. This set was separated into two groups of 3,957 and 3,958 data points for training and testing, respectively. When the linear discriminant three-class classifier was

trained with 3,957 data points, the data was classified into three classes with rejections. As shown in Figure 5, classes 1, 2, and 3 contain 3777, 2, and 0 data points with 178 rejections. The classification rate is 95.45%. In other words, 95.45% of the data in [cl1] falls into class 1. A similar process was employed for [cl2] and [cl3] data sets, and yielded classification rates of 94.99% and 95.95%, respectively. When all three data sets were considered, the linear discriminant three-class classifier yielded a weighted average classification rate of 95.43% for training (Figure 5). Overall, all the classification methods yielded very high classification rates for training as well as for testing.

Linear Discriminant Classification Results							
	Class	1	2	3	Reject	Total	Percent
c11.cxf	1	3777	2	0	178	3957	95.45%
c12.cxf	2	0	1005	5	48	1058	94.99%
c13.cxf	3	0	3	711	27	741	95.95%
Training : 95.43%							
	Class	1	2	3	Reject	Total	Percent
c11.cxf	1	3822	2	0	134	3958	96.56%
c12.cxf	2	0	995	6	57	1058	94.05%
c13.cxf	3	0	7	715	20	742	96.36%
Testing : 96.08%							

Figure 5: Linear discriminant three-class classifier training and testing results

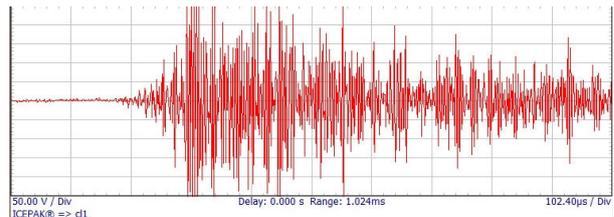
Next, the PLB data was tested against this three-class classifier with a rejection option. The rejection option is triggered when an incoming signal cannot be classified with an acceptable level of confidence. The PLB data file contained 100 data points. The PLB data fell into class 1 and 2 but not class 3, with a lot of rejections. The type of waveforms in the class 1 and 2 along with the rejected ones are shown in Figure 6.

The analysis of waveform characteristics of signals in each class and the rejected group yielded the following observations:

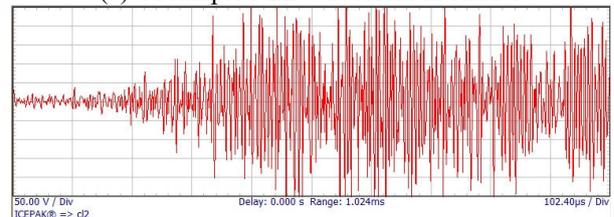
- The waveforms classified as class 1 usually have very fast rise times, relatively quiet pre-trigger portion, and a broad spectral content spanning from 100 to 400 kHz (Figure 7).
- The waveforms classified as class 2 usually have very fast rise times; however, the pre-trigger portion may show some small structure, and the main pocket contains multiple ringing peaks. The spectral content mainly centers around 150 kHz with very little or nothing above 250 kHz, and nothing below 100 kHz (Figure 8).
- The waveforms classified as class 3 usually have a slower rise time, and the spectral content is mainly located below 100 kHz and centered around 50 to 75 kHz. Moreover,

there is absolutely nothing above 200 kHz (Figure 9).

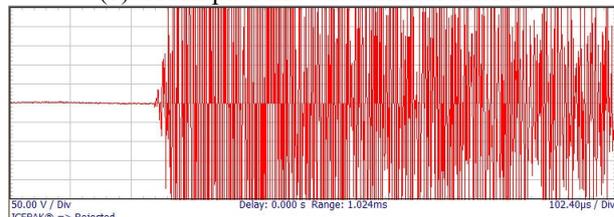
- The waveforms classified as “rejected” are mostly associated with over-saturated clipped waveforms, and some have slow changing, somewhat smooth, waveform centered around 50 kHz (Figure 10).
- The number of waveforms being classified as class 1 is 4 to 6 times more than those of class 2 and class 3 while the sizes of class 2 and class 3 are relatively comparable. In general, there are about 5% of waveforms being rejected.



Note: X-axis: Time (102.4 μs/Div);
Y-axis: Voltage (50 V/Div)
(a) A sample PLB waveform in class 1

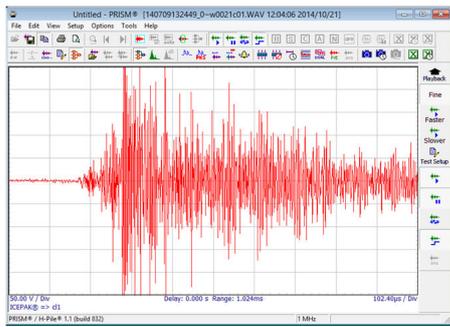


Note: X-axis: Time (102.4 μs/Div);
Y-axis: Voltage (50 V/Div)
(b) A sample PLB waveform in class 2

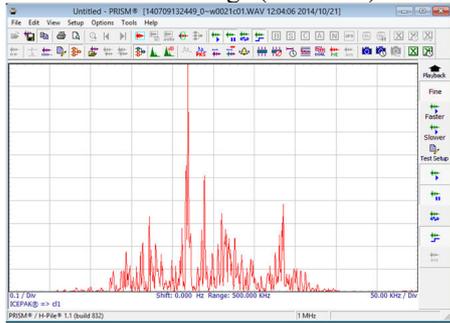


Note: X-axis: Time (102.4 μs/Div);
Y-axis: Voltage (50 V/Div)
(c) A sample rejected PLB waveform

Figure 6: Sample waveforms in class 1, class 2, and the rejected group

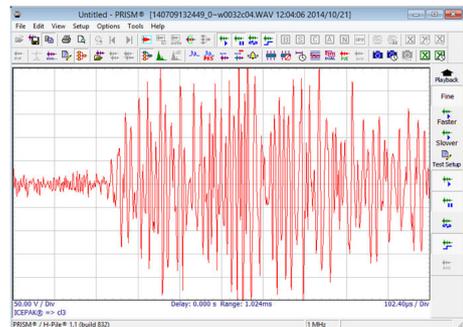


(a) Time domain
X-axis: Time (102.4 μ s/Div)
Y-axis: Voltage (50 V/Div)

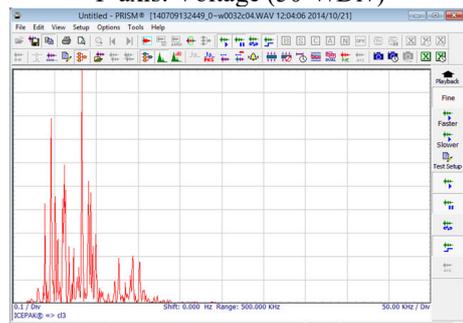


(b) Frequency domain
X-axis: Frequency (50 kHz/Div)
Y-axis: Amplitude (0.1/Div)

Figure 7: A sample class 1 waveform and its power spectrum

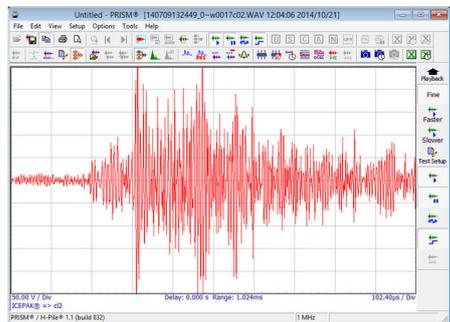


(a) Time domain
X-axis: Time (102.4 μ s/Div)
Y-axis: Voltage (50 V/Div)

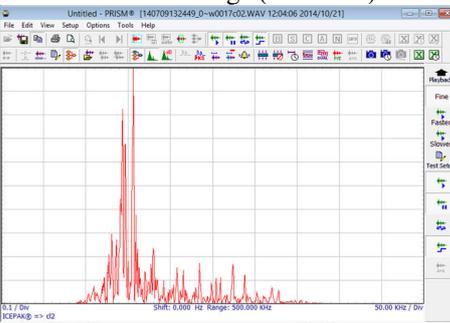


(b) Frequency domain
X-axis: Frequency (50 kHz/Div)
Y-axis: Amplitude (0.1/Div)

Figure 9: A sample class 3 waveform and its power spectrum

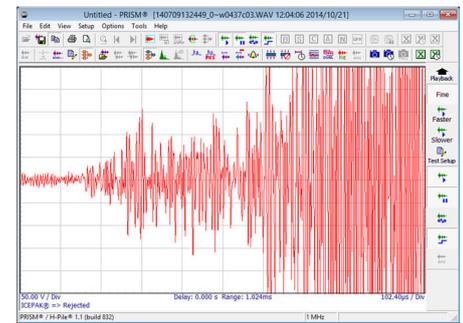


(a) Time domain
X-axis: Time (102.4 μ s/Div)
Y-axis: Voltage (50 V/Div)

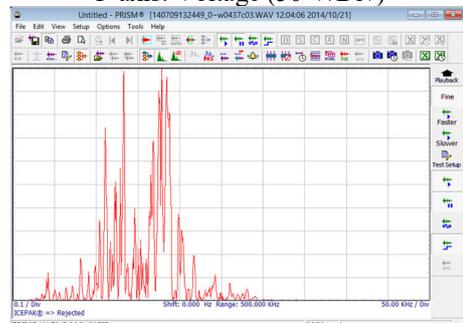


(b) Frequency domain
X-axis: Frequency (50 kHz/Div)
Y-axis: Amplitude (0.1/Div)

Figure 8: A sample class 2 waveform and its power spectrum



(a) Time domain
X-axis: Time (102.4 μ s/Div)
Y-axis: Voltage (50 V/Div)



(b) Frequency domain
X-axis: Frequency (50 kHz/Div)
Y-axis: Amplitude (0.1/Div)

Figure 10: A sample "rejected" waveform and its power spectrum

3.2 Results Interpretation

There are two general types of fracture-related AE activities observed under cyclic loading conditions: crack tip opening during increasing load (upward load cycle) and crack face rubbing during decreasing load (downward cycle).

The observations of sample data analysis show that both class 1 and class 2 waveforms are more structured with a faster rise time at the beginning of the waveform. Thus, they are associated with the AE signals from the crack opening. Even though class 1 and class 2 waveforms represent characteristics of crack opening signals, more accurate characterization requires having access to AE signals that represent properties of steel used in the bridge, component dimensions, exposure conditions, etc. Class 3 waveforms are more slowly rising, and their spectral content is more inline with common background transient noise. The rejected waveforms are more likely due to saturation and clipping of the signal, and the non-saturated ones, centered around 50 kHz are more likely results from structural resonance.

Since class 1 and 2 waveform characteristics closely represented AE signals from the crack opening, the source location plots were analysed. However, during this particular implementation there were no active sources documented within the zone of interest. Therefore, further analysis was not performed. Yet, the data analysis procedures presented in this article demonstrated the possibility of identifying crack type signals from common background transient noise and structural resonance.

4.0 Summary, Conclusions, and Recommendations

Regardless of the causes of cracking, fatigue events (i.e., crack initiation or crack growth) need to be identified to and monitored to assure safety of structures. AE has been successfully implemented in the field and evaluated for continuous monitoring of fatigue-sensitive details.

This article described AE implementation for monitoring a fatigue sensitive detail (local monitoring) and use of cluster analysis, non-linear mapping (NLM), and three-class classifiers to identify the relationship of each cluster to the characteristics of crack opening signals, background noise, and structural resonance. The following are the specific conclusions and recommendations that are derived from this study:

- Cluster analysis and NLM of data can be performed in the time, power, phase, cepstral, and auto-correlation domains independently. The AE data collected from the bridge was examined directly to identify any significant similar AE activity formations. NLM with the spectral power domain produced three significant concentrations. Clustering was performed using the same spectral power domain features, and the clusters were well aligned with the visual presentation of the NLM result. NLM and cluster analysis demonstrated the usefulness of such techniques for understanding AE data.
- Waveform characteristics of pencil lead break (PLB) data and the AE data in three clusters were evaluated. Then, the dominant frequency ranges of each cluster were calculated. The results were used to identify the relationship of each cluster to the characteristics of crack opening signals, background noise, and structural resonance.
- Developing a fatigue cracking signal characteristic database of typical steel and welds used in a specific bridge, and using that database instead of PLB data would allow further refining of the AE data interpretation and more accurately detecting the critical events.
- The signal analysis process presented during this study, further refined with a database of fatigue cracking signals, can be integrated into remote monitoring to minimize receiving false alarms.

Acknowledgements

This project is funded by the Michigan Department of Transportation. The authors would like to acknowledge the support extended by Mr. Steve Kahl and the research advisory panel members for successful completion of the project. Further, the continuous technical support provided by Mr. Roger Chan, TISEC Inc., is greatly appreciated.

Research publication liability disclaimer

This publication is disseminated in the interest of information exchange. The Michigan Department of Transportation (hereinafter referred to as MDOT) expressly disclaims any liability, of any kind, or for any reason, that might otherwise arise out of any use of this publication or the information or data provided in the publication. MDOT further

disclaims any responsibility for typographical errors or accuracy of the information provided or contained within this information. MDOT makes no warranties or representations whatsoever regarding the quality, content, completeness, suitability, adequacy, sequence, accuracy or timeliness of the information and data provided, or that the contents represent standards, specifications, or regulations.

<http://www.tisec.com/shop/enter.html>, Visited, (Date)2010.

References

- [1] Federal Highway Administration (FHWA). Feasibility of Nondestructive Crack Detection and Monitoring for Steel Bridges, *Report No. FHWA-HRT-12-060*, Final Report, NDE Center, McLean, VA 22101, 2012.
- [2] Department for Transport (DFT). Design Manual for roads and bridges, *Report No. BA 86/06*, UK, 2006.
- [3] Nair, A., Cai, C.S. “Acoustic emission monitoring of bridges: Review and case studies”, *Engineering Structures*, Vol. 32, No. 6, pp. 1704-1714, 2010.
- [4] McCrea, A., Chamberlain, D., and Navon, R., “Automated inspection and restoration of steel bridges—a critical review of methods and enabling technologies”, *Journal of Automation in Construction*, Vol. 11, No. 4, pp. 351 – 373, 2002.
- [5] Kaphle, M., Tan, A. C. C., Thambiratnam, D. P., and Chan, T. H. T., “A Study on the Use of Acoustic Emission Technique as a Structural Health Monitoring Tool”, *Journal of Engineering Asset Management and Infrastructure Sustainability*, pp. 459 – 469, 2012.
- [6] Huang, M., Jiang, L., Liaw, P.K., Brooks, C.R., Seeley, R., and Klarstrom, D.L., "Using Acoustic Emission in Fatigue and Fracture Materials Research", *JOM-e*, Member Journal of TMS, Vol. 50, No. 11, 1998.
- [7] Shield, C., McKeefry, J. “Acoustic Emission Monitoring of Fatigue Cracks in Steel Bridge Girders”, *MN/RC – 1999-36*, Minnesota Department of Transportation, St. Paul, MN 55155, 1999.
- [8] AASHTO *LRFD* Bridge Design Specifications, Customary U.S. Units, 7th ed., American Association of State Highway and Transportation Officials, Washington, DC, 2014.
- [9] TISEC. “ICEPAK – Pattern Analysis Software.” TISEC Inc. 601 Route 364, Morin Heights, Quebec, Canada J0R 1H0.