

Fall 2012

## Acoustic Emission Signal Classification for Gearbox Failure Detection

Jun Shishino  
*Embry-Riddle Aeronautical University - Daytona Beach*

Follow this and additional works at: <https://commons.erau.edu/edt>



Part of the [Aerospace Engineering Commons](#), and the [Aviation Commons](#)

---

### Scholarly Commons Citation

Shishino, Jun, "Acoustic Emission Signal Classification for Gearbox Failure Detection" (2012).  
*Dissertations and Theses*. 132.  
<https://commons.erau.edu/edt/132>

This Thesis - Open Access is brought to you for free and open access by Scholarly Commons. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of Scholarly Commons. For more information, please contact [commons@erau.edu](mailto:commons@erau.edu).

# **ACOUSTIC EMISSION SIGNAL CLASSIFICATION FOR GEARBOX FAILURE DETECTION**

By

Jun Shishino

A Thesis Submitted to the Graduate Studies Office in  
Partial Fulfillment of the Requirements for the  
Degree of Master of Science in Aerospace Engineering

Embry-Riddle Aeronautical University

Daytona Beach, Florida

Fall 2012

# ACOUSTIC EMISSION SIGNAL CLASSIFICATION FOR GEARBOX FAILURE DETECTION

by

Jun Shishino

This thesis was prepared under the direction of the candidate's thesis committee chairman, Dr. Fady F. Barsoum, Department of Mechanical Engineering, and has been approved by the members of his thesis committee. It was submitted to the School of Graduate Studies and Research and was accepted in partial fulfillment of the requirements for the degree of Master of Science in Aerospace Engineering

THESIS COMMITTEE:



---

Dr. Fady F. Barsoum  
Chairman




---

Dr. Yi Zhao  
Member



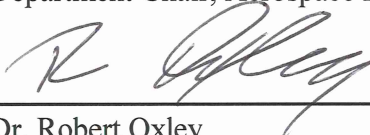
---

Dr. Daewon Kim  
Member



---

Dr. Tasos Lyrintzis  
Department Chair, Aerospace Engineering



---

Dr. Robert Oxley  
Associate Vice President for Academics

12/12/12

---

Date

12/13/12

---

Date

## **Acknowledgements**

Most importantly, I would like to express my deepest appreciation to my thesis adviser and committee chair, Dr. Fady Barsoum for his greatest guidance, encouragements and valuable advices throughout this research. Beyond the research experience, Dr. Barsoum has awarded me with many experiences that I would not have experienced if I had just worked in the lab without his supervision. Also, sharing his ideas about the world, he has taught me more than just being an engineer. Besides my adviser, I would also like to thank Dr. Zhao and Dr. Kim for being my thesis committee, for their guidance, encouragements, and advice for my thesis. Also, I would like to thank Dr. Hill for introducing me to acoustic emission.

I am grateful to my Malaysian brother, Prathikshen Selvadorai for not only encouraging me during the research but the ideas and advices that were given to me were appreciated. The long hours of discussion about AE analysis and the number of times that we went to Starbucks just to have a break will not be forgotten. Also, I thank Ning Leung for introducing me to this topic and the support. I also thank my fellow team members of the Structure Health Monitoring and Nondestructive Evaluation Lab: Rudy Baum and Kiwoong Nam, for their feedback on my presentations and the high motivation that they have brought into the lab. Additionally, I thank Audrey Gbaguidi for helping me out on the criteria codes.

I would like to thank my Japanese engineering friends from Embry-Riddle, Junya Iizuka, Takahiro Kuhara, Fernando Yuji Morita, Yu Ohashi, Kazuya Takabe, Hibiki Kobayshi and Mai Nakajo who has supported and encouraged me throughout the 6 years in Embry-Riddle. I would also like to thank So Ihara who gave me support and advice throughout the editing process of this thesis. Additionally, I would like to thank all my friends around the world that has supported me.

At last, I would like to greatly thank my family: Toshiaki, Keiko and Ken Shishino. Especially to my parents, Toshiaki and Keiko Shishino for what they have taught me, supported me, loved me and raised me. If it was not for them, it would have been impossible to reach to this point.

I also would like to acknowledge the support of the US Navy through STTR.

## **Abstract**

Author: Jun Shishino

Title: Acoustic Emission Signal Classification for Gearbox Failure Detection

Institution: Embry-Riddle Aeronautical University Daytona Beach Campus

Degree: Master of Science in Aerospace Engineering

Year: 2012

The purpose of this research is to develop a methodology and technique to determine the optimal number of clusters in acoustic emission (AE) data obtained from a ground test stand of a rotating H-60 helicopter tail gearbox by using mathematical algorithms and visual inspection. Signs of fatigue crack growth were observed from the AE signals acquired from the result of the optimal number of clusters in a data set. Previous researches have determined the number of clusters by visually inspecting the AE plots from number of iterations. This research is focused on finding the optimal number of clusters in the data set by using mathematical algorithms then using visual verification to confirm it.

The AE data were acquired from the ground test stand that simulates the tail end of an H-60 Seahawk at Naval Air Station in Patuxant River, Maryland. The data acquired were filtered to eliminate durations that were greater than 100,000  $\mu\text{s}$  and 0 energy hit data to investigate the failure mechanisms occurring on the output bevel gear. From the filtered data, different AE signal parameters were chosen to perform iterations to see which clustering algorithms and number of outputs is the best. The clustering algorithms utilized are the Kohonen Self-organizing Map (SOM), k-mean and Gaussian Mixture Model (GMM). From the clustering iterations, the three cluster criterion algorithms were performed to observe the suggested optimal number of cluster by the criterions. The three criterion algorithms utilized are the Davies-Bouldin, Silhouette and Tou Criterions. After the criterions had suggested the optimal number of cluster for each data set, visual verification by observing the AE plots and statistical analysis of each

cluster were performed. By observing the AE plots and the statistical analysis, the optimal number of cluster in the data set and effective clustering algorithms were determined. Along with the optimal number of clusters and effective clustering algorithm, the mechanisms of each cluster can be determined from the statistical analysis as well. From the results, the 5 cluster output using the Kohonen SOM clustering algorithm showed the distinct separation of clusters.

Using the determined number of clusters and the effective clustering algorithms, the AE data sets were analyzed for the fatigue crack growth. Recorded data from the mid test and end test of the data acquisition period were utilized. After each set of clusters were associated with different mechanisms dependent on their AE characteristics. It was possible to detect the increase in the activities of the fatigue crack data points. This indicates that the fatigue crack is growing as the acquisition continued on the H-60 Seahawk ground test stand and that AE has a good potential for early crack detection in gearbox components.

## Table of Contents

Acknowledgements.....	i
Abstract.....	ii
List of Figures.....	vii
List of Tables.....	x
List of Symbols and Abbreviations.....	xii
Chapter 1. Introduction.....	1
1.1. Research Objective.....	1
1.2. H-60 Helicopter.....	1
1.2.1. Basic Information.....	1
1.2.2. Fatigue Problem.....	2
1.3. Approach to the Problem.....	4
Chapter 2. Literature Review.....	5
2.1. Nondestructive Evaluation Techniques.....	5
2.2. Data Analysis Methods for AE Nondestructive Evaluation.....	5
2.3. Other Methods for Diagnosis and Prognosis of Rotating Machinery.....	7
Chapter 3. Nondestructive Evaluation and Acoustic Emission.....	8
3.1. Nondestructive Evaluation and Structural Health Monitoring.....	8
3.2. Acoustic Emission Nondestructive Testing.....	9
3.2.1. Acoustic Emission.....	9
3.2.2. Acoustic Emission Nondestructive Evaluation.....	10
3.2.3. AE Piezoelectric Transducers.....	11
3.2.4. AE Signal Parameters.....	11
3.2.5. Metal Failure Mechanisms and AE signals.....	13
Chapter 4. Data Acquisition.....	15
4.1. Acoustic Emission Signal Data Acquisition.....	15
4.2. Flow Diagram of the AE Signal Acquisition System.....	15
4.3. AE Signal Acquisition System Settings.....	16
4.4. Acquisition Software.....	17
4.4.1. AEWin™.....	17
Chapter 5. Data Clustering and Cluster Verification Methods.....	18
5.1. Data Clustering and Cluster Verification Analysis.....	18
5.2. Structural Pattern Recognition Analysis for Failure Mode Classification.....	19
5.2.1. Artificial Neural Networks and Kohonen Self-organizing Map.....	19

5.2.2.	K-mean Clustering Algorithm .....	22
5.2.3.	Gaussian Mixture Model Clustering Algorithm .....	23
5.3.	Cluster Criterion Algorithms for Cluster Verification .....	24
5.3.1.	Davies-Bouldin Criterion .....	25
5.3.2.	Silhouette Criterion .....	26
5.3.3.	Tou Criterion.....	26
5.4.	Analysis Software .....	27
5.4.1.	NeuralWorks® Professional II/Plus .....	27
5.4.2.	MATLAB® .....	28
Chapter 6.	Data Acquisition Setup of the Rotating Gear Box .....	29
6.1.	Ground Test .....	29
6.1.1.	Setup .....	29
6.1.2.	Transducer.....	30
6.2.	Parameters.....	32
6.2.1.	AE Parameters.....	32
6.2.2.	Loading Profile .....	33
Chapter 7.	Pre-analysis Data Processing.....	34
7.1.	Overview	34
7.2.	Elimination of Duration .....	34
7.3.	Elimination of Energy .....	35
7.4.	Elimination of Channels.....	36
Chapter 8.	Iterations and Inspection .....	38
8.1.	Analysis Iterations.....	38
8.2.	Cluster Inspection .....	38
8.2.1.	Criterion Verification of the Clusters .....	39
8.2.2.	AE plots and Statistical Analysis of Clusters .....	40
Chapter 9.	Results.....	42
9.1.	Clustering Criterion Verification of 2010 Data .....	42
9.1.1.	K-mean Clustering Criterion Verification .....	42
9.1.2.	Gaussian Mixture Model Clustering Criterion Verification .....	44
9.1.3.	Kohonen Self-organizing Map Clustering Criterion Verification.....	45
9.1.4.	Cluster Number Possibility from Clustering Criteria.....	46
9.2.	Visual Verification of 2010 Data .....	47
9.2.1.	K-mean Clustering .....	47



9.2.1.1.	September 7th K-mean.....	47
9.2.1.2.	October 18th K-mean.....	51
9.2.2.	Gaussian Mixture Model.....	56
9.2.2.1.	September 7th Gaussian Mixture Model.....	56
9.2.2.2.	October 18th Gaussian Mixture Model.....	60
9.2.3.	Kohonen SOM.....	65
9.2.3.1.	September 7th Kohonen SOM.....	65
9.2.3.2.	October 18th Kohonen SOM.....	70
9.3.	Cluster Number Possibility from Criteria and Visual Inspection of 2010 Data.....	75
9.4.	Fatigue Crack Growth Observation of 2010 Data.....	75
9.5.	Kohonen SOM Clustering Analysis of 2011 Data.....	77
Chapter 10.	Conclusions and Recommendation.....	79
10.1.	Conclusions.....	79
10.2.	Recommendations.....	81
Appendices	.....	A-1
Appendix A.	MATLAB Code for Filter and Create Input File: InputSeparation.m.....	A-1
Appendix B.	MATLAB Code for Kohonen SOM: nnr2excel.m.....	B-1
Appendix C.	MATLAB Code for Kohonen SOM: Rij_FromKSOMnnpro.m.....	C-1
Appendix D.	MATLAB Code for Kohonen SOM: Silhouette_FromKSOMnnpro.m.....	D-1
Appendix E.	MATLAB Code for Kohonen SOM: Tou_FromKSOMnnpro.m.....	E-1
Appendix F.	MATLAB Code for Kohonen SOM: vPlotKSOM.m.....	F-1
Appendix G.	MATLAB Code for K-mean: Kmeans.m.....	G-1
Appendix H.	MATLAB Code for K-mean: Rij_FromKmeanIDX.....	H-1
Appendix I.	MATLAB Code for K-mean: Silhouette_FromKmeanIDX.m.....	I-1
Appendix J.	MATLAB Code for K-mean: Tou_FromKmeanIDX.m.....	J-1
Appendix K.	MATLAB Code for K-mean: vPlotKmean.m.....	K-1
Appendix L.	MATLAB Code for GMM: GMM.m.....	L-1
Appendix M.	MATLAB Code for GMM: Rij_FromGMMIDX.m.....	M-1
Appendix N.	MATLAB Code for GMM: Silhouette_FromGMMIDX.m.....	N-1
Appendix O.	MATLAB Code for GMM: Tou_FromGMMIDX.m.....	O-1
Appendix P.	MATLAB Code for GMM: vPlotGMM.m.....	P-1
Appendix Q.	MATLAB Code for Visual Verification: AEPlot.m.....	Q-1
Appendix R.	MATLAB Code for Visual Verification: StatisticalAnalysisIDX.m.....	R-1
Appendix S.	MATLAB Code for Visual Verification: StatisticalAnalysisNNpro.m.....	S-1

## List of Figures

Figure 1 MH-60R Seahawk [2].....	2
Figure 2 (a) Tail of H-60 Seahawk [4] (b) Tail Gear Box [3] (c) Output Bevel Gear Test Specimen [3]....	2
Figure 3 X-Ray Image of Bevel Gear Spline Cracks From Fretting [5] .....	3
Figure 4 Crack Growth around the Gear Spline [6] .....	4
Figure 5 Sample Representation of the AE Testing Setup [19].....	10
Figure 6 Sample Representation of AE Piezoelectric Transducers [20].....	11
Figure 7 AE Signal Parameters [21] .....	12
Figure 8 Fatigue Crack Modes [25] .....	14
Figure 9 Flow diagram of AE Data Acquisition .....	16
Figure 10 Flow Diagram of the Analysis .....	19
Figure 11 Representation of (a) one dimensional (b) two dimensional Kohonen SOM neural network....	20
Figure 12 Representation of one dimensional Kohonen SOM neural network with 4 input neurons, 5x1 Kohonen layer and 2 output neurons.....	22
Figure 13 Flow Diagram of K-mean Clustering Algorithm .....	22
Figure 14 (a) Example of Gaussian Distribution (b) Multiple Gaussian Distribution from X-Y View .....	24
Figure 15 Flow Diagram of Gaussian Mixture Model Clustering Algorithm.....	24
Figure 16 3D Drawing of Ground Test Setup [3].....	30
Figure 17 Transducer Position for (a) 2010 [41] (b) 2011.....	32
Figure 18 Hub Moment Load and Rotational Speed of the Bevel Gear in (a) September 7, 2010 (b) October 18, 2010 .....	33
Figure 19 Sorting Method for AE Data.....	34
Figure 20 October 18th AE Hit Data with Duration vs. Amplitude (a) all signals (b) signal duration $\leq 100,000\mu\text{s}$ .....	35
Figure 21 Duration vs. Counts of October 18, 2010 (a) Channel 1 (b) Channel 2.....	36

Figure 22 Duration vs. Counts AE Signals from TGBOBG0002 of 2011 Acquisition for Different Channels .....	37
Figure 23 Cluster Inspection Flow .....	39
Figure 24 Example of Criterion Verification Plot .....	40
Figure 25 Flow of the Criterion Voting Plot .....	40
Figure 26 October 18, 2010 Duration vs. Count (a) Input: EAF Output 3 (b) Input: DEAF Output 4 .....	41
Figure 27 September 7th K-mean (Input: Duration, Energy, Amplitude, Avg Freq.) Cluster Criterion ....	43
Figure 28 October 18th K-mean (Input: Duration, Energy, Amplitude, Avg Freq.) Cluster Criterion .....	43
Figure 29 September 7th GMM (Input: Duration, Energy, Amplitude, Avg Freq.) Cluster Criterion.....	44
Figure 30 October 18th GMM (Input: Duration, Energy, Amplitude, Avg Freq.) Cluster Criterion .....	45
Figure 31 September 7th K-mean (Input: Duration, Energy, Amplitude, Avg Freq.) Cluster Criterion ....	45
Figure 32 October 18th K-mean (Input: Duration, Energy, Amplitude, Avg Freq.) Cluster Criterion .....	46
Figure 33 September 7th - 4 Clusters K-mean Clustering AE Plots.....	48
Figure 34 September 7 <sup>th</sup> - 5 Clusters K-mean Clustering AE Plots.....	49
Figure 35 September 7th - 6 Clusters K-mean Clustering AE Plots.....	50
Figure 36 October 18th - 4 Clusters K-mean Clustering AE Plots .....	52
Figure 37 October 18th - 5 Clusters K-mean Clustering AE Plots .....	53
Figure 38 October 18th - 6 Clusters K-mean Clustering AE Plots.....	55
Figure 39 September 7th - 4 Clusters GMM Clustering AE Plots.....	57
Figure 40 September 7th - 5 Clusters GMM Clustering AE Plots.....	58
Figure 41 September 7th - 6 Clusters GMM Clustering AE Plots.....	59
Figure 42 October 18th - 4 Clusters GMM Clustering AE Plots.....	61
Figure 43 October 18th - 5 Clusters GMM Clustering AE Plots.....	62
Figure 44 October 18th - 6 Clusters GMM Clustering AE Plots.....	64
Figure 45 September 7th – 4 Clusters Kohonen SOM Clustering AE Plots.....	66
Figure 46 September 7th – 5 Clusters Kohonen SOM Clustering AE Plots.....	67

Figure 47 September 7th – 6 Clusters Kohonen SOM Clustering AE Plots.....	69
Figure 48 October 18th – 4 Cluster Kohonen SOM Clustering AE Plots.....	70
Figure 49 October 18th – 5 Cluster Kohonen SOM Clustering AE Plots.....	72
Figure 50 October 18th – 6 Cluster Kohonen SOM Clustering AE Plots.....	74
Figure 51 September 7th – Kohonen SOM 5 Cluster AE plots .....	76
Figure 52 October 18th – Kohonen SOM 5 Cluster AE plots .....	77
Figure 53 TGBOBG0002 Kohonen SOM Clustering .....	78
Figure 54 Analysis Flow to Obtain Optimal Number of Cluster in a Data Set.....	80

## List of Tables

Table 1. Different NDE Technique and its Detection Objective [18].....	9
Table 2 AE Signal Parameter Description .....	13
Table 3 Timing Values Depending on Material [22].....	13
Table 4 Characteristics of AE Data and Failure Mechanisms in Metals [12].....	14
Table 5 AE System and Transducer Used.....	16
Table 6 PDT, HDT and HLT Settings for 2010 and 2011.....	17
Table 7 AE Signal Acquisition System Settings for 2010 and 2011 .....	17
Table 8 Transducer and Channels for each Test [38] [39] [40].....	31
Table 9 Recorded AE Hit Signal Parameters .....	32
Table 10 AE Data with Zero Energy.....	35
Table 11 AE Data with Energy Greater than Zero .....	35
Table 12 AE Input Parameters and Output Iterations.....	38
Table 13 Criterion Peaks of Each Date using Different Cluster Algorithms for 4th Iteration .....	46
Table 14 September 7th – 4 Clusters K-mean Statistical Analysis (a)Amplitude (b)Duration (c)Energy .	48
Table 15 September 7th – 5 Clusters K-mean Statistical Analysis (a)Amplitude (b)Duration (c)Energy .	49
Table 16 September 7th – 6 Clusters K-mean Statistical Analysis (a)Amplitude (b)Duration (c)Energy .	50
Table 17 October 18th – 4 Clusters K-mean Statistical Analysis (a)Amplitude (b)Duration (c)Energy ...	52
Table 18 October 18th – 5 Clusters K-mean Statistical Analysis (a)Amplitude (b)Duration (c)Energy ...	54
Table 19 October 18th – 6 Clusters K-mean Statistical Analysis (a)Amplitude (b)Duration (c)Energy ...	56
Table 20 September 7th – 4 Clusters GMM Statistical Analysis (a)Amplitude (b)Duration (c)Energy ....	57
Table 21 September 7th – 5 Clusters GMM Statistical Analysis (a)Amplitude (b)Duration (c)Energy ....	58
Table 22 September 7th – 6 Clusters GMM Statistical Analysis (a)Amplitude (b)Duration (c)Energy ....	60
Table 23 October 18th – 4 Clusters GMM Statistical Analysis (a)Amplitude (b)Duration (c)Energy .....	61
Table 24 October 18th – 5 Clusters GMM Statistical Analysis (a)Amplitude (b)Duration (c)Energy .....	63

Table 25	October 18th – 6 Clusters GMM Statistical Analysis (a)Amplitude (b)Duration (c)Energy .....	65
Table 26	September 7th – 4 Clusters KSOM Statistical Analysis (a)Amplitude (b)Duration (c)Energy...	66
Table 27	September 7th – 5 Clusters KSOM Statistical Analysis (a)Amplitude (b)Duration (c)Energy...	68
Table 28	September 7th – 6 Clusters KSOM Statistical Analysis (a)Amplitude (b)Duration (c)Energy...	69
Table 29	October 18th – 4 Clusters KSOM Statistical Analysis (a)Amplitude (b)Duration (c)Energy .....	71
Table 30	October 18th – 5 Clusters KSOM Statistical Analysis (a)Amplitude (b)Duration (c)Energy .....	72
Table 31	October 18th – 6 Clusters KSOM Statistical Analysis (a)Amplitude (b)Duration (c)Energy .....	74
Table 32	Number of Clusters by Criterion and Visual Inspection.....	75

## List of Symbols and Abbreviations

TGB	Tail Gearbox
OBG	Output Bevel Gear
AE	Acoustic Emission
SOM	Self-organizing Map
GMM	Gaussian Mixture Model
NDE	Nondestructive Evaluation
SHM	Structure Health Monitoring
ANN	Artificial Neural Network
PDT	Peak Definition Time
HDT	Hit Definition Time
HLT	Hit Lockout Time
PAC	Physical Acoustic Corporation
PE	Processing Element
MHD	Multiple Hit Data

# **Chapter 1. Introduction**

## **1.1. Research Objective**

Based on the few problems reported from the H-60 helicopter community on its tail gearbox (TGB) output bevel gear (OBG), the purpose of this research is to develop a methodology and technique to determine the optimal number of clusters in acoustic emission (AE) data obtained from a ground test stand of a rotating H-60 Seahawk helicopter tail gearbox. Results of the optimal number of cluster in a data set using mathematical algorithms and visual inspection were used to observe the signs of fatigue crack growth from the AE signals acquired. Also, this research will observe the relationship between the AE hit signal characteristics to the failure mechanisms. Previous researches have determined the number of clusters by visually inspecting the AE plots from multiple iterations. This research is focused on finding the optimal number of cluster in the data set by using mathematical algorithms and then using visual verification to confirm it to be used for further analysis. Success of these tasks will establish additional understanding of the diagnosis and prognosis of rotating machinery using multiple AE hit signal parameters and pattern recognition algorithms.

## **1.2. H-60 Helicopter**

### **1.2.1. Basic Information**

The H-60 helicopter is a four bladed twin engine utility helicopter manufactured by the Sikorsky Aircraft Corporation. Introduced to the United States Army in the mid 1970's as the Blackhawk, it has been the one of the most important rotary aircraft in its fleet. Beginning with the UH-60A Blackhawk, many different models have evolved to serve its needs as a multi-mission support aircraft for air assault, border security, command and control, drug interdiction, search and rescue and medical evacuation [1]. For example, the SH-60 Seahawk operated by the US Navy has evolved and redesigned to serve the needs of the Navy. The H-60 helicopter is operated internationally, and in the case of the Blackhawk, over 26 governments and militaries are operating it around the world.

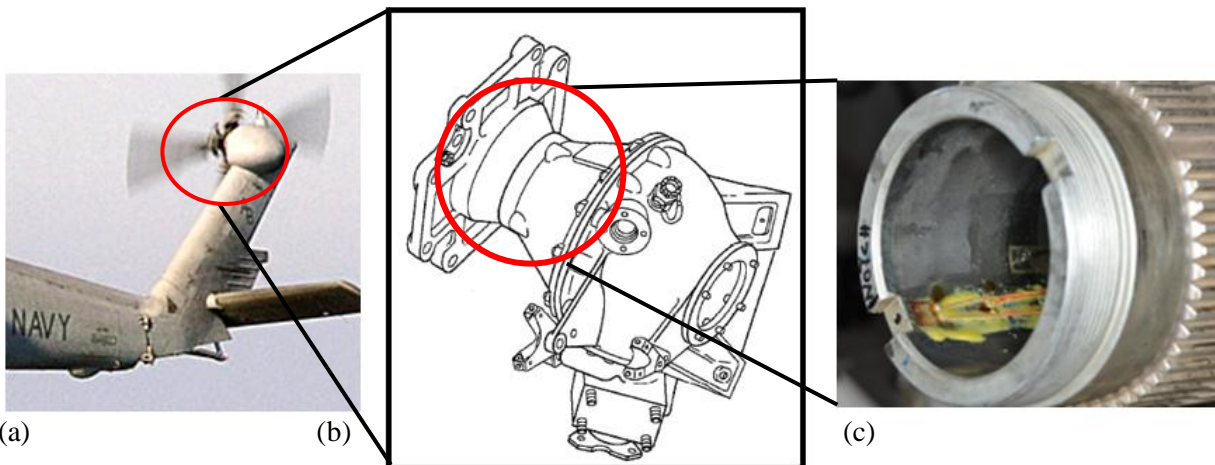




**Figure 1 MH-60R Seahawk [2]**

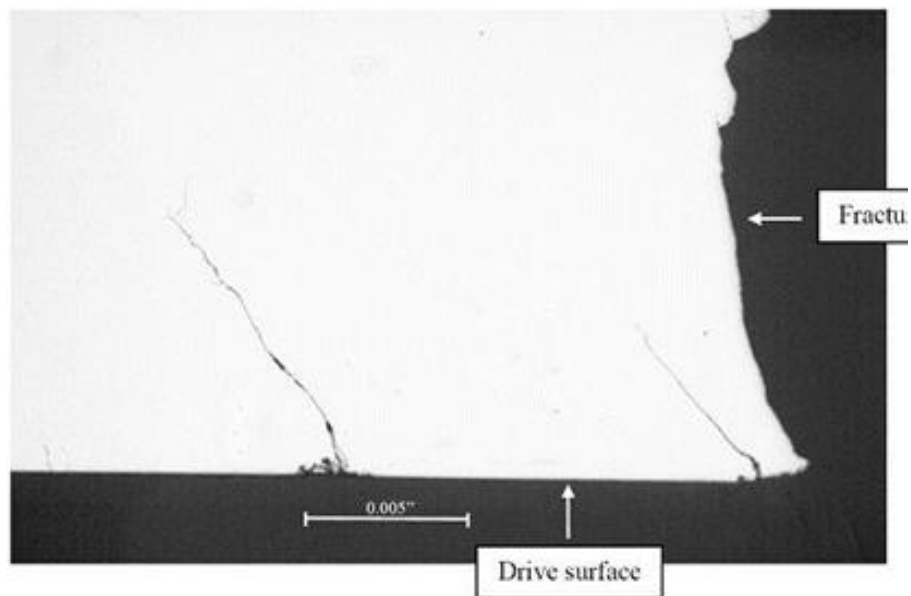
### 1.2.2. Fatigue Problem

Five crack incidents of the TGB OBG have been reported from the H-60 helicopter community since 2008. Of these incidents, one was from a United States Army Blackhawk, one from the Japan Maritime Self-Defense Force, and three failures from the United States Navy Seahawks [3]. Figure 2(a) shows the tail end of the H-60 Seahawk and the Figure 2(b) shows the TGB. In the TGB, there is an OBG inside similar to Figure 2(c) where it is circled in Figure 2(b).



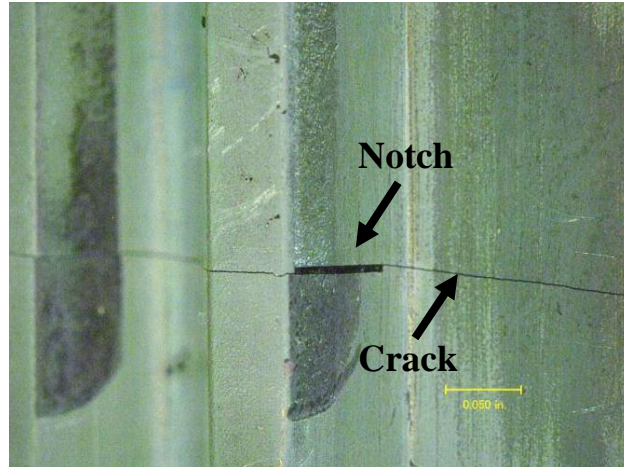
**Figure 2 (a) Tail of H-60 Seahawk [4] (b) Tail Gear Box [3] (c) Output Bevel Gear Test Specimen [3]**

These incidents have created an interest in investigating the cause of the failures as well as diagnosis and prognosis methods for the failed components. Complete root cause is still under investigation, but it has been recognized that fretting wear takes the protective coating layer of silver-plate off from the splines. After the protective layer is gone, bare metal to metal rubbing continues and triggers galvanic corrosion. When a galvanic couple forms, one of the metals in the couple becomes the anode and corrodes quicker than on its own, while the other becomes the cathode and corrodes more slowly. This provides conditions for crack initiation to the splines of the OBG. Figure 3 shows an x-ray image of OBG spline cracks from fretting.



**Figure 3 X-Ray Image of Bevel Gear Spline Cracks From Fretting [5]**

In addition to the fretting wear and metal to metal rubbing, the bending hub moment occurring on the OBG is another potential initiation of the fatigue crack. This once-per-revolution cyclic loading causes the crack to grow circumferentially around the OBG splines as it is shown in Figure 4.



**Figure 4 Crack Growth around the Gear Spline [6]**

### **1.3. Approach to the Problem**

Two acoustic emission data acquisitions were performed on a ground test stand at Naval Air Station in Patuxant River, Maryland in 2010 and 2011. Different mathematical clustering algorithms are utilized to determine the significant difference in clustering the acoustic emission data for further analysis of the rotating machinery. These mathematical models are Kohonen Self-organizing map (SOM), k-mean, and Gaussian Mixture Models (GMM). To determine the efficiency and accuracy of the classification methods, three cluster verification methods are performed. The three verification methods are Silhouette, Davies-Bouldin, and Tou criterion. From the result of the criterion verification, the optimal number of clusters in the data set can be suggested. From the suggested number of clusters, visual verification is performed by observing AE plots and statistical analysis of the each set of clusters. The data sets are clustered again with the result from the visual verification of optimal number of clusters for the metal failure diagnosis of rotary machinery component. Moreover, the research will be further extended in the future for real-time structure health monitoring of the helicopter rotary machinery.

## **Chapter 2. Literature Review**

### **2.1. Nondestructive Evaluation Techniques**

Nondestructive evaluation (NDE) and structural health monitoring (SHM) systems are of great interest for aging structures, as well as new innovative structures. Liu et al. [7] has indicated some perspectives in the research area of the SHM. Much research on analysis techniques such as statistical pattern recognition and vibration analysis has been used to detect, quantify and localize the damage. In addition to the analysis techniques, research on sensing techniques such as using piezoelectric transducers, optic fiber Bragg gratings, and accelerometers is in progress. Combinations of these advanced techniques have influenced the improvement of the evaluation techniques and the monitoring systems. Liu et al. states that the most useful SHM techniques developed for Department of Defense and aerospace applications are vibration-based approaches and guided wave propagation-based approaches. Montalvão et al. [8] has created a review on the vibration based SHM on composite materials. Belsak et al. [9] monitored vibration to detect the cracks in the tooth root of gears. They have conducted different methods of time signals analysis for the recorded signals. They have concluded that the time-frequency analysis, especially the typical spectrogram pattern of a short-time-frequency analysis, was capable of fault detection compared to the classical frequency analysis under low level of noise condition. Another approach that Liu et al. indicated is the guided wave propagation to inspect structures for flaws. They state that current diagnosis techniques require initial baseline on undamaged information of the structure so the damaged information can be compared to estimate the damage state of the structure. However, Liu et al. also states that since the reference signals will not be always available, development of techniques for the unsupervised damage detection are one of the important area of study for NDE and SHM.

### **2.2. Data Analysis Methods for AE Nondestructive Evaluation**

It has been shown that AE NDE used with pattern recognition algorithms and statistical analysis is capable of classifying different failure mechanisms in various materials under loading. Mahamed et al. [10] states that AE and vibration technique to detect the bearing failure are similar except for the

detection frequency. Since AE is capable of making detections simply and more quickly, Mahamed et al has used AE and artificial neural network (ANN) for diagnosis and prognosis of bearing failure in rotating machinery. Yella et al. [11] has researched different types of pattern recognition algorithms. They have used support vector machines, learning vector quantization, multi-layer perception, radial basis function networks, and the Gaussian mixture models to observe difference in the clustering rate. Acquiring the data from wooden beams, Yella et al. has concluded that the Gaussian mixture model achieved the highest classification rate. This was followed by support vector machines and learning vector quantization.

In 1998, Vaughn [12] demonstrated in-flight fatigue crack monitoring using AE and was able to classify different failure mechanisms on Piper PA-28 engine cowling by using SOM. Also in 1998, Rovik [13] examined fatigue crack on the vertical tail section of a Cessna T-303 Crusader aircraft. Rovik acquired the AE data from a controlled lab environment to train the SOM and used AE signals from a flying test bed to test the trained SOM. Rovik was able to successfully classify the AE signal data from the flying test bed into fatigue cracks to develop a fatigue crack monitoring system. Other than using the Kohonen SOM, there are other pattern recognition algorithms that can be used to classify the failure mechanisms. Meriaux et al [14] used k-mean to cluster the AE signals from fretting fatigue test. From the analysis, three different steps in the crack propagation were observed.

While classifying the failure mechanisms using a pattern recognition algorithm, it is necessary to understand and correlate the AE signals with the clusters that the algorithms have created. Sause et al. [15] has correlated AE signal types to microscopic failure mechanisms on fiber reinforced plastics. This was done by correlating fracture surface microstructure of different specimens with frequency of occurrence of distinct signal types, and by comparing experimental signals and signals based from finite element simulations. From these comparisons, the feature values of three different failure mechanisms coincided well with the simulated signals and experimental signals showing the applicability of AE signal classification methods for failure analysis of fiber reinforced plastics. Vaughn [12] has demonstrated the correlation between AE signals and clusters from pattern recognition algorithms on metal specimen.

Vaughn has acquired acoustic emission signals from the metal fatigue test specimen in the lab experiment which were then analyzed by performing Kohonen SOM. While the test was observed in real-time on duration vs. amplitude plot, Vaughn was able to observe each of the failure mechanisms occurring at different times and strength of amplitude, showing that failure mechanisms during the fatigue failure have different AE characteristics that can be applied for the clustering of the pattern recognition algorithms.

### **2.3. Other Methods for Diagnosis and Prognosis of Rotating Machinery**

Other than using nondestructive evaluation, there have been many studies on methods that can be used for the prognosis of rotating machinery. The numerical methods using finite element analysis are the most common method performed. Glodež et al. [16] has performed finite element analysis to determine the service life of a gear tooth under bending fatigue. Strain-life method in finite element method was used to determine the number of stress cycles required for fatigue crack initiation, and simple Paris equation was used for further simulation of the fatigue crack growth. By comparing the number of loading cycles with the computational analysis result to the experimental result, it has shown that the computational result correspond well with the experimental data. Flasker et al. [17] observed the effect of contact area to the direction of crack propagation and service life of gears with crack in a tooth root while applying different load conditions numerically and experimentally. From the experiment, the results showed that the speed and direction of the crack growth and the service life are influenced by contact are of the gear tooth by direct comparison with experiment and the finite element model.

## **Chapter 3. Nondestructive Evaluation and Acoustic Emission**

### **3.1. Nondestructive Evaluation and Structural Health Monitoring**

NDE is a method that can be used for investigating a part, material or system without damaging it for future usage. Using different NDE techniques, it is possible to detect variations in structures, minute changes in surface finish, the presence of crack or discontinuities, and to determine material/structure property/characteristics [18]. The main reasons for many manufacturers to use the nondestructive evaluations are to ensure product integrity, to avoid failure, and to control manufacturing process. All of these reasons are related to lowering the manufacturing cost and the safety of the product. Also by using some of the NDE methods such as acoustic emission, ultrasonic, and vibration, real-time structural health monitoring can be performed. By observing the structure in real-time, it is possible to detect change in the structure before severe failure. This real-time structural health monitoring is performed on many large and safety dependent structures such as bridges, buildings, and gas tanks.

The National Materials Advisory Board (NMAB) Ad Hoc Committee on Nondestructive Evaluation has classified the NDE techniques into six major categories: visual, penetrating radiation, magnetic-electrical, mechanical vibration, and thermal and chemical [18]. Table 1 shows different NDE techniques depending on the NMAB categories. Types of NDE techniques can be chosen depending on the detection objective, test environment and the budget.

**Table 1. Different NDE Technique and its Detection Objective [18]**

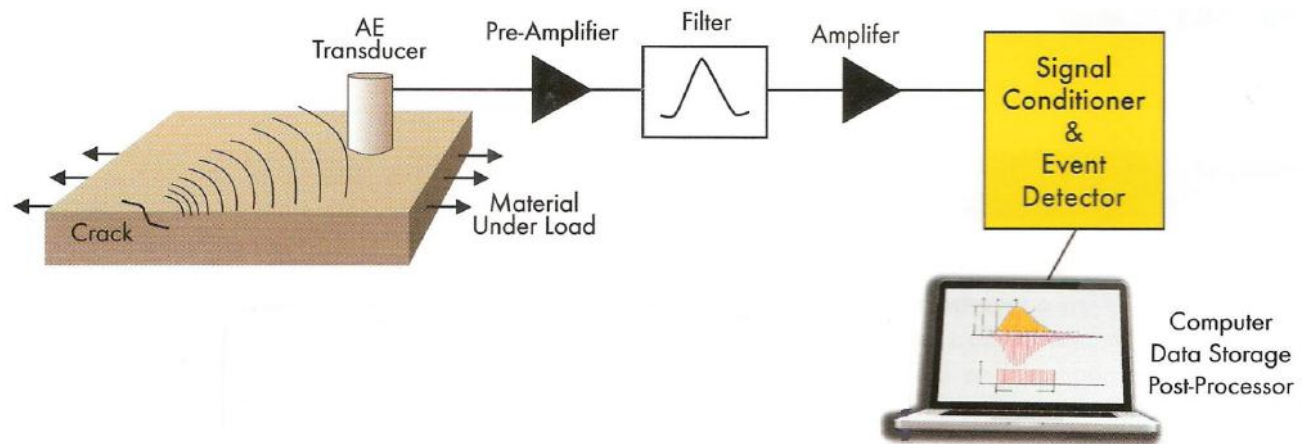
<b>NDE Technique</b>	<b>NMAB Category</b>	<b>Detection Objective</b>
<b>Visual</b>	Visual	Surface discontinuities, cracks, misalignments, warping, corrosion
<b>Liquid Penetrant</b>	Visual	Surface discontinuities, seams, cracks, laps, porosity, leak paths
<b>Magnetic Particle</b>	Magnetic-electrical	Surface discontinuities, seams, cracks, laps
<b>Eddy Current</b>	Magnetic-electrical	Cracks, dents, holes, grain size and hardness, coating, material thickness, dimensions, composition, conductivity
<b>Radiography</b>	Penetrating Radiation	Inclusions, lack of fusion, cracks, corrosion, porosity, leak paths, debris
<b>Acoustic Emission</b>	Mechanical Vibration	Internal discontinuities, material property, leak detection, chemical reaction, phase transformation
<b>Ultrasonic</b>	Mechanical Vibration	Surface discontinuities, internal discontinuities, material property
<b>Leak Testing</b>	Thermal and Chemical	Cracks, discontinuities, leak paths
<b>Infrared and Thermal</b>	Thermal and Chemical	Fluid flow, disconnects, discontinuities, heat transfer characteristic

### **3.2. Acoustic Emission Nondestructive Testing**

#### **3.2.1. Acoustic Emission**

Acoustic emission is elastic wave generated by the rapid energy release from localized source of stress due to failure in a structure under loading. When a structure is under loading, the material will begin to fail on both microscopic and macroscopic levels, thus generating elastic wave. This wave will propagate through the structure and are detected at the surface using AE piezoelectric transducers. These transducers convert the captured elastic wave into a voltage versus time electrical signal that can be recorded by AE data acquisition system after the signals are amplified and filtered. Figure 5 shows a sample representation of the setup. The analysis performed after the data acquisition can determine the location of the failure, failure mechanism classification, and condition monitoring of the structure.





**Figure 5 Sample Representation of the AE Testing Setup [19]**

### **3.2.2. Acoustic Emission Nondestructive Evaluation**

AE NDE is one of the passive volumetric NDE techniques that are capable of performing pressure testing, diagnostics, condition monitoring, and leak detection through detection and analysis of AE signals. AE testing and modern technologies allow high speed large scale monitoring of the material and structures. In a typical AE test, equal or slightly greater controlled mechanical loads compared to the practiced application load are applied to simulate the AE signals from the test sample. From these tests, an AE system can be carefully instrumented as well as having a better understanding of the relationship between the AE signals and material/structure mechanisms. The major advantage of AE NDE is the capability of passive volumetric monitoring, which allows it to be used to obtain AE data from the whole test objects during its regular service. Utilizing multiple sensors can distinguish the location of the different sources of the emission waves. Another advantage of AE NDE is the high sensitivity to detect changes in structures, using the high frequency transducers. Despite having various advantages, there are also disadvantages of AE NDE as well. Due to its high sensitivity, noises, electromagnetic interference, and other environment noise sources are acquired during the signal acquisition as AE signals. These noise signals are not related to the structure deformation and must be removed for the further analysis. Also, since AE is an irreversible process, it will not generate more AE signals unless the discontinuity grows or the structure is applied with higher loads.

### 3.2.3. AE Piezoelectric Transducers

AE Piezoelectric transducers are sensors that convert elastic energy wave to voltage to be recorded as electrical signal. The transducers usually have a ceramic plate that is coupled to the test structure for dynamic surface motion to propagate into the piezoelectric element. This piezoelectric element is connected to the AE acquisition system by electrical lead, and the piezoelectric element is covered by damping material to block the AE signal reflection to the active element. All components are covered by the casing and protected. Some transducers have integrated preamplifier inside the transducers. This is to reduce the amount of noise coming in from the surrounding environment to the electrical lead while traveling undistorted AE signal through long distances. A sample representation of AE piezoelectric transducer can be seen as Figure 6. Typical operating frequency of the transducer is in the range of 30 kHz to 1MHz. Depending on the usage environment and operation requirements, different types of AE piezoelectric transducers are available from many manufactures.

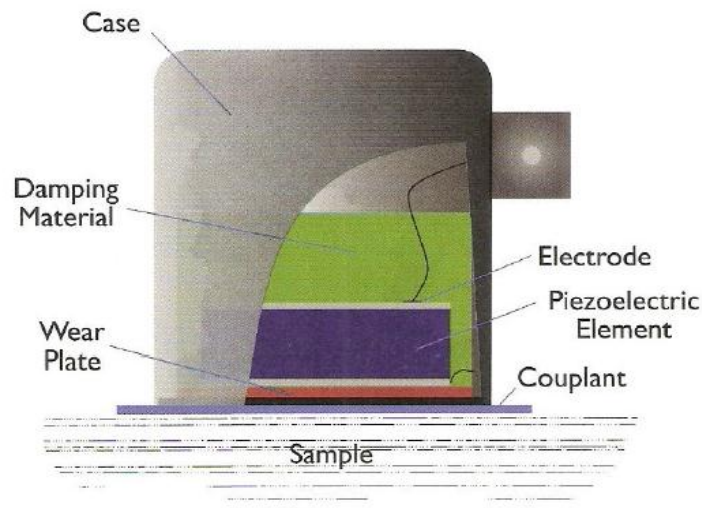
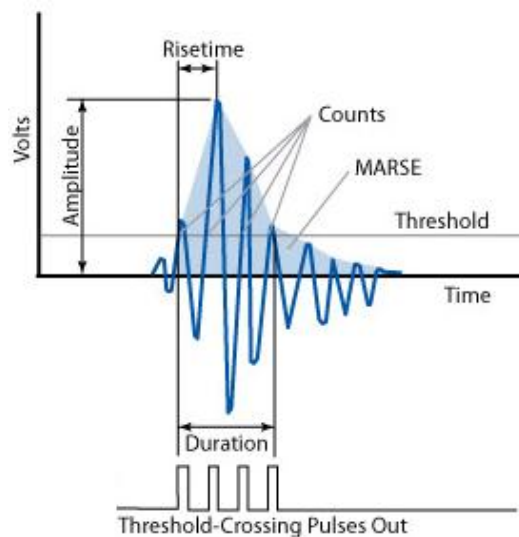


Figure 6 Sample Representation of AE Piezoelectric Transducers [20]

### 3.2.4. AE Signal Parameters

The threshold value is set to record the AE signals coming into the transducers. The signals that have amplitude greater than the threshold value are recognized as an AE hit and recorded by the AE data acquisition system. From the recorded AE hits, 6 different parameters are quantified: amplitude, rise time,

duration, counts, counts to peak, and measured area under the reflected signal envelope (MARSE), or energy. AE signal parameter representation can be seen in Figure 7 and the description in Table 2. When the AE acquisition system records the AE hits, it needs to set the timing values; which control how AE hit events are characterize. Different timing values include peak definition time (PDT), hit definition time (HDT), and hit lockout time (HLT). As seen in Table 3, the material and environment of the test object cause these timing values to differ. The PDT is time after the first crossing of the threshold to the highest peak of the signal. It is needed to avoid false measurements of high velocity and low amplitude signals. This time should be set as short as possible in order to discriminate between separate events, but long enough to capture the true peak of an event wave form. The HDT is time after the signal has fallen below the threshold. It is a measure of the time between events. The HDT is used to determine the end of an event to store the parameters of that event for later use. It must be long enough to record the event as well as low enough to separate individual events. The HLT is the time between the past signal to the new signal which closes out the measurement process and stores the AE parameters from the waveform. The HLT is set to prevent the measurement of reflected or late arriving parts of the AE signal to eliminate false event echoes and so that the data acquisition rate can be optimized. The values of HDT and HLT parameters are selected from the pencil lead break tests from different locations in the specimen.



**Figure 7 AE Signal Parameters [21]**

**Table 2 AE Signal Parameter Description**

<b>AE Parameter</b>	<b>Description</b>	<b>Units</b>
Amplitude	Strength of the maximum signal voltage	Decibel (dB)
Rise Time	Time elapsed from initial threshold crossing to peak amplitude	Microsecond ( $\mu$ s)
Duration	Time elapsed between initial to final threshold crossing	Microsecond ( $\mu$ s)
Counts	Number of times signal positively crosses the threshold	# of hits
Counts to Peak	Number of counts until signal reaches peak amplitude	# of hits
Energy	Also called as Measured Area Under the Reflected Signal Envelope (MARSE)	Energy Counts

**Table 3 Timing Values Depending on Material [22]**

<b>Material</b>	<b>PDT (<math>\mu</math>s)</b>	<b>HDT (<math>\mu</math>s)</b>	<b>HLT (<math>\mu</math>s)</b>
Composites, Non-Metals	20-50	100-200	300
Small Metal Specimens	300	600	1,000
Metal Structures (high damping)	300	600	1,000
Metal Structures (low damping)	1,000	2,000	20,000

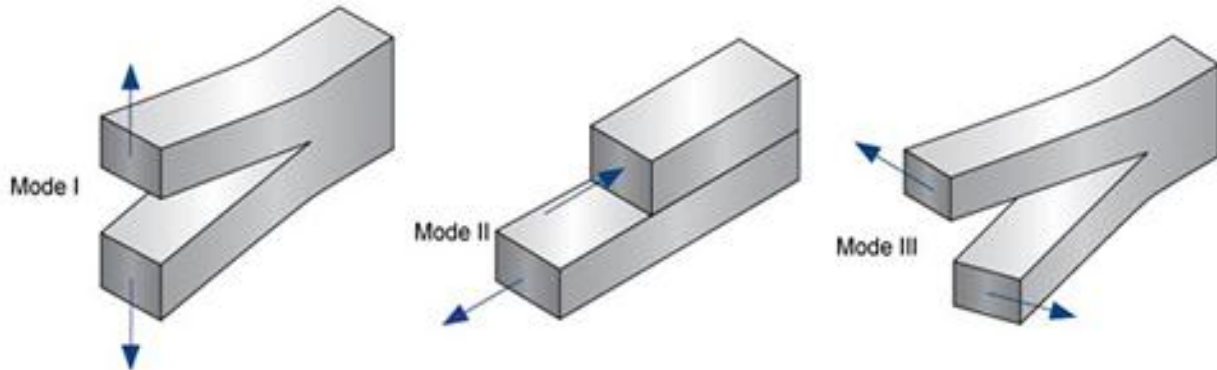
### 3.2.5. Metal Failure Mechanisms and AE signals

AE signals can originate from different sources. It is important to know the different failure mechanisms that a metal structure undergoes while it proceeds to the stage of catastrophic failure. The common failure mechanisms in metal structures are plastic deformation and fatigue cracking.

Plastic deformation occurs when a certain location undergoes some loading which is higher than the elastic load. Sufficient stress resulting in incapability for the structure to elastically deform back to its original shape causes plastic deformation. When a certain location undergoes strain hardening after multiple plastic deformations due to a microscopic crack or small imperfection, crack initiation occurs at that location. While the loads are constantly applied (cyclic or uniaxial loading), the crack will grow until the amount of undamaged material is insufficient to carry the load. The structure then undergoes catastrophic failure.

Three basic types of stress can be defined for crack-tip stress analysis as shown in Figure 8. Plane strain or Mode I is the most commonly encountered mode in fatigue failure. Plane strain or Mode I tears open the crack and continue to grow the crack. Mode II is developed while crack surface slide over each other

in perpendicular direction to the edge of the crack. While the crack grows, plane stress or Mode III that characterizes crack surfaces sliding respect to each other in parallel direction to the edge of the crack continues to grow the crack. Superposition of these three modes describes the general case of the local crack tip deformation. Addition to the three modes, plastic deformation occurs at the tip and the process returns back to the modes [23] [24].



**Figure 8 Fatigue Crack Modes [25]**

Table 4 shows the characteristics of AE signals related to failure mechanisms of metal materials. By observing the amplitude, duration, and energy of the cluster created by the clustering algorithms, different failure mechanisms can be determined. Other than the failure mechanisms, noises must be categorized even though the noise itself does not involve in the failure of the material. The two major types of noise are electric noise (environmental noises) and mechanical noise (friction or fretting noises). Most of the time, frequency of these noise can be utilized to distinguish between the different types of noise.

**Table 4 Characteristics of AE Data and Failure Mechanisms in Metals [12]**

<b>Mechanism</b>	<b>Amplitude</b>	<b>Duration</b>	<b>Energy</b>
<b>Fatigue Cracking</b>	High	Medium	High
<b>Plastic Deformation</b>	Low	Short	Low
<b>Noises</b>	Low-Medium	Long	Medium

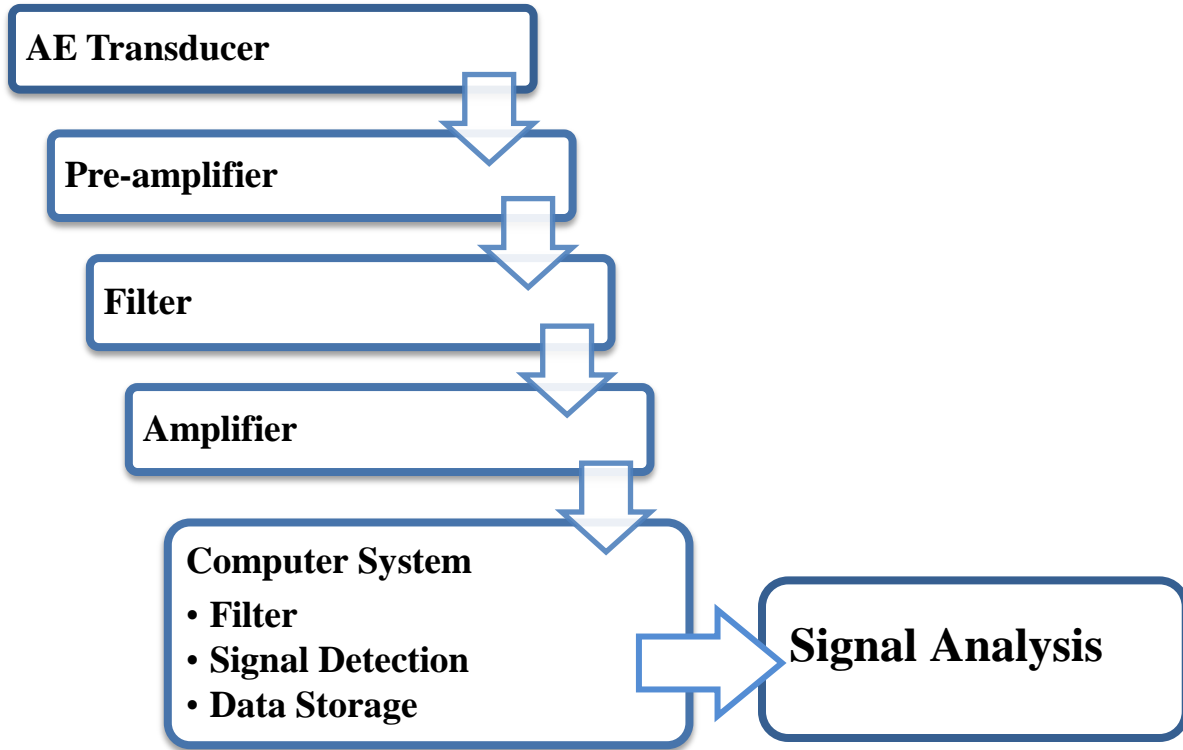
## **Chapter 4. Data Acquisition**

### **4.1. Acoustic Emission Signal Data Acquisition**

Two tests were conducted. Dr. Fady F. Barsoum, professor of Mechanical Engineering Department at Embry-Riddle Aeronautical University in Daytona Beach, Florida acquired AE data during the first test in 2010. Dr. Barsoum had conducted the data acquisition while serving as a faculty research fellow at US Navy Naval Air Station Patuxent River, Maryland. The AE data recording was conducted on the side while the test was performed for its original test plan. The second acquisition was conducted by the engineers at US Navy Naval Air Station Patuxant River in 2011.

### **4.2. Flow Diagram of the AE Signal Acquisition System**

Figure 9 shows the flow diagram of AE data acquisition. After the elastic wave acoustic emission is captured by the transducers, the emission is converted to a voltage signal. This voltage signal is better to be pre-amplified before traveling the signal for long distances. The pre-amplifier is recommended to be as near as possible to the transducers since it will reduce the signal interference from the environmental noises. After the AE signal is amplified, it is filtered to eliminate unwanted signals. Types of filters that can be applied are low pass, high pass, and band-pass filters. After the filtering, the signal can be amplified again to strengthen the desired AE signals. The signals are then sent to computer system to be filtered again and detect the AE hit signal parameters. After the detection, the AE hit data are stored to be analyzed.



**Figure 9 Flow diagram of AE Data Acquisition**

### 4.3. AE Signal Acquisition System Settings

In 2010, a 4 channel Physical Acoustic Corporation (PAC)  $\mu$ DISP/NB-8 AE system was connected to two PAC R15I AE transducers through a BNC connector cable. The AE hit data was then recorded and analyzed using the AEwin™ software. In 2011, the PAC PCI-2 system was connected to two PAC WD AE transducers and two PAC Micro30S AE transducers. Two different transducer types were used to observe the difference in the signal acceptance between the transducers.

**Table 5 AE System and Transducer Used**

Acquisition Year	AE System	Transducers
2010 (Channel 1-2)	PAC $\mu$ DISP/NB-8	2 PAC R15I Transducer
2011 (Channel 1-4)	PAC PCI-2	2 PAC WD and 2 PAC Micro30S Transducers

Data recording setting of the acquisition system must to be set to acquire the correct and enough AE signal parameters from the source. Table 6 and 7 shows acquisition system setting parameters.

**Table 6 PDT, HDT and HLT Settings for 2010 and 2011**

<b>PDT (<math>\mu</math>s)</b>	<b>HDT (<math>\mu</math>s)</b>	<b>HLT (<math>\mu</math>s)</b>
200	800	1000

**Table 7 AE Signal Acquisition System Settings for 2010 and 2011**

<b>Acquisition Year</b>	<b>Threshold (dB)</b>	<b>Pre-amplifier (dB)</b>	<b>Lower (kHz)</b>	<b>Upper (MHz)</b>	<b>Sample Rate (MSPS)</b>
<b>2010 (Channel 1-2)</b>	45	40	10	2	10
<b>2011 (Channel 1-4)</b>	65	40	100	1	2

In both data acquisitions, filtering was not performed at the acquisition stage. This was to understand how AE signals can be captured from the rotating ground test stand. Therefore, the entire AE signal was acquired during the 2010 and 2011 test. From the experiences obtained, an AE acquisition system parameter setting as well as equipment can be planned for the further AE signal acquisition tests.

#### **4.4. Acquisition Software**

##### **4.4.1. AEwin™**

AEwin™ [26] is a 32bit Windows program that is capable of controlling PAC's AE systems. Using the PAC's standard data files, DTA files, the AEwin™ allows users to not only record the AE data, but to replay and analyze the AE files that have been collected previously. Using AEwin™, location analysis, waveform analysis, and AE activity analysis can be performed. For this research, AEwin™ was used to acquire and record the data as well as to convert the AE signal parameters into ASCII data file that can be used with other software. Having the acquisition, graphing, and analysis capabilities imbedded in an AE system, along with many features to ease data analysis and visualization tasks, AEwin™ is one of the important software packages used for this research.

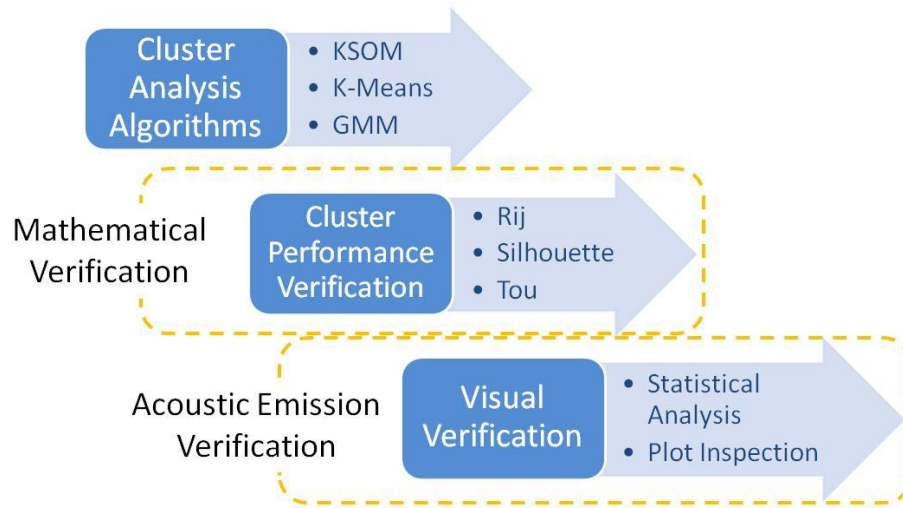


## **Chapter 5. Data Clustering and Cluster Verification Methods**

### **5.1. Data Clustering and Cluster Verification Analysis**

AE signals can originate from different failure sources in the structure. To observe the change in a structure and to determine which signal parameters are from specific mechanisms, correct analysis technique must be performed. The wave analysis was popular for AE analysis due to its simple observation. However, with the improvement of hardware and technique to observe AE signals in more detail, mathematical pattern recognition analysis is more popular today.

In this research, there are three main steps in determining the optimal number of clusters in an AE signal data set obtained. The flow of the analysis is shown in Figure 10. First, the cluster analysis is performed by using three pattern recognition algorithms. The three pattern recognition algorithms are Kohonen SOM, k-mean, and GMM. These three are unsupervised clustering techniques that are capable of partitioning a data point from the input of number of cluster that the user specifies. Second, from the cluster analysis, three internal criteria are measured to observe how separated the clusters are. The three internal criteria are Davies-Bouldin Criterion also known as  $R_{ij}$ , Silhouette Criterion and Tou Criterion. These criteria measures the properties of compactness of the clusters and how well separated groups are by outputting indices. At the end, AE verification is performed by statistical analysis for each cluster as well as inspecting the AE plots.



**Figure 10 Flow Diagram of the Analysis**

## 5.2. Structural Pattern Recognition Analysis for Failure Mode Classification

Unsupervised structural pattern recognition technique uses mathematical algorithm to determine the pattern and classification from a given data set. It is very difficult for humans to observe and analyze the complex patterns with many different parameters. Therefore, a mathematical algorithm is used to recognize the pattern from the given data sets. For the analysis of the AE signals, multi-parameter analysis would show a much more accurate and desirable results from the relationship of the AE signal characteristics and failure mechanisms.

The unsupervised pattern recognition technique is often performed to separate between the noise and the desired failure mechanism AE signals. This technique is effective for data separation, but it has disadvantages. One of the major problems is the initial indication of the number of the cluster by the user. The algorithm itself does not have ability to automatically specify the number of clusters in the data set. Another is the difference in the classification result depending on the different algorithms used [27].

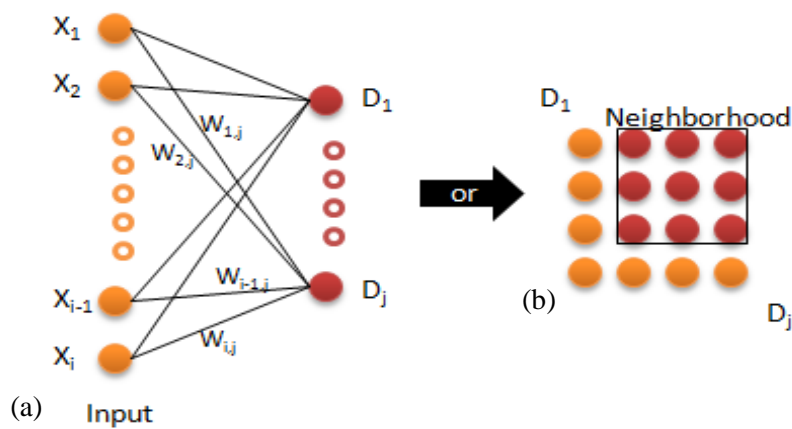
### 5.2.1. Artificial Neural Networks and Kohonen Self-organizing Map

ANNs are mathematical algorithms that function similar to the human brain by identifying complex patterns in nonlinear data space. ANNs are composed of artificial neurons, or processing elements (PEs),

that construct the system. Input layer of the ANN consist of each PEs representing specific AE signal parameter. Based on the number of the classifications defined by the user, the output can be in the form of binary or x-y coordinate representing each classification. The output layer is connected to the input layer through weight functions which are adjusted during the training phase of the network [28]. When understanding neural networks, it is necessary to observe the architecture, competitiveness, and supervision of the network.

The Kohonen SOM is a one type of unsupervised learning artificial neural network which is competitive and single layered that creates discrete representation of low-dimensional map from multi-dimensional input space [29]. From multiple iterations, the Kohonen SOM is capable of sorting large amounts of data and accurately divides them into different clusters with common characteristics.

A typical SOM architecture consists of an input layer, one or two dimensional Kohonen layers, and an output layer. Figure 11 shows an example of one dimensional Kohonen SOM. The PEs in the input layer are independent of each other, but all PEs are connected to all neurons with their own weights in the Kohonen layer.



**Figure 11 Representation of (a) one dimensional (b) two dimensional Kohonen SOM neural network**

The Kohonen SOM and unsupervised learning neural network allows the winning neuron weight to be updated, as well as the weights of the adjacent neurons that are linked through a neighborhood function to optimally exhibit the input space. The weights are calculated as shown in Equation 1.

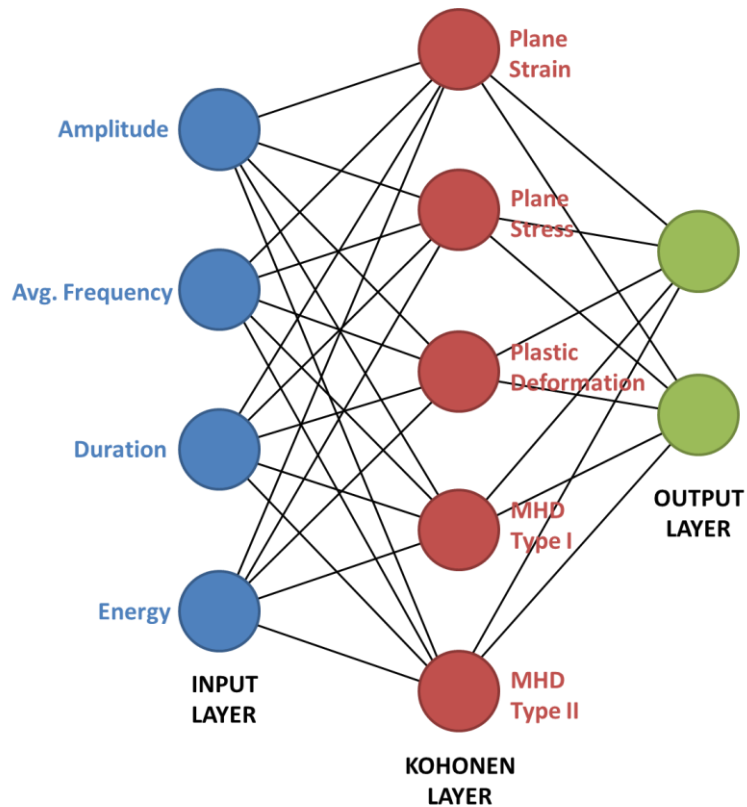
$$W_{i,j}(new) = W_{i,j}(old) + \alpha(X_i - W_{i,j}(old)) \quad (1)$$

Where  $W_{ij}$  are the weight vectors,  $\alpha$  is the learning coefficient,  $X_i$  is the input vector. Typical SOM learns by minimizing the Euclidean distance between the weight and input vectors. Euclidean distance can be calculated with following Equation 2.

$$D_j = \sqrt{\sum (W_{i,j} - X_i)^2} \quad (2)$$

Where  $D_j$  is the Euclidean distance. The network locates the neuron with the closest weight vectors to the input vectors taken from the data set, and assigns specific coordinate to this input vector. As learning and updating of the weights continues, the network will attempt to cluster the input vectors into “like” data clusters as an output.

In this case of Figure 12, the 4 input parameters are amplitude, average frequency, duration, and energy. Each PE in the input layer represents specific AE signal parameters that are used to classify the AE signal parameters to the failure mechanisms. Each PE in the Kohonen layer represents different failure mechanism classification specified by the user. At the end, output layer represents the classification by two dimensional coordinates.

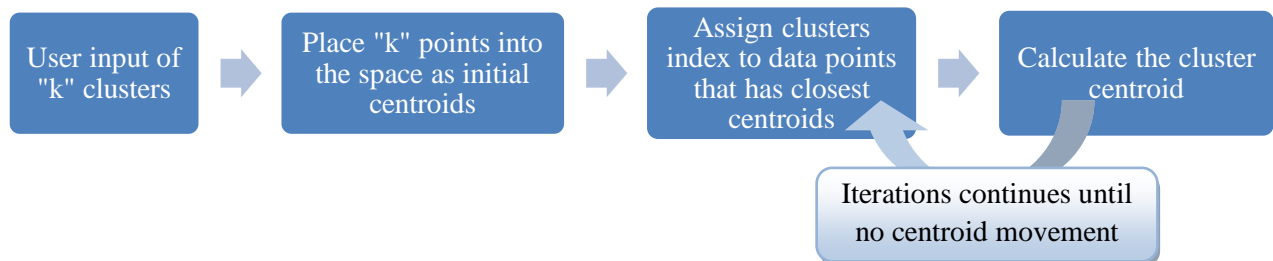


**Figure 12 Representation of one dimensional Kohonen SOM neural network with 4 input neurons, 5x1 Kohonen layer and 2 output neurons**

### 5.2.2. K-mean Clustering Algorithm

The k-mean clustering algorithm [30] is one of the simplest pattern recognition algorithms that are capable of generating a specific number of flat non-hierarchical clusters. This method is numerical, unsupervised, non-deterministic, and iterative technique. The final goal of the k-mean is to partition the data set into “ $k$ ” number clusters to have similar characteristics with other data points in the same cluster.

Figure 13 shows a flow diagram of the k-mean clustering algorithm.



**Figure 13 Flow Diagram of K-mean Clustering Algorithm**

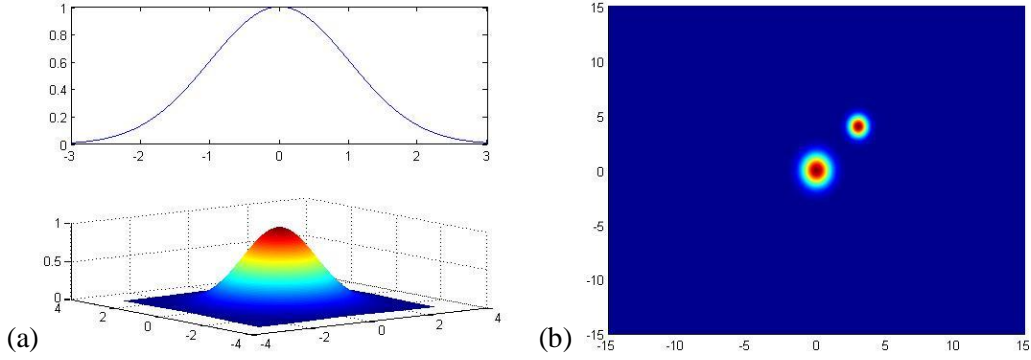
After the input of number of “ $k$ ” cluster wanted, the data points are partitioned into “ $k$ ” clusters with centroids in each. Based on the centroids, the distances between the centroids and data points are calculated to minimize the distance. The distance is minimized by the objective function (J) as shown in Equation 3.

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (3)$$

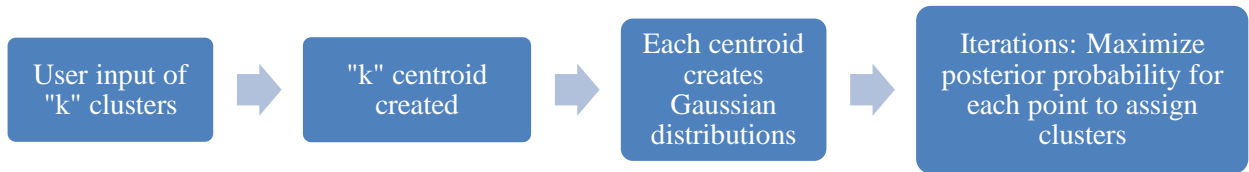
Where  $x_i^{(j)}$  is data points,  $c_j$  is cluster centroid. After the new cluster indexes are assigned to each data points, the centroid for the cluster is recalculated. Then, the distance between the centroids and the data points is calculated again. The k-mean algorithm is very simple for data clustering but has its disadvantages as an algorithm does not necessarily find the most optimal configuration. The algorithms are very sensitive to the initial randomly selected centroid points. This problem can be reduced by multiple iterations of the k-mean algorithm.

### 5.2.3. Gaussian Mixture Model Clustering Algorithm

The GMM clustering algorithm [31 - 33] is another type of pattern recognition algorithm that is capable of generating a specific number of the clusters from a given data set from unsupervised learning. The GMM is one of the mode-based approaches which use certain models for clustering and attempting to optimize the data and model. For this model, it considers clusters as Gaussian distribution centered on its centroid as seen in red from Figure 14 (b). The final goal of GMM is to cluster the data into a “ $k$ ” number of cluster by maximizing the posterior probability of each data point. Figure 15 shows the flow of the GMM algorithm.



**Figure 14 (a) Example of Gaussian Distribution (b) Multiple Gaussian Distribution from X-Y View**



**Figure 15 Flow Diagram of Gaussian Mixture Model Clustering Algorithm**

The posterior probability is calculated by updating the likelihood with the prior probability. The likelihood is calculated as shown in Equation 4.

$$\log p(X|\theta) = \sum_{i=1}^m \log \sum_{j=1}^k \pi_j \mathcal{N}(x_i | \mu_j, \Sigma_j) \quad (4)$$

Where  $\log p(X|\theta)$  is log-likelihood,  $\pi_j$  is mixing coefficients that is greater than zero,  $\mu$  is mean of respective Gaussian distributions,  $\Sigma$  is variance parameters of the respective Gaussian distributions,  $X$  is data points from  $x_1, \dots, x_m$  and  $\theta$  is set of parameters  $\{\pi, \mu, \Sigma\}$ . The set of parameters  $\{\pi, \mu, \Sigma\}$  are iterated until they converge by expectation-maximization algorithm. The GMM algorithms are more appropriate than k-mean when clusters have different sizes and correlation within the clusters.

### 5.3. Cluster Criterion Algorithms for Cluster Verification

Even when the cluster algorithm itself is very simple, it is very difficult to find the optimal cluster of configuration when performing a clustering analysis. This is because the clustering is performed without an initial understanding of the internal structure of the data and requiring the input number of cluster to be

partitioned. When the cluster verification is performed, it describes each cluster by the strong internal similarities between the data points within the same cluster. It uses criterion function, such as the sum of squared distance from the cluster centroids, to maximize the criterion functions.

### 5.3.1. Davies-Bouldin Criterion

The Davies-Bouldin criterion [34], also known as Rij criterion, is based on the ratio of average within-cluster distance to between-group distances. The average measure of such ratio is calculated using all of the different pairs of classes. Equation 5 shows the equation to calculate the Davies-Bouldin index to evaluate the quality of a given data partition.

$$DB = \frac{1}{k} \sum_{i=1}^k R_i \quad (5)$$

Where DB is the Davies-Bouldin index,  $R_i$  is maximum of  $R_{ij}$ . To obtain the  $R_i$ ,  $R_{ij}$  must be explained.

The  $R_{ij}$  can be written as shown in Equation 6.

$$R_{ij} = \frac{S_i - S_j}{M_{ij}} \quad (6)$$

Where  $S_i$  and  $S_j$  are dispersions of clusters  $i$  and  $j$  and  $M_{ij}$  are the distance between two centroids. The equations  $S_i$  and  $M_{ij}$  are shown below.

$$S_i = \left\{ \frac{1}{T_i} \sum_{j=1}^{T_i} |X_j - A_i|^q \right\}^{1/q} \quad (7)$$

Where  $T_i$  is number of vectors in cluster  $i$  and  $A_i$  is the centroid of cluster  $i$ . When  $q = 2$ , the equation is the standard deviation of the distance of samples in a cluster to the respective cluster centroid.

$$M_{ij} = \left\{ \sum_{k=1}^N |a_{ki} - a_{kj}|^p \right\}^{1/p} \quad (8)$$



Where  $a_{ki}$  is the  $k$ th component of the  $n$ -dimensional vector  $a_j$  with centroid  $i$ . When  $p = 2$ , the equation is the Euclidian distance between centroids.

Minimizing the  $R_i$  for all cluster will minimize the Davies-Bouldin index. Therefore, to have a compact and separate cluster, the Davies-Bouldin index will have small values.

### 5.3.2. Silhouette Criterion

The Silhouette criterion is based on the ratio of the intra-group distance and the inter-group distance. In this research, the Simplified Silhouette criterion [35] was utilized. The Simplified Silhouette criterion was chosen since the original Silhouette criterion needs to calculate the distance between all of the data points which will result in increase of computational time. This Simplified Silhouette will calculate the distance between points and cluster centroids. Equation 9 shows the equation to calculate the silhouette index to evaluate the compactness and quality of the cluster.

$$S_{x(j)} = \frac{b_{p,j} - a_{p,j}}{\max(a_{p,j}, b_{p,j})} \quad (9)$$

Where  $a_{p,j}$  is the distance of the  $i$ th object to the centroid of its cluster,  $b_{p,j}$  is the distance of the  $i$ th object to the centroid of its closest neighboring cluster.

By computing distances among all objects to the centroids, the Simplified Silhouette criterion indicates a better quality by having a higher Silhouette index.

### 5.3.3. Tou Criterion

The Tou criterion [36] corresponds to the ratio of the minimum distance between any pair of classes to the maximum of the average within class distances shown in Equation 10.

$$\lambda(N_c) = \frac{\min D_{ij}}{\max D_j} \quad (10)$$

Where  $D_{ij}$  is minimum distance between pair of classes and  $D_i$  is an average distance between the classes.

Following Equations 11-12 show  $D_{ij}$  and  $D_i$ .

$$D_{ij} = \|z_i - z_j\| \quad (11)$$

Where  $z_i$  are centroids for cluster  $i$ .

$$D_i = \sqrt{2 \sum_{i=1}^n (\sigma_{ij})^2} \quad (12)$$

Where  $\sigma_{ij}$  is denoted as Equation 13.

$$\sigma_{ij} = \sqrt{\frac{1}{N_j} \sum_{x_k \in C_j} (x_{ik} - z_{ij})^2} \quad (13)$$

Where  $x_{ik} - z_{ij}$  is the distance between the data point to the centroid.

From Equation 10, the optimal clusters occur when  $\lambda(N_c)$  reaches a maximum peak.

## 5.4. Analysis Software

### 5.4.1. NeuralWorks® Professional II/Plus

The NeuralWorks® Professional II/Plus [37] is software utilized for neural network analysis. It can run on multiple operating systems and have the file compatibility between the operating systems. This software allows the user to quickly operate one of the 28 standard neural networks architectures in the software. While a user can utilize the standard neural networks to design, build, train, and test to solve complex problems, NeuralWorks® Professional II/Plus has capabilities for the user to create their own network by writing customized script files. In this research, the Kohonen SOM from the one of the 28 standard neural networks architectures was utilized to perform a clustering analysis on the AE signals.

#### **5.4.2. MATLAB®**

The MATLAB® is powerful numerical computing software that can be used to analyze data and develop algorithms and models. It can be used in a wide range of applications, but in this research, MATLAB® is utilized for data processing, clustering analysis using k-mean and GMM functions, and plotting of the results.

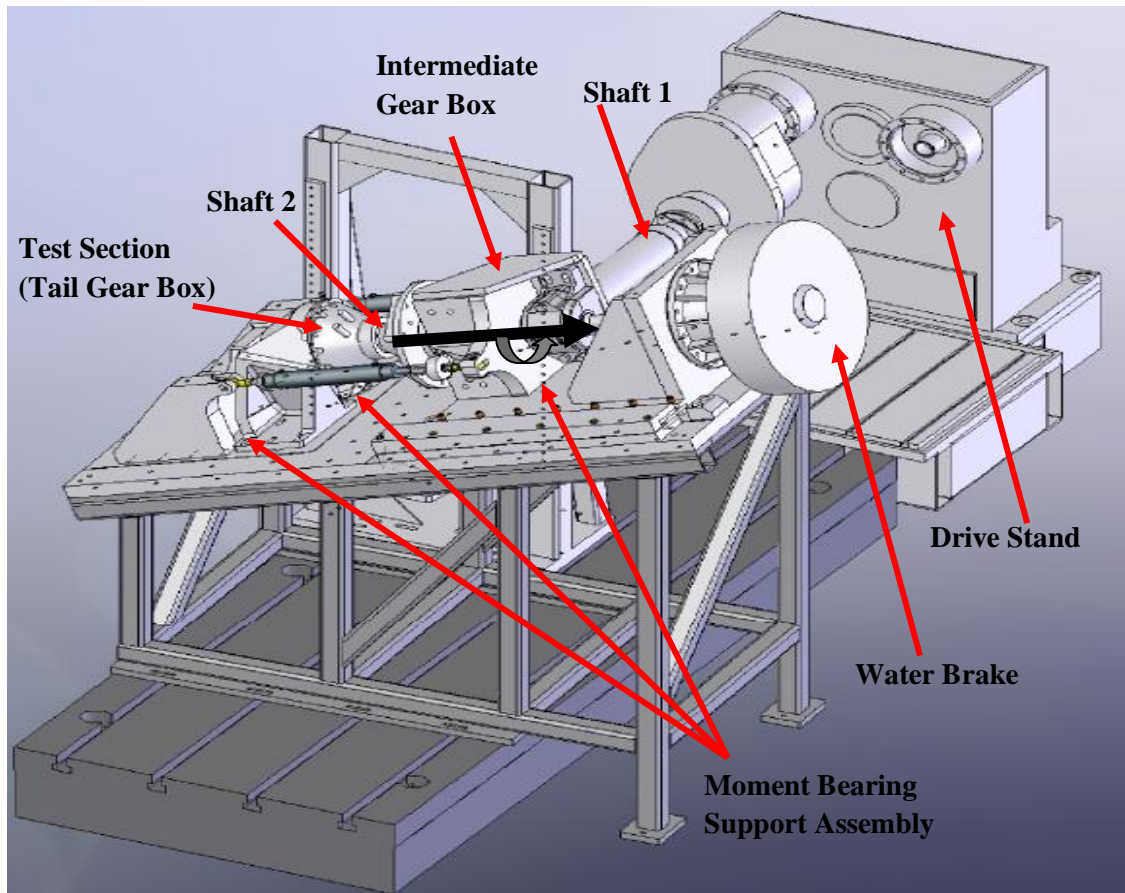
## **Chapter 6. Data Acquisition Setup of the Rotating Gear Box**

### **6.1. Ground Test**

The acquisition was performed on a ground test stand at Naval Air Station Patuxant River, Maryland in 2010 and 2011. The test stand was capable of simulating the service load conditions (hub moment and thrust) on the H-60 helicopter tail gearbox by loading actuators. The maximum continuous moment load can be as high as 70,000 in-lb<sub>f</sub>, while the thrust load can be as high as 2,000 lb<sub>f</sub>. While the output shaft of the tail gearbox is a rotating motion, these two loads can be applied separately or simultaneously. During the test, the AE transducers are mounted on the casing of the tail gearbox to record the signals that are subsequently analyzed using the processing software.

#### **6.1.1. Setup**

Figure 16 shows the 3D CAD drawing of the complete ground test stand. The setup consists of a drive stand, intermediate gearbox, moment bearing support assembly, water brake, and the tail gearbox, where bevel gear is housed. The drive stand supplies the torque to the intermediate gearbox through shaft 1. The shaft 2 is connected between the intermediate gearbox and the tail gearbox as well as the water brake for the rotational load absorption to simulate the loading effect of the H-60 helicopter tail rotor. The maximum rotating speed is 1,189 RPM. The black arrow in the Figure 16 shows the direction of the rotor.



**Figure 16 3D Drawing of Ground Test Setup [3]**

### **6.1.2. Transducer**

To observe the different reactions of the transducers, different transducers and number of the channel were used in the two data acquisitions as shown in Table 8. From the observation, the transducer to be used in further research was to be determined.

In 2010, a PAC R15I transducer was utilized and was attached to the casing of the gear box. The two transducers were placed 90 degrees apart from each other as shown in Figure 17(a). The PAC R15I transducer has an integrated 40dB preamplifier inside the transducer casing. The integrated preamplifier allows the transducer to have high sensitivity and capability of connecting long cables without additional preamplifiers. This type of transducer provides a good mixture of high sensitivity and high low frequency rejection to be useful to monitoring many common structures such as pipeline, bridges, and storage tanks [38].

In 2011, a PAC WD transducer and a PAC Micro30S transducer were used. Two of each transducer were attached to the casing of the tail gear box as shown in Figure 17(b). Two different transducers were placed next to each other as a set, and the two sets were spaced 90 degrees apart as same as the 2010 test. The PAC WD transducer is a differential wideband sensor that has very high sensitivity and bandwidth. The differential transducer differs from a general sensor by having two sensing elements with opposing polarization direction built-in the casing. The two signals from the two sensing elements are feed into differential preamplifier to eliminate the common noise. WD sensors are used for structural health monitoring of large structures like storage tanks and pipeline, as well as research applications where high AE response is needed [39]. The PAC Micro30S transducer is a small sized sensor with good sensitivity to be used in applications with size and weight limitations. With its small size and the sensitivity, it is ideal for structural health monitoring of structures like aircraft and storage tanks [40].

**Table 8 Transducer and Channels for each Test [38] [39] [40]**

<b>Acquisition Year</b>	<b>Channel</b>	<b>Transducer</b>	<b>Resonant Frequency (kHz)</b>	<b>Operating Frequency Range (kHz)</b>
<b>2010</b>	1, 2	PAC R15I Transducer	150	80-200
<b>2011</b>	1, 3	PAC WD Transducer	450	125-1000
	2, 4	PAC Micro30S Transducers	225	150-400

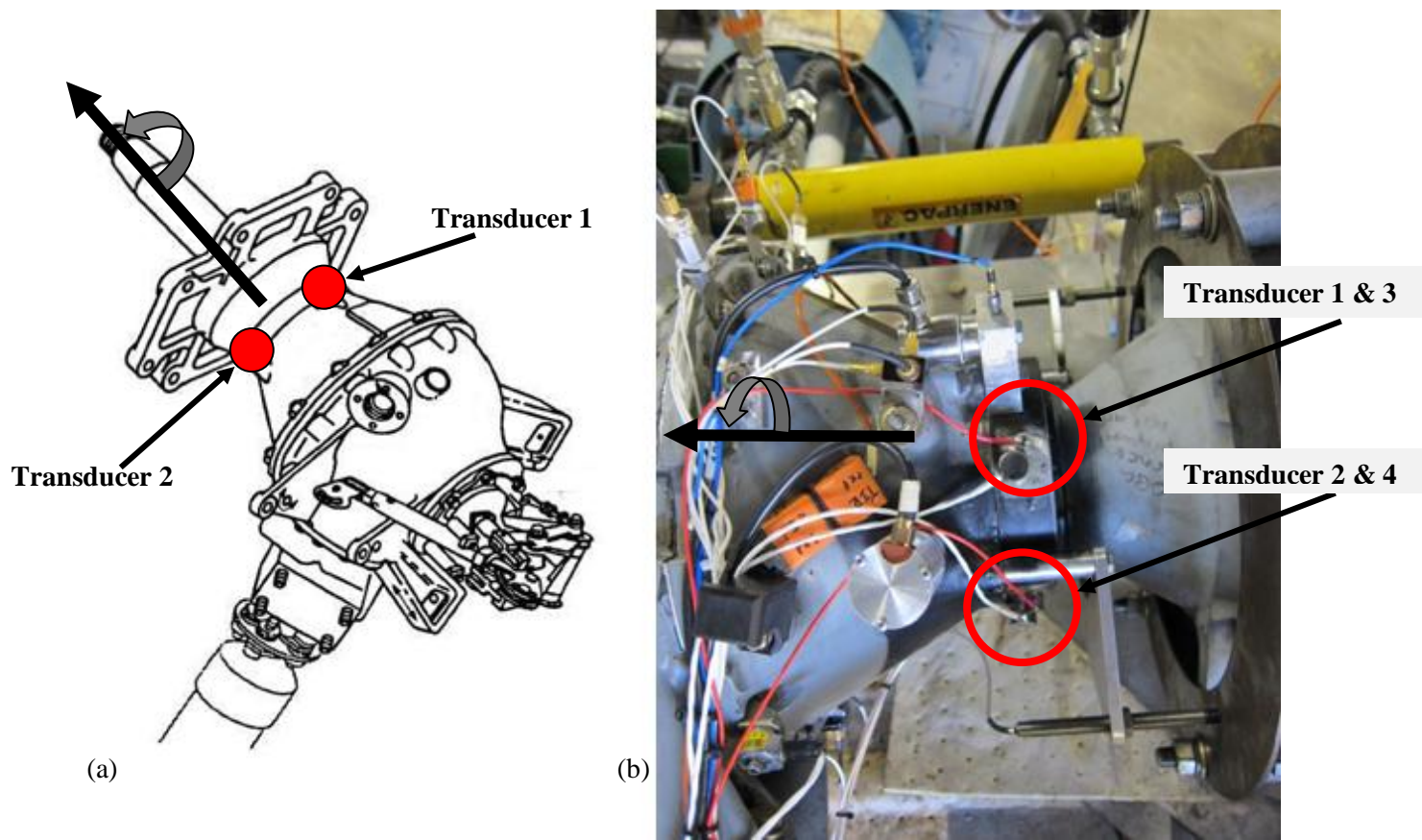


Figure 17 Transducer Position for (a) 2010 [41] (b) 2011

## 6.2. Parameters

### 6.2.1. AE Parameters

With the improvement of the recording hardware in past decades, it is possible to record many AE signal hit data using recent equipment. In the two acquisition years, different amounts of the AE hit parameters were recorded as seen in Table 9. Of the many parameters, four were selected: energy, duration, amplitude, and counts. The parameter to be used for analysis were chosen by experiences of the past work and research to observe the best classification performed by the clustering algorithms.

Table 9 Recorded AE Hit Signal Parameters

Acquisition Year	Recorded AE Hit Parameters
2010	Rise time, Counts, Energy, Duration, Amplitude, Channel
2011	Rise time, Counts, Energy, Duration, Amplitude, Average Frequency, Initialization Frequency, Signal Strength, Absolute Energy, Centroid Frequency, Peak Frequency, Channel, Partial Power

### 6.2.2. Loading Profile

For the analysis of the fatigue crack in the bevel gear, the mission profile of the H-60 helicopter was simulated in the data acquisition. Figure 18(a) (b) shows the hub moment loading and the rotational speed of the bevel gear on September 7th and October 18th, 2010. The rotating speed ranged from 0 to 1,189 RPM and a different bending moment load was applied simultaneously. On different date, hub moment load, rotational speed, and acquisition time were different.

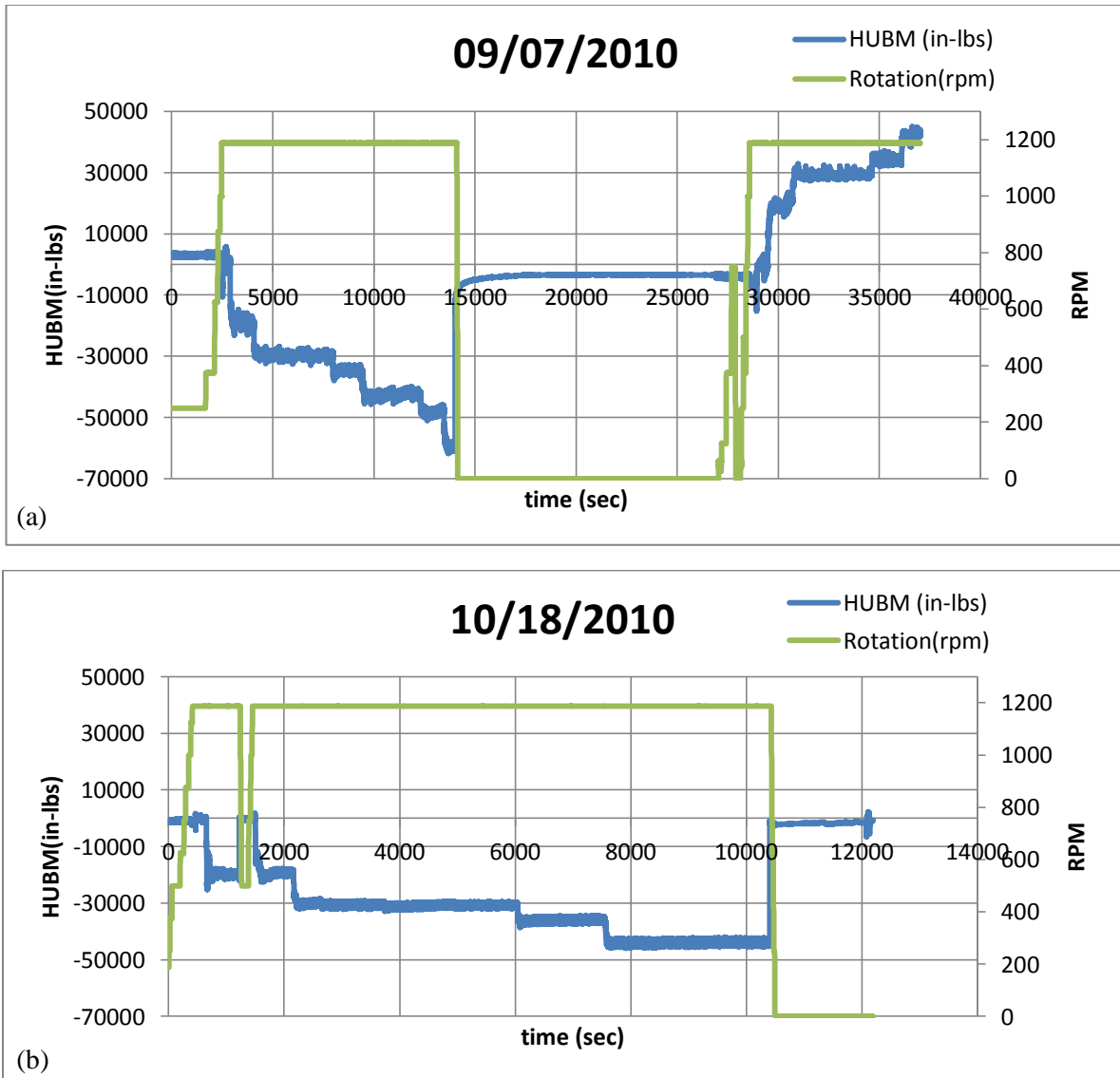


Figure 18 Hub Moment Load and Rotational Speed of the Bevel Gear in (a) September 7, 2010 (b) October 18, 2010



## Chapter 7. Pre-analysis Data Processing

### 7.1. Overview

To prevent incorrect classification of the failure mechanisms by the clustering algorithm, the data needed to be sorted before the classification analysis. The raw data was checked to see if it contained enough AE signal to be analyzed. Figure 19 shows the sorting methods that were applied to the acquired AE signal data.

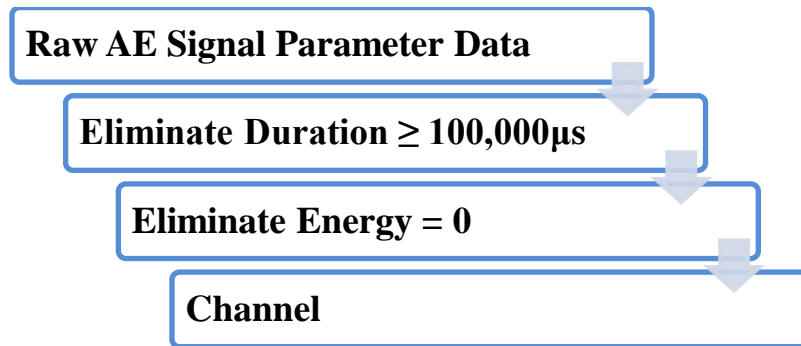
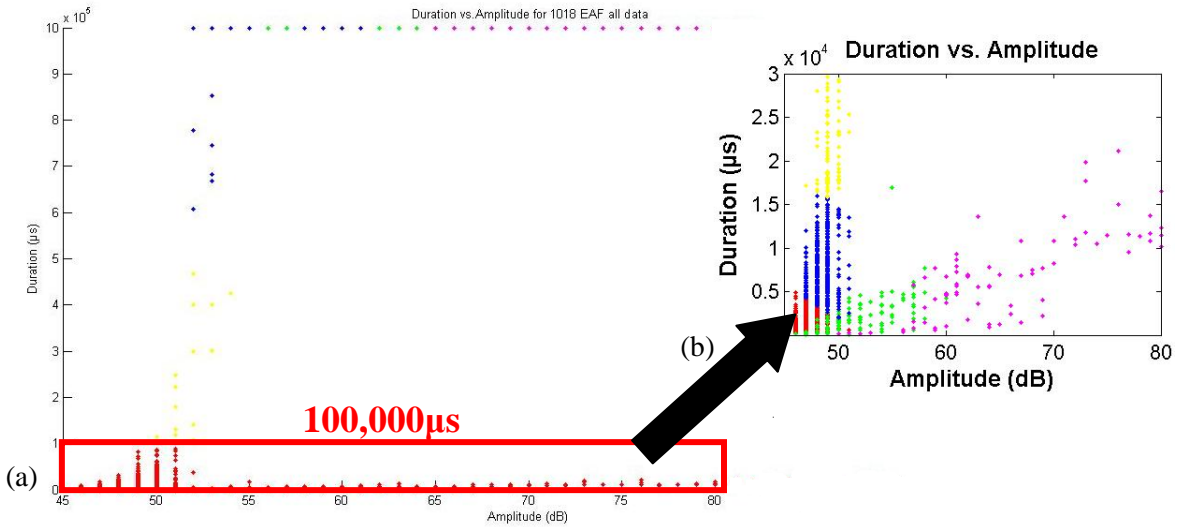


Figure 19 Sorting Method for AE Data

### 7.2. Elimination of Duration

Durations greater than 100,000  $\mu\text{s}$  were sorted out first since these long duration data are considered as noises or multiple hit data (MHD). MHD are AE hits in which several AE signals are closely spaced in overlapping time, making its statistics very different compared to single hit data. Rubbing or friction noises are typical example of MHD. Since the goal of this analysis is to observe the fatigue crack growth in the bevel gear, only AE signal with duration  $\leq 100,000 \mu\text{s}$  which are boxed red in Figure 20(a) are used. The cutoff point of duration of 100,000  $\mu\text{s}$  was chosen since the plot from Figure 20 (a) indicates there is only one cluster under 100,000  $\mu\text{s}$ . Figure 20 (b) shows the analysis with only the signal duration  $\leq 100,000 \mu\text{s}$ . Comparing (a) and (b), it can be seen that the analysis is much more in detail.



**Figure 20** October 18th AE Hit Data with Duration vs. Amplitude (a) all signals (b) signal duration  $\leq 100,000\mu\text{s}$

### 7.3. Elimination of Energy

Zero energy AE hit data were also eliminated since it was not a meaningful AE data. When the signal are received by the AE system, the energy reading less than 0.5 are round down as zero energy hits. This is because of the conversion between analog and digital signals by the acquisition system. Table 10 and 11 shows the comparison of AE hit with zero energy and energy other than zero. AE signal data with zero energy hit have lower duration time and counts compared to the AE hit with energy greater than 0. This indicates that the AE signals with energy less than zero can be eliminated from the analysis.

**Table 10** AE Data with Zero Energy

Counts	Energy	Duration ( $\mu\text{s}$ )	Amplitude (dB)
1	0	1	45
1	0	1	47
2	0	38	46

**Table 11** AE Data with Energy Greater than Zero

Counts	Energy	Duration ( $\mu\text{s}$ )	Amplitude (dB)
16	42	5267	48
98	42	3589	52
12	43	5834	47

#### 7.4. Elimination of Channels

After the sorting was performed, AE signals were separated into channels to observe the different analysis result of the clustering analysis. For the 2010 data acquisition, it showed that the Channel 1 had more fatigue crack signal compared to Channel 2, as shown in Figure 21. One possible reason for this is that the fatigue cracks could have been occurring closer to Channel 1; hence Channel 2 data were put aside for further analysis for the 2010 acquisition data.

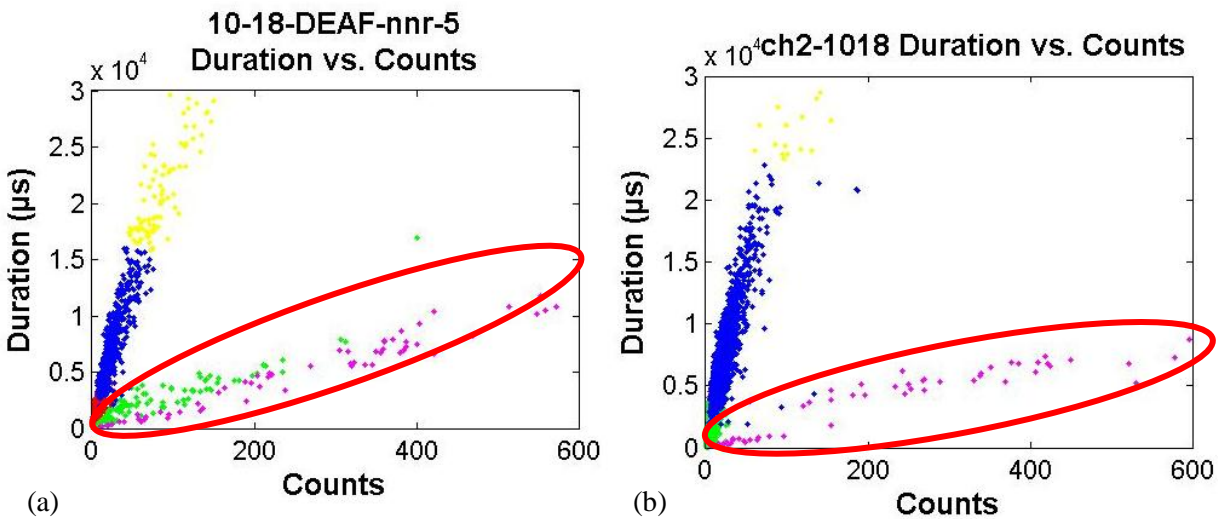
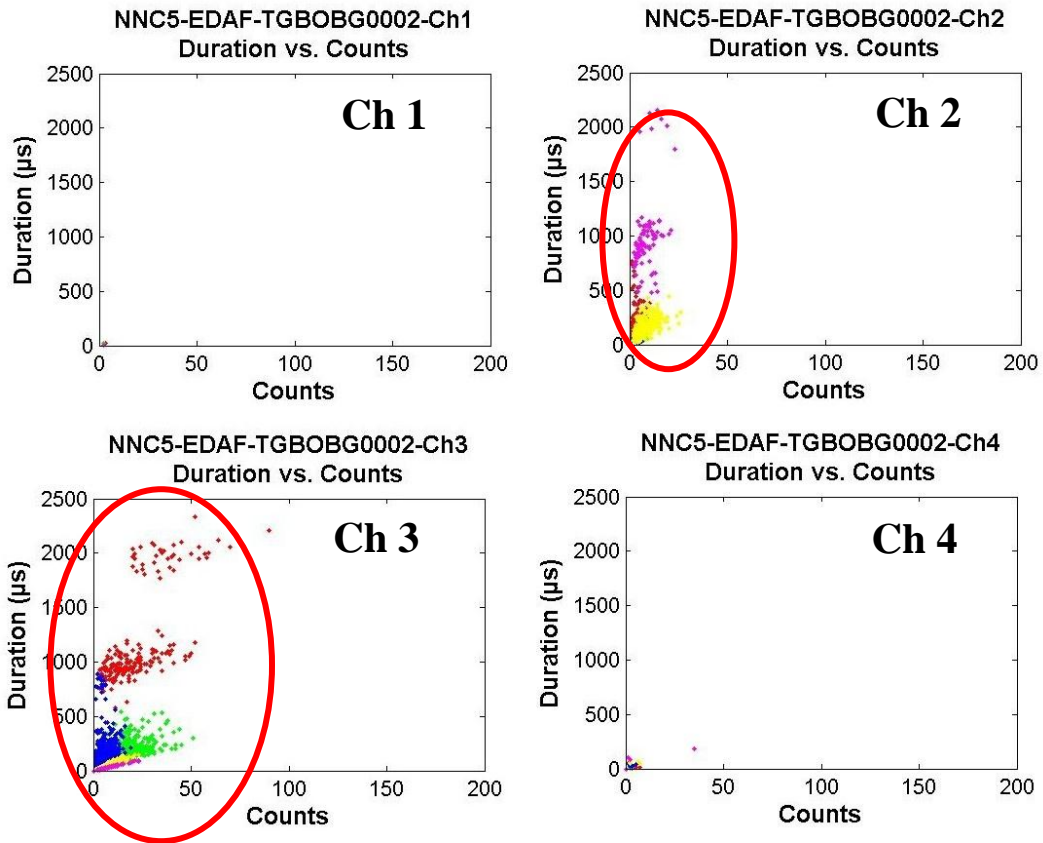


Figure 21 Duration vs. Counts of October 18, 2010 (a) Channel 1 (b) Channel 2

In 2011, four transducers were used to acquire AE signals. From the four transducers, Channel 2 and 3 showed the most AE activities as shown in Figure 22. Since Channel 3 showed more AE activity compared to others, it was chosen for the further analysis.



**Figure 22 Duration vs. Counts AE Signals from TGBOBG0002 of 2011 Acquisition for Different Channels**

## Chapter 8. Iterations and Inspection

### 8.1. Analysis Iterations

To perform the clustering analysis, numbers and parameters for input and number of cluster for an output must be selected. The input AE parameters utilized were counts, duration, energy, amplitude, and average frequency. From these five parameters, different combinations were utilized to see the output classification result of the three clustering algorithms used. The output numbers were varied from 3-7 clusters to observe the performance of each clustering techniques using the clustering verification criteria. Table 12 shows a list of all the iterations performed.

**Table 12 AE Input Parameters and Output Iterations**

Iteration	Input Parameters	Number of Input Parameter	Number of Output Clusters
1	Duration, Energy, Amplitude	3	3 - 7
2	Counts, Duration, Energy, Amplitude	4	3 - 7
3	Energy, Amplitude, Average Frequency	3	3 - 7
4	Duration, Energy, Amplitude, Average Frequency	4	3 - 7

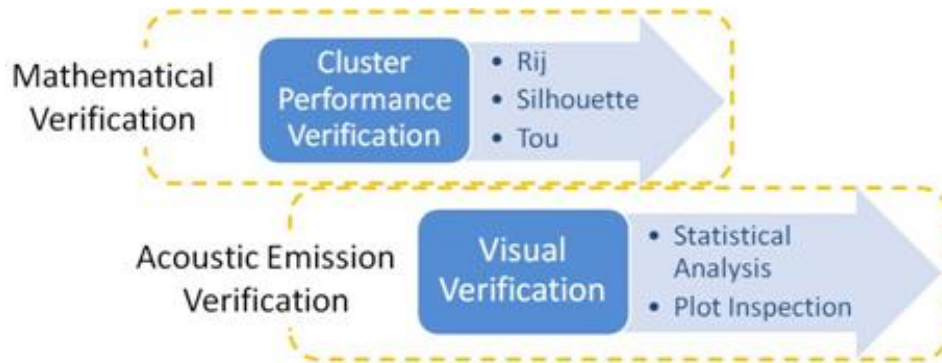
The first two analyses were performed by using the native AE signal parameters obtained. Since the first two clustering show inconclusive results from the clustering criterions, average frequency was introduced to the clustering analysis as a new input parameter to observe the classification performance from the third iteration on. Out of the five input parameters, average frequency is only parameter that is not native from the AE signals. Equation 14 shows the equation of the average frequency.

$$\text{Average Frequency} = \frac{\text{Counts}}{\text{Duration}} \quad (14)$$

### 8.2. Cluster Inspection

The AE data points clustered by the three clustering algorithms are inspected in two steps, as shown in Figure 23. The two steps are cluster performance verification by using mathematical model and visual verification. From the performance verification, a possible number of clusters in the data set will be suggested. To verify the possible number of clusters in the data set, visual verification by performing

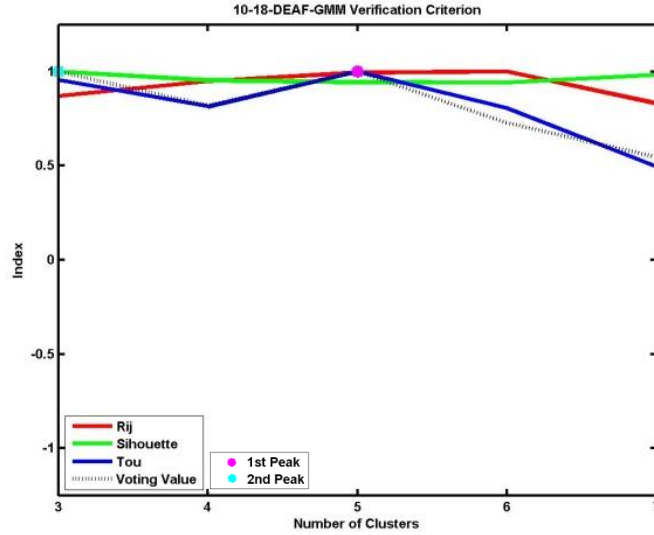
statistical analysis on each cluster and plotting the acoustic emission plots are performed. From the two, the number of clusters can be verified by the AE characteristics of each mechanism.



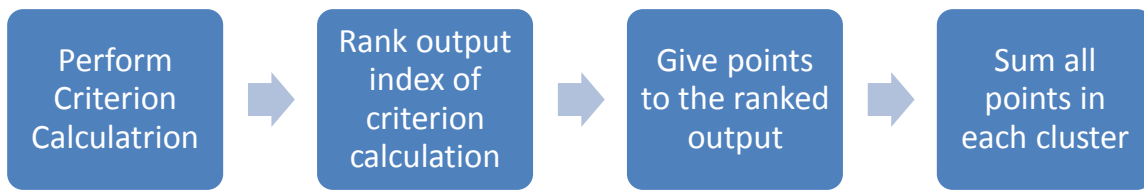
**Figure 23 Cluster Inspection Flow**

### 8.2.1. Criterion Verification of the Clusters

Based on the three mathematical cluster performance analyses, the results can show the optimal number of the cluster in a data set. Figure 24 shows the example plot of criteria depending on the clustering techniques. In this plot, all numbers were normalized to have them fall between -1 to 1. Also, the criterion index value of the Davies-Bouldin criterion has been inversed to have the maximum criterion index value as the optimal number of clusters. Using each criterion index value calculated, each cluster index values for each criterion are given a point regarding to their rank within. In this research, the 5 different clusters are observed in between 3 through 7, the highest peak will be given 10 points, and the second highest peak will be given 9 points and so on. From the points given, all three points per cluster number are summed to include the voting plot in the verification plot. The cluster with the highest point are indicated with magenta dot and the second highest is shown as light blue dot. Figure 25 shows the flow of the voting plot.



**Figure 24 Example of Criterion Verification Plot**

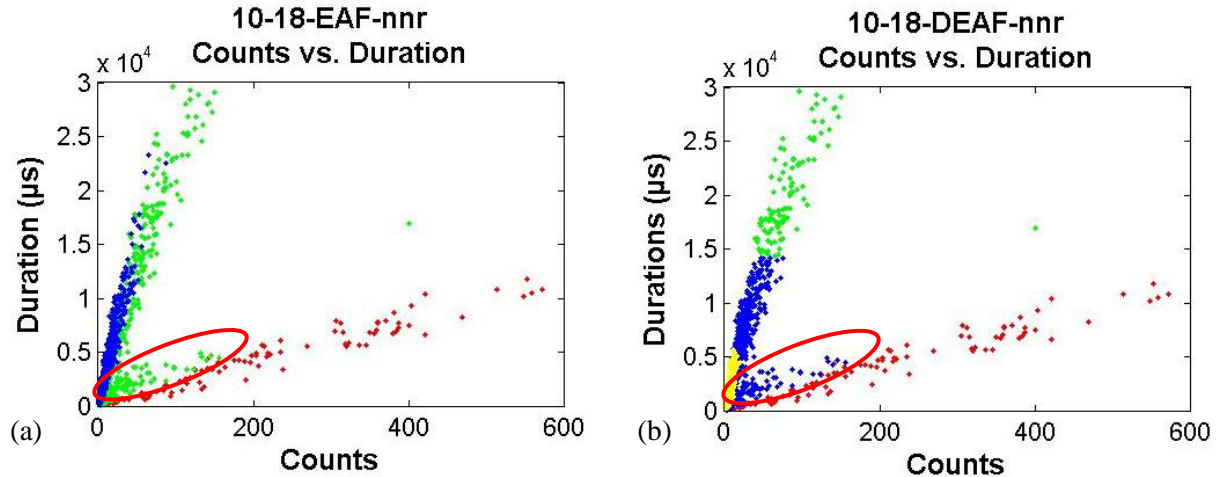


**Figure 25 Flow of the Criterion Voting Plot**

From each clustering algorithm, the possible number of clusters in the data set will be predicted by the cluster criteria. To be sure the result from the criterion is the best possible number of the cluster in the data set, further verification of acoustic emission plots and statistical analysis of the each cluster will be necessary.

### 8.2.2. AE plots and Statistical Analysis of Clusters

Clustering can be verified by visual inspection. Figure 26 shows two examples of improper classification. Figure 26(a) is shown using Kohonen SOM with input AE parameters as energy, amplitude, and average frequency. This analysis shows the possibility of two types of signals within the green set of data by part of the green data points extending into the same direction as red area (circled in red).



**Figure 26 October 18, 2010 Duration vs. Count (a) Input: EAF Output 3 (b) Input: DEAF Output 4**

Additionally, Figure 26(b) shows improper classification with Kohonen SOM clustering into 4 clusters with inputs of duration, energy, amplitude, and average frequency as shown in red circle. Initial expectation of this iteration was that SOM would be able to further dissect the fatigue crack data and separate it into two different types: plane strain and plane stress cracking. However, the output result shows that this iteration was able to separate the plastic deformation cluster as shown in yellow. Still, the classification output showed that the blue data points are extending toward the red area which is plane stress. Hence, further increase of clusters was needed.

After the plots have been checked to have a proper classification, statistical analysis of each clusters are performed to observe the relationship between the clusters. The parameters to be investigated are the minimum, maximum, average, standard deviation, and number of hits of the each cluster. From observing all parameters in statistical analysis of each cluster with the relationship of the each cluster mechanisms with AE signal parameters shown in Table 4, it is possible to determine the accuracy and mechanism of each cluster.

With the optimal cluster number possibility suggested from the clustering verification criteria, the visual inspection of the acoustic emission plots and the statistical analysis can conclude the optimal number of cluster in the data set.



## **Chapter 9. Results**

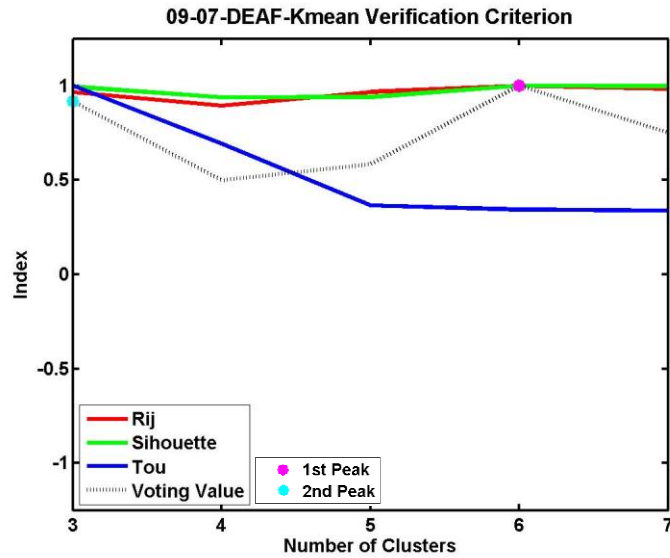
The 2010 data collected with lower threshold of 45 dB had more significant hits than the 2011 data with higher threshold of 65 dB. The difference in the threshold value resulted in the different amount of AE hits to be analyzed. Results presented through 9.1 and 9.4 are for 2010 data and the results for 2011 are presented at 9.5.

### **9.1. Clustering Criterion Verification of 2010 Data**

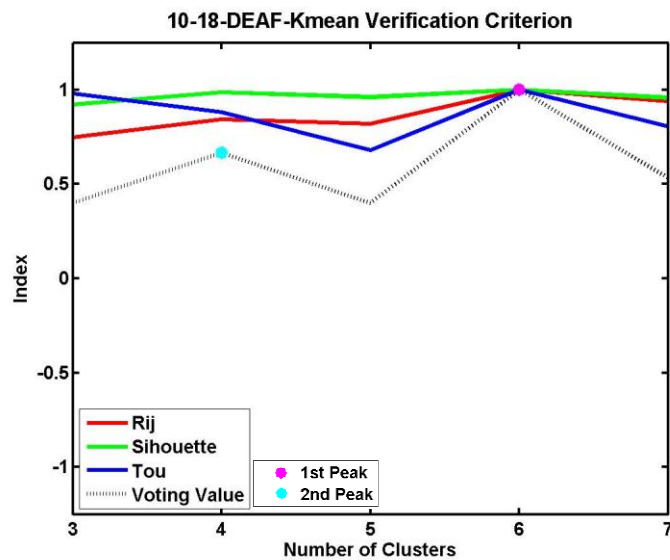
Two of the acquisition dates were chosen to demonstrate the methodology and techniques established in this research. The two dates are September 7th and October 18th. These dates represent a midpoint data set and final date data set from the data acquired, respectively. From the data set clustered by the three clustering algorithm, cluster criterion verification was performed to obtain a suggested optimal number of clusters, followed by visual verification to confirm the suggested number of optimal clusters in a data set.

#### **9.1.1. K-mean Clustering Criterion Verification**

Figures 27 and 28 demonstrate the clustering criterion verification of the two dates after K-mean clustering algorithm. From the voting value plot of Figure 27, it indicates that the September 7th data has maximum peak at 6 clusters and next highest peak at 3 clusters. This indicates the possibility of the data having 6 clusters or 3 clusters in the September 7th data set.



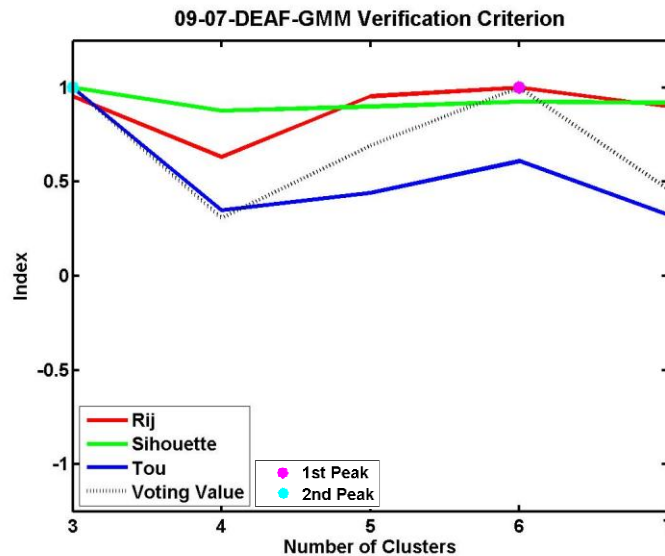
**Figure 27 September 7th K-mean (Input: Duration, Energy, Amplitude, Avg Freq.) Cluster Criterion**  
 For October 18th data set, it indicates the maximum peak at 6 clusters and next highest peak at 4 clusters as shown in Figure 28. From this date, it indicates the possibility of 6 or 4 clusters in the data being optimal.



**Figure 28 October 18th K-mean (Input: Duration, Energy, Amplitude, Avg Freq.) Cluster Criterion**

### 9.1.2. Gaussian Mixture Model Clustering Criterion Verification

Figures 29 and 30 show the clustering criterion verification of the two dates after GMM clustering algorithm. From the voting value plot of Figure 29, it indicates that the September 7th data has maximum peak at 6 clusters and next highest peak at 3 clusters.



**Figure 29 September 7th GMM (Input: Duration, Energy, Amplitude, Avg Freq.) Cluster Criterion**

For the October 18th data set after GMM, it indicates the maximum peak at 5 clusters and next highest peak at 3 clusters as shown in Figure 30. From this date's data set, it indicates the possibility that 5 or 3 clusters in the data are optimal.

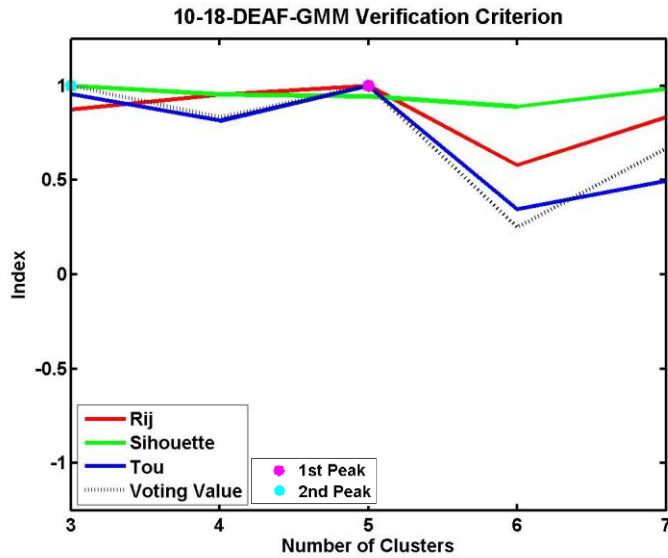


Figure 30 October 18th GMM (Input: Duration, Energy, Amplitude, Avg Freq.) Cluster Criterion

### 9.1.3. Kohonen Self-organizing Map Clustering Criterion Verification

Figures 31 and 32 show the clustering criterion verification of the two dates after Kohonen Self-organizing map clustering algorithm. As shown in Figure 31, the voting plot indicates the maximum peak of the criterion at 3 clusters and next highest at 4 clusters.

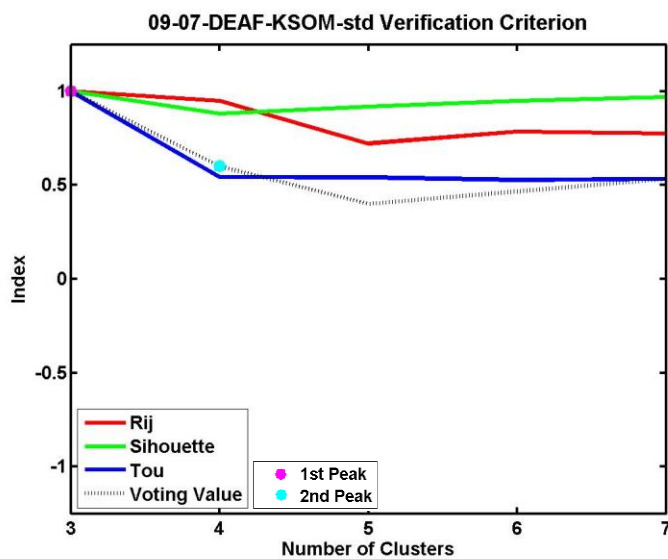
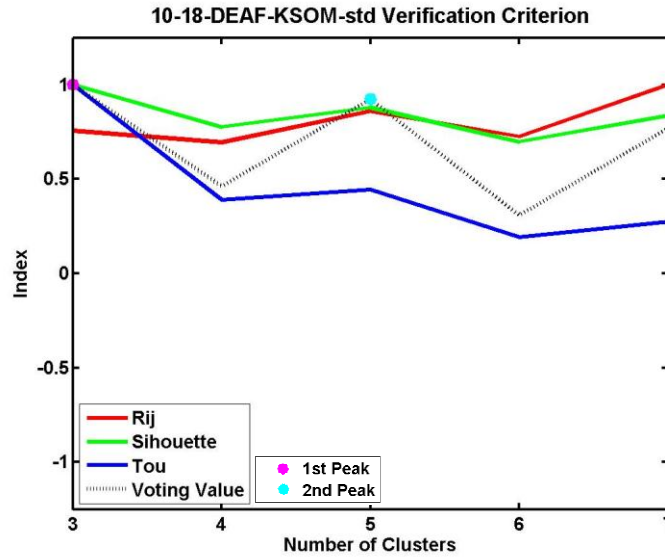


Figure 31 September 7th K-mean (Input: Duration, Energy, Amplitude, Avg Freq.) Cluster Criterion

Figure 32 indicates the criterion analysis performed on October 18th from Kohonen SOM. It indicates that the criteria have the maximum peak at 3 clusters and next highest at the 5 clusters.



**Figure 32 October 18th K-mean (Input: Duration, Energy, Amplitude, Avg Freq.) Cluster Criterion**

#### 9.1.4. Cluster Number Possibility from Clustering Criteria

From the each clustering algorithms, the criterion peaks are tabulated as shown in Table 13. For each date, the 1st and 2nd peak clusters are listed as well as the highest possible clusters that the criteria are suggesting. The table cells with orange color indicate the highest peak for that date.

**Table 13 Criterion Peaks of Each Date using Different Cluster Algorithms for 4th Iteration**

KMeans	Peak	1st	2nd	High
	907	6	3	6
	1018	6	4	6
GMM	Peak	1st	2nd	High
	907	6	3	6
	1018	5	3	5
KSOM	Peak	1st	2nd	High
	907	3	4	4
	1018	3	5	5

Since the metal structures should have at least 3 clusters for these data sets (plastic deformation, fatigue crack, and noise), the higher peak was considered to be the optimal number of clusters suggested by the

criteria. From Table 13, it indicates that the highest criterion peak occurred in between 4 and 6 clusters. To perform the visual verification, the clusters between 4 through 6 were investigated to observe the performance of the clustering by statistical analysis and acoustic emission plots.

## **9.2. Visual Verification of 2010 Data**

After the cluster criterion verification, the visual verification is done by performing statistical analysis and observation of the acoustic emission plots to suggest the number of clusters in the data set. Different colors in the plots indicate the different clusters separated by the clustering algorithms. After the visual verification suggesting single number of possible cluster in the data set, different mechanisms can be determined by comparing each mechanism with the AE characteristics of each cluster.

### **9.2.1. K-mean Clustering**

#### **9.2.1.1. September 7th K-mean**

Figures 33-35 and Tables 14-16 show the AE plots and the statistical analysis performed on September 7th using the k-mean clustering algorithm. From the cluster criteria, it has provided with a possibility that 6 clusters in the data set are the optimal number of cluster.

By visually observing Figures 33-35, it is possible to detect one of the clusters is extending in two directions from the duration versus counts plot. This indicates that there are two different average frequencies in one cluster. Due to this reason, it is possible to state that the k-mean clustering algorithms were not able to classify the September 9th data into definite clusters. The cluster with two different average frequencies is circled in red.

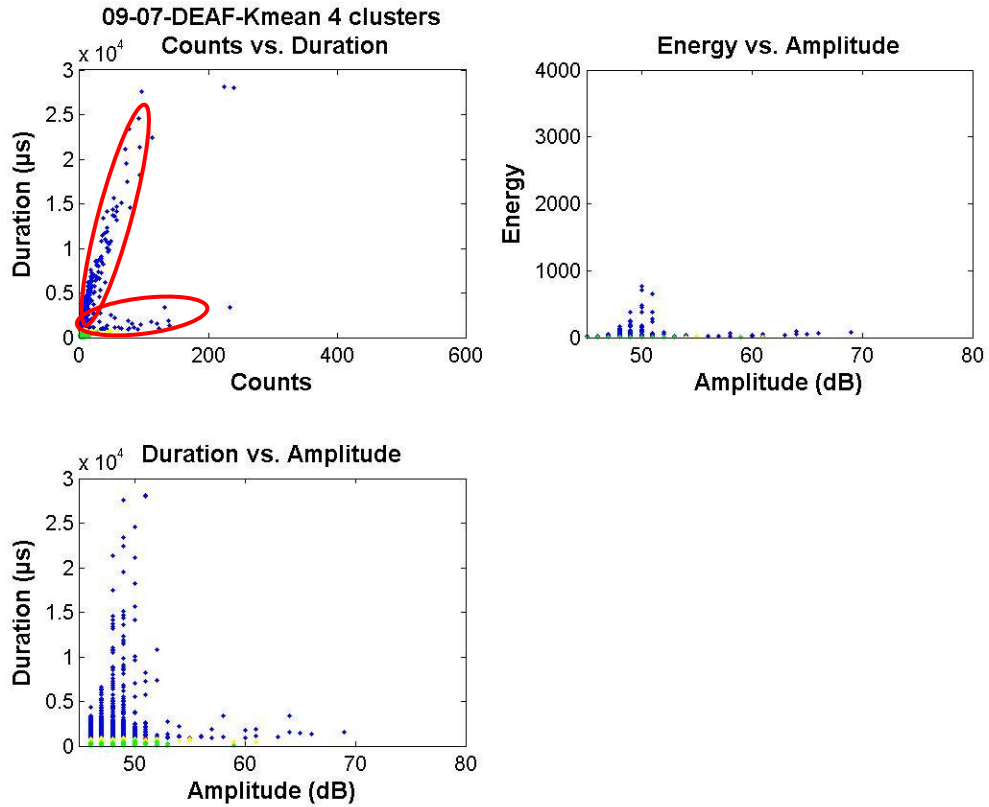


Figure 33 September 7th - 4 Clusters K-mean Clustering AE Plots

Table 14 September 7th – 4 Clusters K-mean Statistical Analysis (a)Amplitude (b)Duration (c)Energy

Amplitude							Duration					
	Min	Max	Average	SD	# of Hits	Rank	Min	Max	Average	SD	# of Hits	Rank
	45	51	46.31	1.01	7086	4	480	1189	729.72	148.50	7086	2
	45	69	47.05	1.64	3071	1	897	97406	1991.71	3770.72	3071	1
	45	61	46.37	1.14	4618	2	228	878	432.45	99.10	4618	3
	45	59	46.34	1.15	3748	3	45	540	215.31	62.70	3748	4

Energy						
	Min	Max	Average	SD	# of Hits	Rank
	2	7	4.20	1.00	7086	2
	6	772	13.77	32.56	3071	1
	1	15	2.46	0.85	4618	3
	1	7	1.11	0.38	3748	4

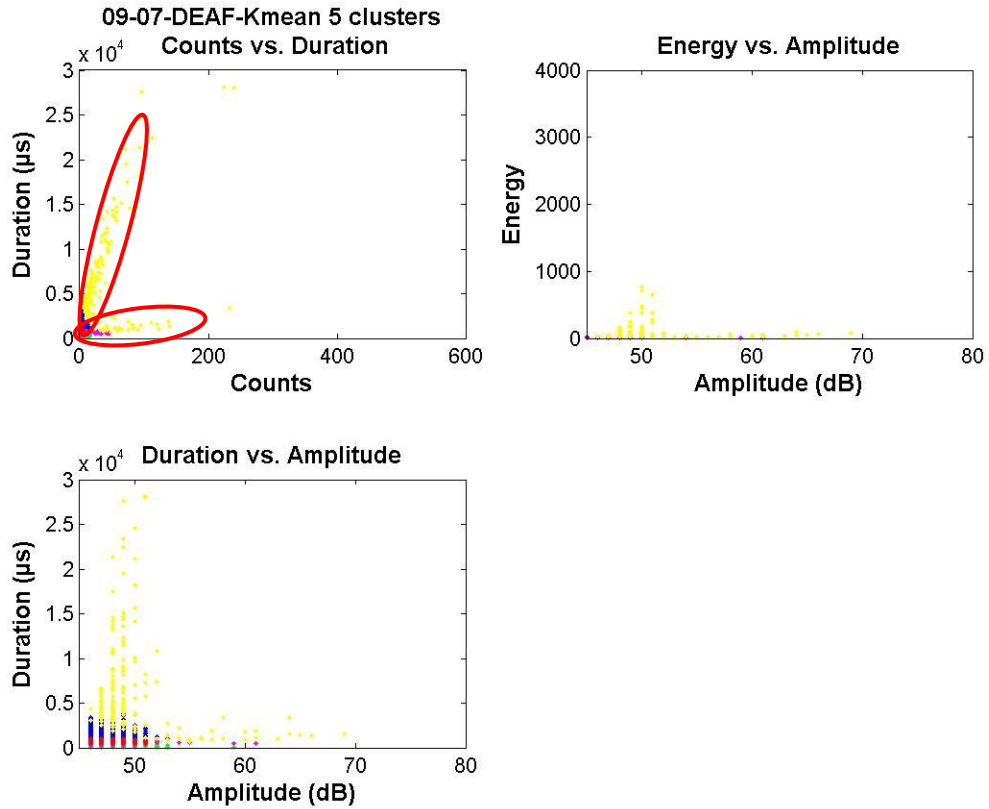


Figure 34 September 7<sup>th</sup> - 5 Clusters K-mean Clustering AE Plots

Table 15 September 7<sup>th</sup> – 5 Clusters K-mean Statistical Analysis (a)Amplitude (b)Duration (c)Energy

Amplitude							Duration					
	Min	Max	Average	SD	# of Hits	Rank	Min	Max	Average	SD	# of Hits	Rank
	45	53	46.78	1.07	3590	2	873	3766	1411.72	435.58	3590	2
	45	59	46.34	1.15	3695	4	45	493	213.63	61.38	3695	5
	45	61	46.37	1.13	4429	3	225	878	426.38	97.70	4429	4
	45	51	46.29	1.01	6609	5	464	1052	690.73	120.36	6609	3
	46	69	50.03	3.95	200	1	653	97406	8902.35	12856.91	200	1

Energy						
	Min	Max	Average	SD	# of Hits	Rank
	5	23	8.79	2.93	3590	2
	1	7	1.10	0.36	3695	5
	1	13	2.40	0.81	4429	4
	2	6	3.97	0.84	6609	3
	11	772	75.24	110.33	200	1



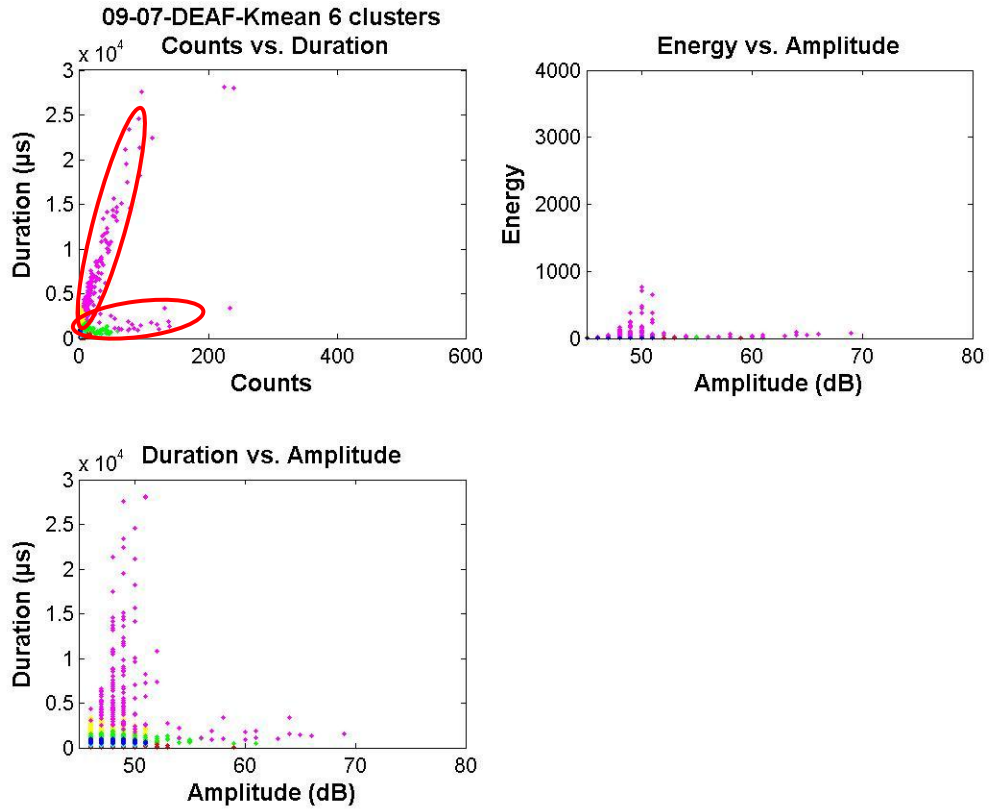


Figure 35 September 7th - 6 Clusters K-mean Clustering AE Plots

Table 16 September 7th – 6 Clusters K-mean Statistical Analysis (a)Amplitude (b)Duration (c)Energy

Amplitude							Duration					
	Min	Max	Average	SD	# of Hits	Rank	Min	Max	Average	SD	# of Hits	Rank
	45	51	46.14	0.97	4932	6	467	1067	661.74	111.51	4932	4
	45	51	46.72	1.06	2916	3	902	3766	1491.02	447.90	2916	2
	45	61	46.93	1.26	2702	2	442	1970	837.03	186.70	2702	3
	45	51	46.27	1.01	4145	5	225	656	407.45	76.49	4145	5
	46	69	49.88	3.97	184	1	948	97406	9542.17	13211.20	184	1
	45	59	46.36	1.16	3644	4	45	487	212.28	60.83	3644	6

Energy						
	Min	Max	Average	SD	# of Hits	Rank
	2	6	3.73	0.72	4932	4
	6	23	9.23	3.05	2916	2
	3	17	5.29	1.54	2702	3
	1	4	2.21	0.50	4145	5
	16	772	80.39	113.59	184	1
	1	5	1.11	0.37	3644	6

### **9.2.1.2. October 18th K-mean**

Figures 36-38 and Tables 17-19 show the AE plots and the statistical analysis performed on October 18th using the k-mean clustering algorithm. Again, cluster criteria have provided with a possibility that 6 clusters in the data set are optimal. To verify that having 6 clusters in the data set is optimal, 4 through 6 clusters were checked.

The verification must be done by observing Figure 36 and Table 17 at the same time. From Table 4, the plastic deformation should have the lowest amplitude, shortest duration, and lowest energy. By observing Table 17, “red” color in the AE plot indicates that it satisfies the characteristics of the plastic deformation. Therefore, the red cluster can be concluded to be plastic deformation. Next, the “yellow” shows it has longest duration and the rest in the medium. This indicates that the yellow cluster is noise. Thirdly, the “green” indicates that it satisfies the acoustic emission characteristics of the fatigue cracking. Finally, the “blue” cluster indicates the two different average frequency form the counts versus duration plot. This indicates that there is a possibility of two different clusters in the blue cluster. From these reasons, there is a possibility of cluster number increasing.

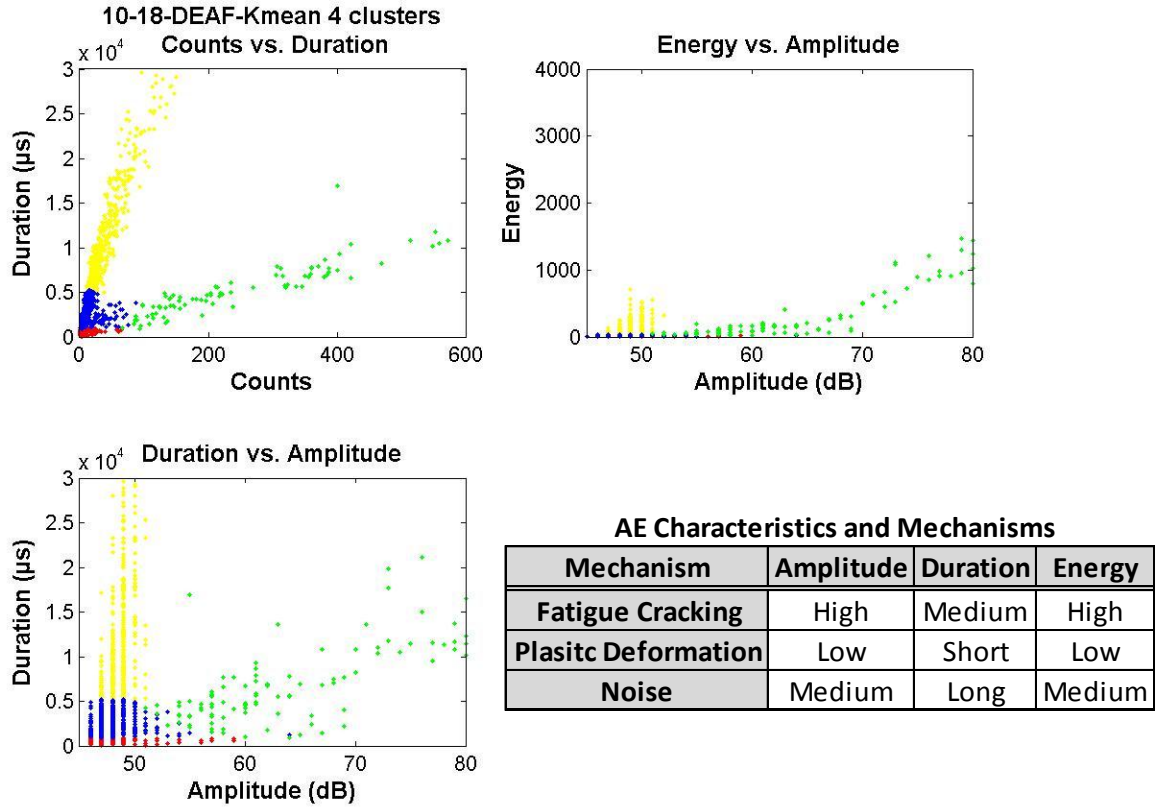


Figure 36 October 18th - 4 Clusters K-mean Clustering AE Plots

Table 17 October 18th – 4 Clusters K-mean Statistical Analysis (a)Amplitude (b)Duration (c)Energy

Amplitude							Duration					
	Min	Max	Average	SD	# of Hits	Rank	Min	Max	Average	SD	# of Hits	Rank
	45	59	47.30	2.06	352	4	78	1142	509.97	248.35	352	4
	47	52	48.74	0.92	444	2	4948	82363	14227.57	10942.41	444	1
	45	64	47.84	1.45	625	3	922	5244	2544.28	1174.80	625	3
	51	94	67.14	11.36	124	1	869	48491	8517.85	6430.69	124	2

Energy						
	Min	Max	Average	SD	# of Hits	Rank
	1	20	4.02	2.59	352	4
	40	713	119.86	95.77	444	2
	6	43	20.68	9.70	625	3
	26	5528	714.07	1096.97	124	1

By comparing the 4 cluster and 5 cluster AE plots, it is possible to observe that the “blue” cluster from the 4 cluster plot has separated in to two different clusters in the Figure 37. Again, by observing Figure 37

and statistical analysis on Table 18, it is possible to distinguish each cluster with each mechanism. The “red” cluster has lowest amplitude, duration, and energy, indicating it is possibly a plastic deformation. Secondly, looking at the “yellow” clusters, it has high duration, mid – high amplitude, and energy, indicating as one type of noise. In addition, the “blue” cluster can be considered to be noise as well from the AE characteristic of the noise signal. Looking at “pink” cluster, it has high amplitude and energy, with medium duration. This indicates that the “pink” cluster is one type of fatigue crack.

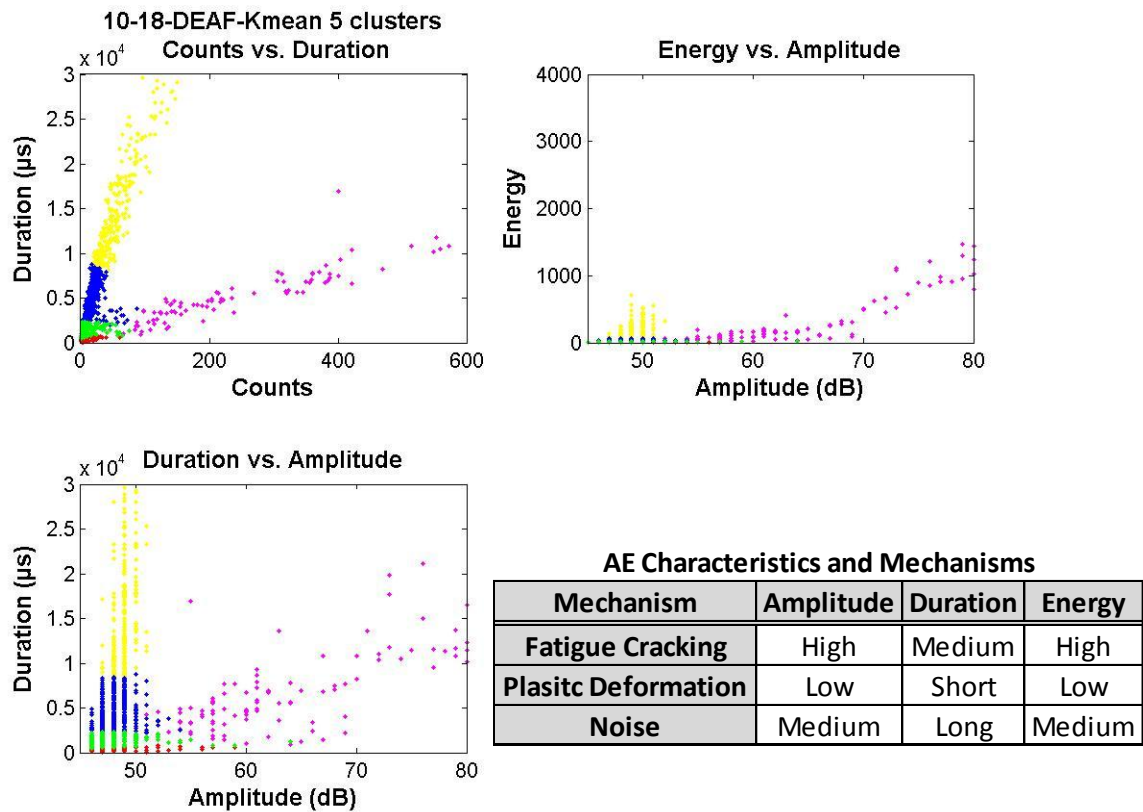


Figure 37 October 18th - 5 Clusters K-mean Clustering AE Plots

**Table 18 October 18th – 5 Clusters K-mean Statistical Analysis (a)Amplitude (b)Duration (c)Energy**

Amplitude							Duration					
	Min	Max	Average	SD	# of Hits	Rank	Min	Max	Average	SD	# of Hits	Rank
	51	94	67.14	11.36	124	1	869	48491	8517.85	6430.69	124	2
	45	59	47.39	2.26	214	5	78	797	345.56	149.58	214	5
	47	52	48.99	0.85	283	2	8125	82363	18597.29	11608.76	283	1
	46	54	48.13	1.07	473	3	2013	8773	4555.40	1673.46	473	3
	45	64	47.50	1.69	451	4	495	2333	1319.37	483.97	451	4

Energy						
	Min	Max	Average	SD	# of Hits	Rank
	26	5528	714.07	1096.97	124	1
	1	17	2.79	2.07	214	5
	68	713	157.72	102.01	283	2
	17	70	37.11	13.61	473	3
	3	27	10.58	4.41	451	4

Figure 38 shows the 6 cluster plots by the k-mean algorithm. From observing Table 19 it can be observed that the “red” clusters is plastic deformation since it has lowest amplitude, duration, and energy. The “light blue” and “yellow” clusters indicate having the very close AE characteristics of noise; the two clusters can be classified as noise cluster. In observing the AE characteristic of “green” and “pink” clusters, both shows the characteristic of fatigue crack. As the cluster criteria have suggested, the 6 cluster for k-mean algorithm shows a decent clustering. However, there is a possibility that some point in “blue” cluster could be from the “yellow” cluster by observing the similar value of the amplitude in two clusters. Therefore, the number of cluster in this data set could be 5. However, if we observe the 5 cluster AE plot, the “green” cluster does not have the distinct characteristics of the fatigue crack or noise. Due to this reason, 5 clusters in the data set is optimal in the data set, but the k-mean clustering algorithms may not have the capabilities to classify the data well enough to be analyzed for failure mechanisms inspection.

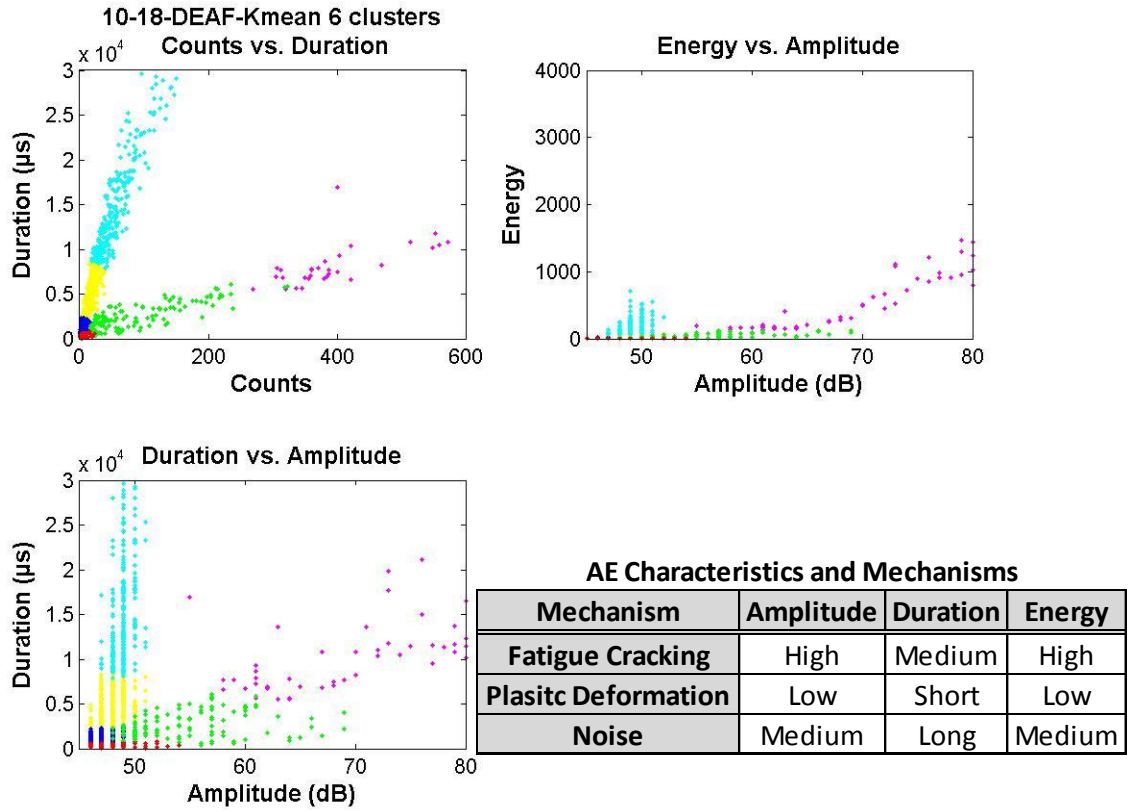


Figure 38 October 18th - 6 Clusters K-mean Clustering AE Plots

**Table 19 October 18th – 6 Clusters K-mean Statistical Analysis (a)Amplitude (b)Duration (c)Energy**

Amplitude							Duration					
	Min	Max	Average	SD	# of Hits	Rank	Min	Max	Average	SD	# of Hits	Rank
	55	94	74.11	10.13	70	1	5582	48491	12341.33	6186.35	70	2
	46	51	48.02	0.89	456	4	2219	8773	4603.88	1667.20	456	3
	45	50	47.15	0.93	413	5	495	2327	1283.58	483.31	413	5
	48	69	54.43	4.90	122	2	347	6125	2560.48	1420.88	122	4
	45	54	47.06	1.46	200	6	78	797	331.26	142.54	200	6
	47	52	48.99	0.86	284	3	7893	82363	18559.60	11605.63	284	1

Energy						
	Min	Max	Average	SD	# of Hits	Rank
	133	5528	1211.31	1252.38	70	1
	17	69	37.24	13.72	456	4
	3	18	9.87	3.98	413	5
	7	129	41.48	32.01	122	3
	1	7	2.44	1.29	200	6
	68	713	157.42	101.96	284	2

## 9.2.2. Gaussian Mixture Model

### 9.2.2.1. September 7th Gaussian Mixture Model

Figures 39-41 and Tables 20-22 show the AE plots and the statistical analysis performed on September 7th using the Gaussian Mixture Model clustering algorithm. From the cluster criteria, the analysis has provided a possibility that 6 clusters in the September 7th data are the optimal number of clusters.

By visually observing the Figures 33-35, it is possible to detect that one of the clusters is extending in two directions from the duration versus counts plot. This indicates that there are two different average frequencies in one cluster. Due to this reason, it is possible to say that the Gaussian Mixture Model clustering algorithms were not able to classify the September 9th data into definite clusters. The cluster with two different average frequencies is circled in red.

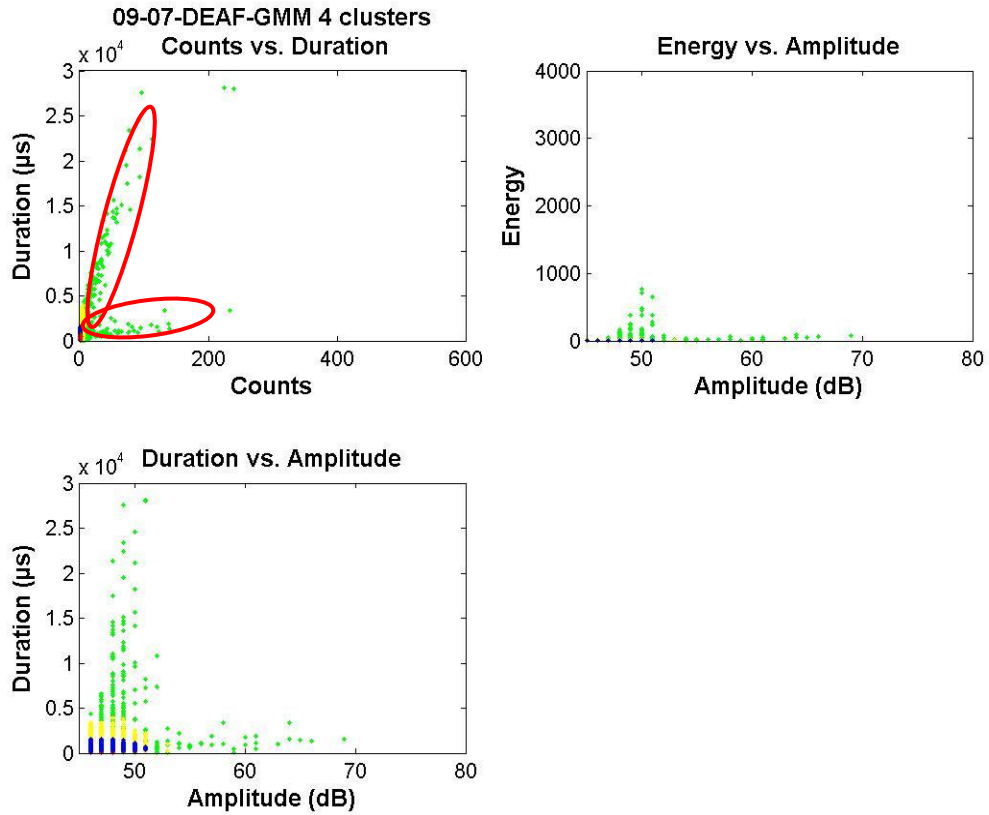


Figure 39 September 7th - 4 Clusters GMM Clustering AE Plots

Table 20 September 7th – 4 Clusters GMM Statistical Analysis (a)Amplitude (b)Duration (c)Energy

Amplitude							Duration					
	Min	Max	Average	SD	# of Hits	Rank	Min	Max	Average	SD	# of Hits	Rank
	45	69	48.66	2.94	591	1	45	97406	3407.99	8458.13	591	1
	45	50	46.08	0.95	2763	4	77	656	218.15	60.30	2763	4
	45	51	46.14	0.96	8419	3	211	1592	605.74	224.53	8419	3
	45	53	46.82	1.05	6750	2	128	3906	944.40	559.35	6750	2

Energy						
	Min	Max	Average	SD	# of Hits	Rank
	1	772	28.65	72.35	591	1
	1	1	1.00	0.00	2763	4
	2	10	3.40	1.39	8419	3
	1	26	5.81	3.59	6750	2



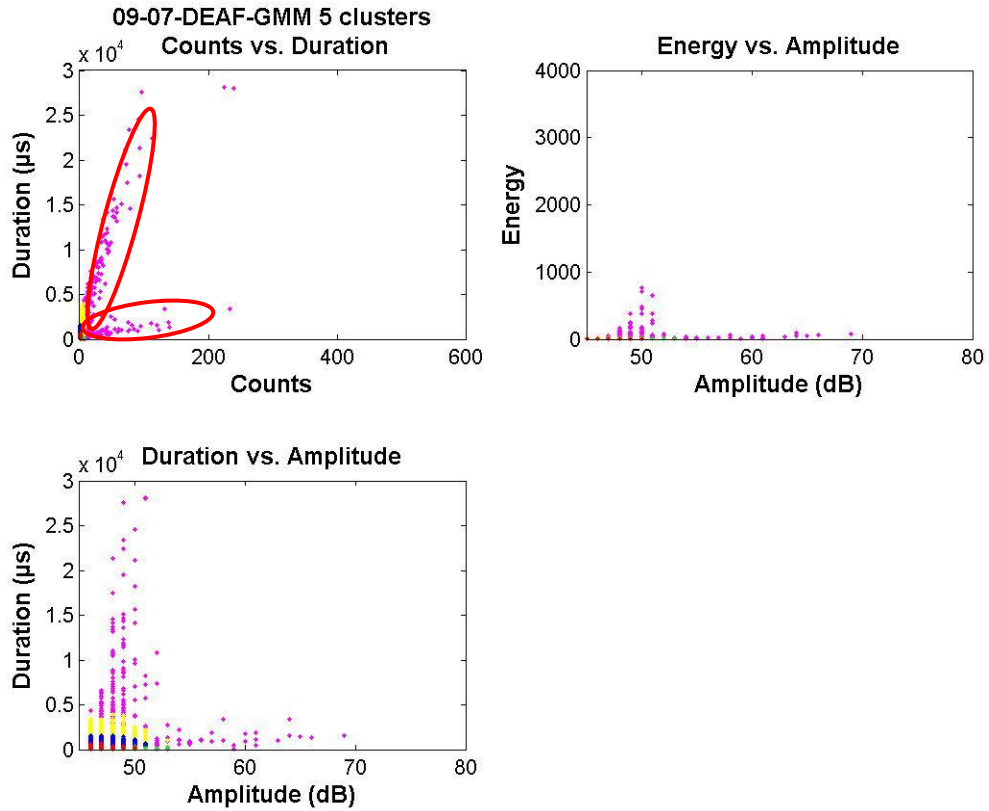


Figure 40 September 7th - 5 Clusters GMM Clustering AE Plots

Table 21 September 7th – 5 Clusters GMM Statistical Analysis (a)Amplitude (b)Duration (c)Energy

Amplitude							Duration					
	Min	Max	Average	SD	# of Hits	Rank	Min	Max	Average	SD	# of Hits	Rank
	45	50	46.08	0.95	2762	5	77	656	218.05	60.07	2762	4
	45	53	46.82	1.04	6289	3	201	3906	1003.06	538.45	6289	2
	45	51	46.14	0.96	8419	4	211	1592	605.74	224.53	8419	3
	45	53	46.95	1.18	666	2	61	496	179.36	60.71	666	5
	46	69	49.39	3.29	387	1	45	97406	5068.73	10062.94	387	1

Energy						
	Min	Max	Average	SD	# of Hits	Rank
	1	1	1.00	0.00	2762	4
	2	26	6.20	3.45	6289	2
	2	10	3.40	1.39	8419	3
	1	1	1.00	0.00	666	4
	1	772	42.68	86.18	387	1

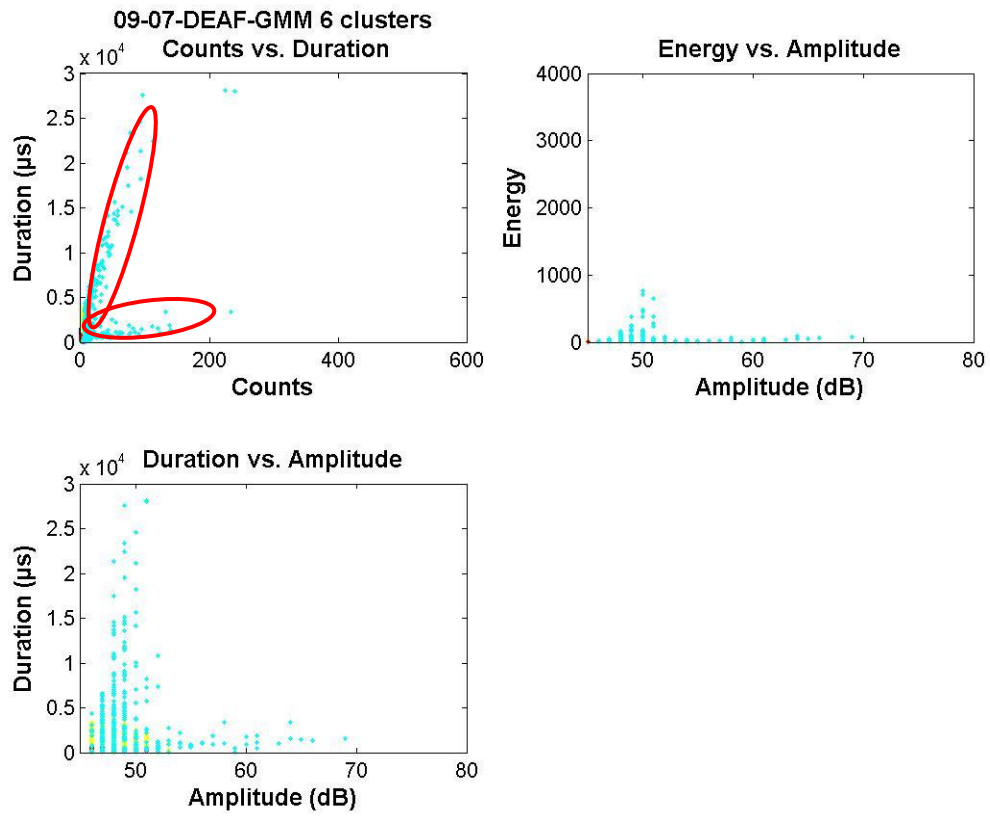


Figure 41 September 7th - 6 Clusters GMM Clustering AE Plots

**Table 22 September 7th – 6 Clusters GMM Statistical Analysis (a)Amplitude (b)Duration (c)Energy**

Amplitude							Duration						
	Min	Max	Average	SD	# of Hits	Rank		Min	Max	Average	SD	# of Hits	Rank
	45	51	46.12	0.97	2604	5		211	789	384.59	65.11	2604	4
	45	50	46.08	0.95	2762	6		77	656	218.05	60.07	2762	5
	45	51	46.14	0.95	5815	4		345	1592	704.78	198.44	5815	3
	45	53	46.82	1.04	6288	3		201	3906	1003.07	538.49	6288	2
	45	53	46.95	1.18	666	2		61	496	179.36	60.71	666	6
	46	69	49.39	3.29	388	1		45	97406	5058.09	10052.1	388	1

Energy						
	Min	Max	Average	SD	# of Hits	Rank
	2	2	2.00	0.00	2604	4
	1	1	1.00	0.00	2762	5
	3	10	4.03	1.23	5815	3
	2	26	6.20	3.45	6288	2
	1	1	1.00	0.00	666	5
	1	772	42.59	86.09	388	1

**9.2.2.2. October 18th Gaussian Mixture Model**

Figures 42-44 and Tables 23-25 show the AE plots and the statistical analysis performed on October 18th using the Gaussian Mixture Model clustering algorithm. From the cluster criteria, the analysis has provided a possibility of 5 clusters in the October 18th data being the optimal number of clusters.

From observing Figure 42 and Table 23, it is possible to match each cluster with the mechanisms. For example, the “red” cluster has the lowest amplitude, duration, and energy, indicating that it is possibly a plastic deformation signal. The “yellow” cluster can be considered as one type of noise since it has medium amplitude and energy with long duration. Examining the “green” and “blue” clusters with the three acoustic emission characteristics, it is possible to conclude that both is two different types of fatigue crack. From the observation, the 4 clusters in this data show good clustering. However, since the criterion has suggested that 5 clusters is the optimal number of cluster, it is important to check and observe those plots and statistical analysis as well, as some of the clusters could be separated even further.

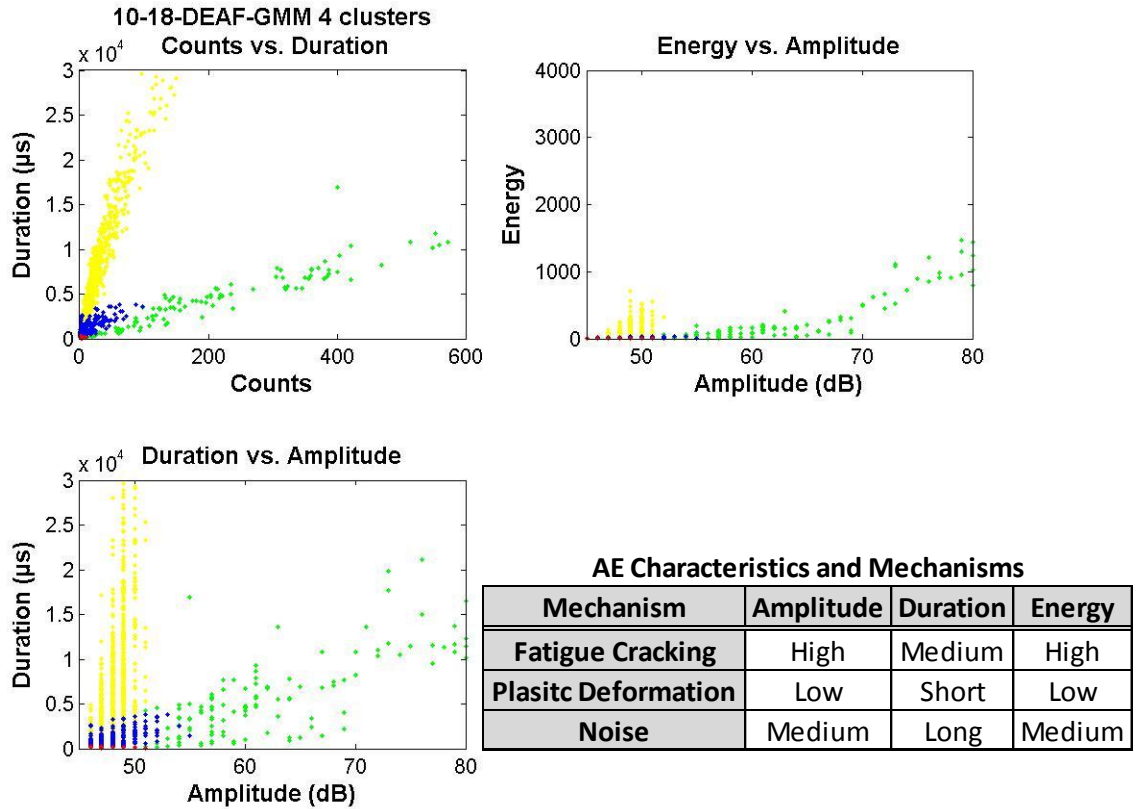


Figure 42 October 18th - 4 Clusters GMM Clustering AE Plots

Table 23 October 18th – 4 Clusters GMM Statistical Analysis (a)Amplitude (b)Duration (c)Energy

Amplitude							Duration					
	Min	Max	Average	SD	# of Hits	Rank	Min	Max	Average	SD	# of Hits	Rank
	45	55	47.65	1.63	354	3	180	3794	988.19	739.78	354	3
	50	94	65.55	11.39	143	1	145	48491	7448.52	6579.30	143	2
	45	52	48.05	1.12	989	2	358	82363	7803.22	9393.60	989	1
	45	51	46.68	1.25	59	4	78	281	179.59	52.12	59	4
Energy												
	Min	Max	Average	SD	# of Hits	Rank						
	2	42	8.40	7.52	354	3						
	2	5528	620.44	1048.78	143	1						
	3	713	65.06	81.37	989	2						
	1	1	1.00	0.00	59	4						

Figure 43 and Table 24 show the AE plots and the statistical analysis of the October 18th data with 5 clustering using GMM. Again, by performing a comparison between each cluster with the statistical analysis results, it is possible to distinguish the failure mechanisms for each cluster. Signal characteristics matches with the characteristics of each cluster, and therefore shows that “red” is plastic deformation, “yellow” and “blue are noise, while “green” and “pink” are two different types of fatigue cracking. Comparing Figure 42 and 43, it can be seen that the fatigue crack from the Figure 42 has separated further. This separation has indicated the more detailed and precise difference between the data point characteristics within the cluster.

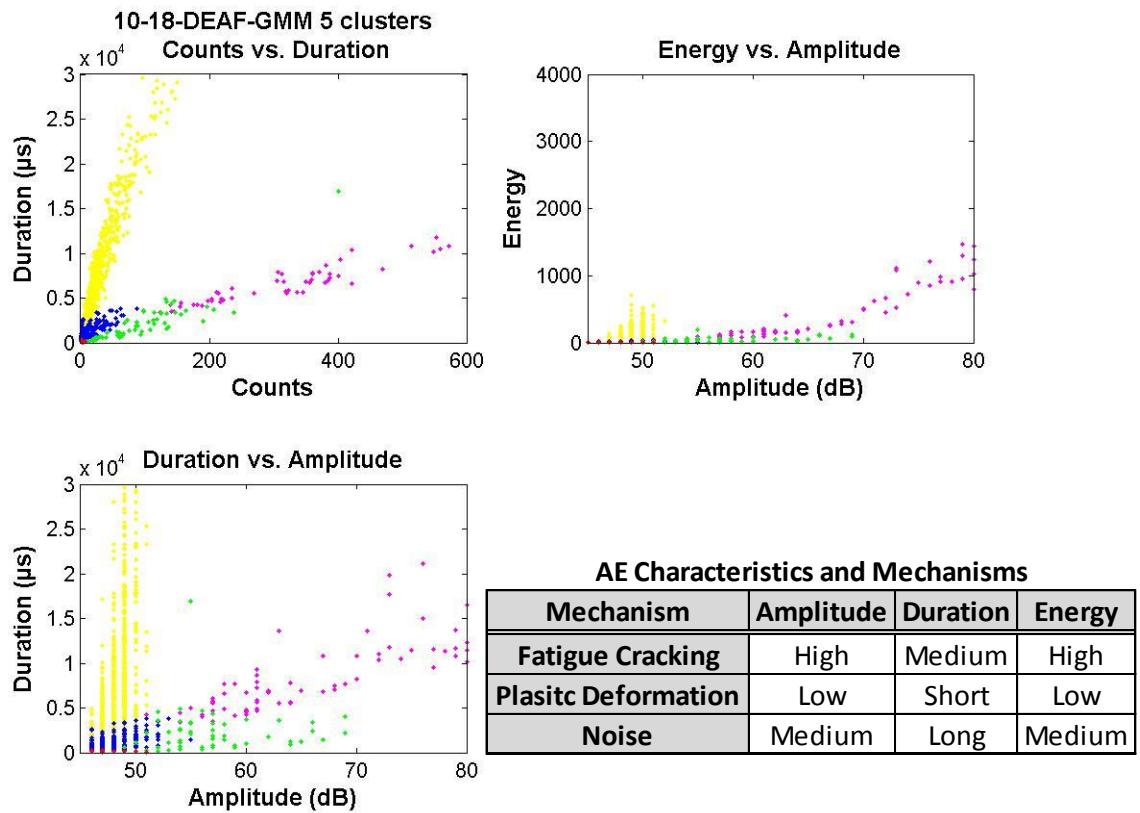


Figure 43 October 18th - 5 Clusters GMM Clustering AE Plots

**Table 24 October 18th – 5 Clusters GMM Statistical Analysis (a)Amplitude (b)Duration (c)Energy**

Amplitude							Duration						
	Min	Max	Average	SD	# of Hits	Rank		Min	Max	Average	SD	# of Hits	Rank
	45	52	48.08	1.10	970	3		387	82363	7939.24	9433.93	970	2
	51	94	70.04	11.25	90	1		3364	21824	10013.31	4894.58	90	1
	45	51	46.62	1.18	58	5		78	281	180.52	52.09	58	5
	45	55	47.54	1.57	368	4		180	3794	967.43	717.49	368	4
	49	91	57.20	6.60	59	2		126	48491	2952.47	6481.07	59	3

Energy						
	Min	Max	Average	SD	# of Hits	Rank
	3	713	66.21	81.74	970	3
	48	5528	909.33	1136.00	90	1
	1	1	1.00	0.00	58	5
	2	40	8.08	7.10	368	4
	1	4669	118.75	603.69	59	2

Figure 44 and Table 25 shows the AE plots and the statistical analysis of the October 18th data with 6 clustering using GMM. When trying to determine which cluster is plastic deformation, the “red” cluster indicates having the lowest duration and energy but not amplitude. On the other hand, the “blue” cluster has the lowest amplitude but second to lowest duration and energy. From this reason, it is possible to state that the two clusters were forced to separate by the clustering algorithm. The other mechanisms can be assigned to the clusters as “yellow” and “light blue” identified as noise and the “green” and “pink” as fatigue crack from the AE characteristics of the each clusters. From the visual verification, the Gaussian Mixture Model was capable of clustering this data set into 5 clusters, as it was suggested by the clustering criterions.

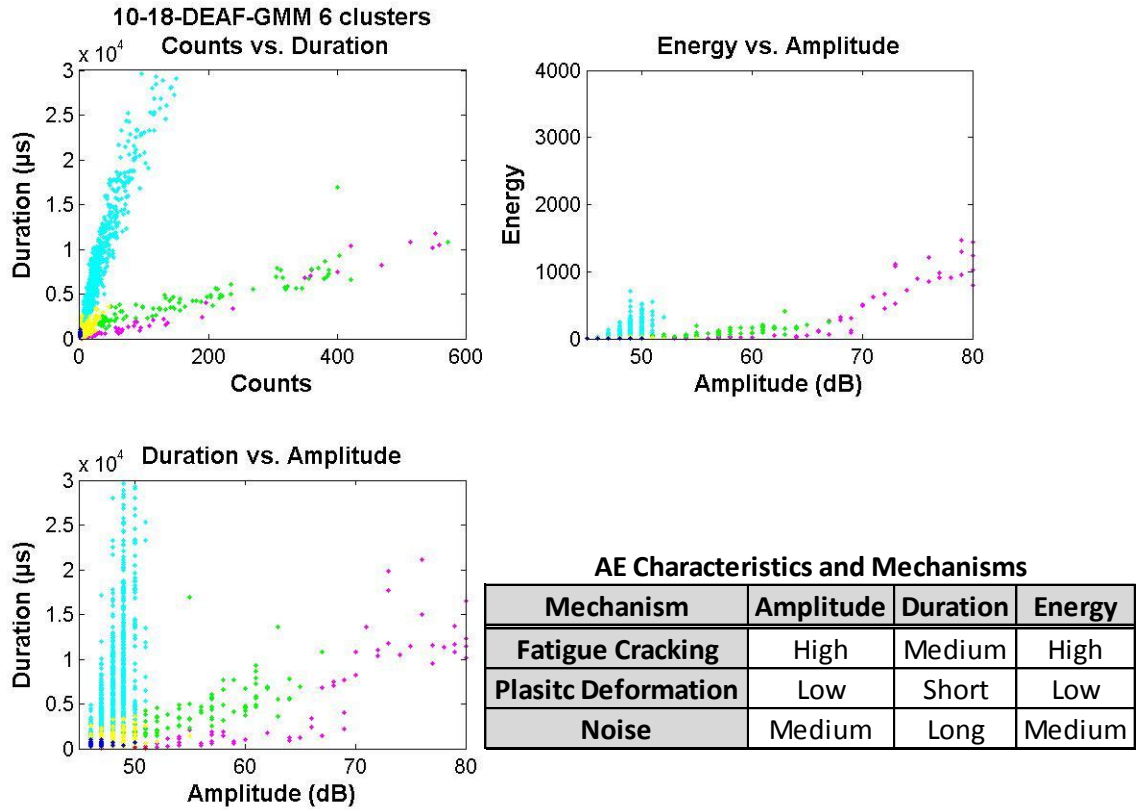


Figure 44 October 18th - 6 Clusters GMM Clustering AE Plots

**Table 25 October 18th – 6 Clusters GMM Statistical Analysis (a)Amplitude (b)Duration (c)Energy**

Amplitude							Duration						
	Min	Max	Average	SD	# of Hits	Rank		Min	Max	Average	SD	# of Hits	Rank
	45	50	46.34	0.86	76	6		248	1033	559.43	193.01	76	5
	49	94	69.97	12.71	86	1		126	48491	8556.22	8096.92	86	1
	45	52	48.10	1.09	963	3		618	82363	7992.80	9447.25	963	2
	45	51	46.58	1.15	57	5		78	281	181.19	52.29	57	6
	48	67	55.84	4.66	91	2		1143	16931	4409.44	2631.30	91	3
	45	55	47.53	1.27	272	4		180	3589	940.91	643.26	272	4

Energy						
	Min	Max	Average	SD	# of Hits	Rank
	2	7	3.68	1.30	76	5
	1	5528	952.87	1246.78	86	1
	5	713	66.66	81.87	963	3
	1	1	1.00	0.00	57	6
	14	405	82.86	68.39	91	2
	2	33	7.66	5.64	272	4

### 9.2.3. Kohonen SOM

#### 9.2.3.1. September 7th Kohonen SOM

Figures 45-47 and Tables 26-28 show the AE plots and the statistical analysis performed on September 7th using the Kohonen Self-organizing Map clustering algorithm. From the cluster criteria, the analysis has suggested a possibility of 4 clusters in the September 7th data as the optimal number of clusters.

Comparing the September 7th result by Kohonen SOM with other clustering algorithms, it can be clearly seen that each of the clusters are not extending in two directions from the duration versus counts plot. This indicates that Kohonen SOM was successful in clustering the data points into different clusters.

Figure 45 and Table 26 show the AE plot and statistical analysis of the 4 clustering of the September 7th data set. From the observation of AE characteristics of the each cluster, it shows that the “red” is plastic deformation, “yellow” as one of the fatigue cracks, and “green” and “blue” as two different types of noise. To further investigate on the number of clusters, 5 and 6 cluster plots were investigated.



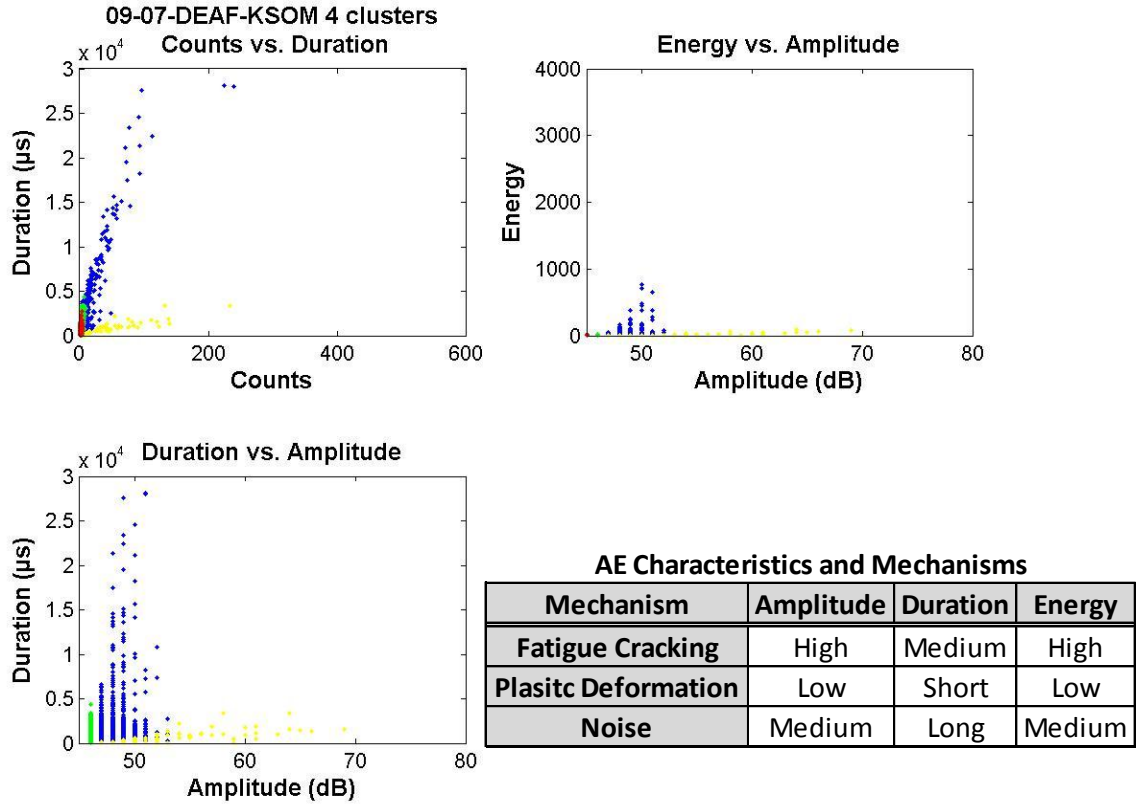


Figure 45 September 7th – 4 Clusters Kohonen SOM Clustering AE Plots

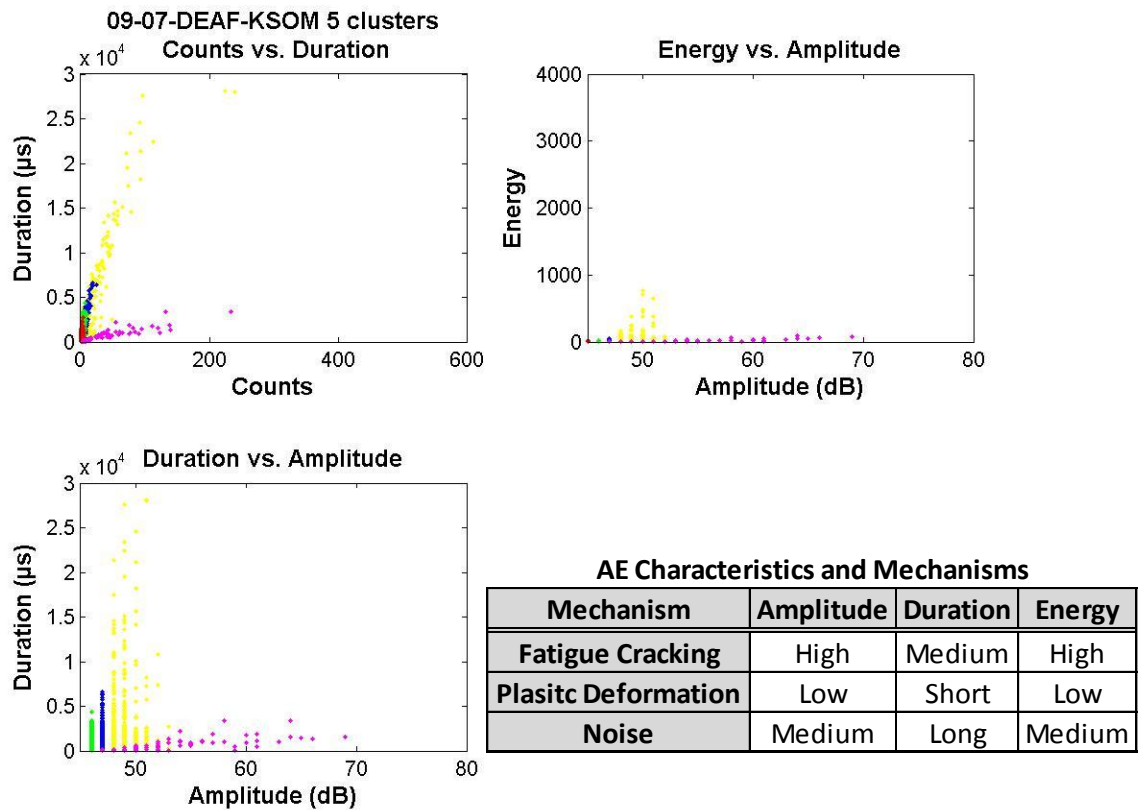
Table 26 September 7th – 4 Clusters KSOM Statistical Analysis (a)Amplitude (b)Duration (c)Energy

Amplitude							Duration						
	Min	Max	Average	SD	# of Hits	Rank		Min	Max	Average	SD	# of Hits	Rank
	45	45	45.00	0.00	3104	4		88	2724	517.03	286.13	3104	4
	46	46	46.00	0.00	8236	3		75	4352	649.28	428.71	8236	2
	47	53	47.54	0.86	7086	2		76	97406	999.82	2592.95	7086	1
	47	69	52.59	4.97	97	1		45	3417	559.30	656.11	97	3

Energy						
	Min	Max	Average	SD	# of Hits	Rank
	1	16	2.849871134	1.77	3104	4
	1	28	3.791767848	2.76	8236	3
	1	772	6.551368896	22.05	7086	2
	1	87	11.34020619	17.05	97	1

Figure 46 and Table 27 show the AE plot and statistical analysis of the 5 clustering of September 7th data set. Comparing it with the Figure 45 with 4 clusters, it can be observed that some of the noise cluster from Figure 45 has separated. From the Table 27, it indicates that the “red” cluster has the lowest amplitude, duration, and energy, showing it as the plastic deformation. The “yellow” and “blue” clusters have the AE characteristics of noise. The “pink” cluster can be identified as fatigue crack since it has high amplitude, energy, and medium duration.



**Figure 46 September 7th – 5 Clusters Kohonen SOM Clustering AE Plots**

**Table 27 September 7th – 5 Clusters KSOM Statistical Analysis (a)Amplitude (b)Duration (c)Energy**

Amplitude							Duration						
	Min	Max	Average	SD	# of Hits	Rank		Min	Max	Average	SD	# of Hits	Rank
	45	45	45.00	0.00	3104	5		88	2724	517.03	286.13	3104	5
	46	46	46.00	0.00	8231	4		75	4352	649.62	428.62	8231	3
	46	47	47.00	0.03	4574	3		77	6617	774.64	593.65	4574	2
	48	53	48.52	0.76	2520	2		76	97406	1405.77	4244.53	2520	1
	47	69	52.69	5.01	94	1		45	3417	570.98	663.17	94	4

Energy						
	Min	Max	Average	SD	# of Hits	Rank
	1	16	2.85	1.77	3104	5
	1	28	3.79	2.76	8231	4
	1	52	4.71	4.11	4574	3
	1	772	9.88	36.33	2520	2
	1	87	11.64	17.23	94	1

Figure 47 and Table 28 show the AE plot and statistical analysis of the 6 clustering of September 7th data set. It can be observed that the noise cluster again separated into different clusters. Also, by comparing the 5 cluster and the 6 cluster, it can be observed that the “blue” cluster from 6 cluster data shows fewer hits compared to the other clusters. One may argue that the “pink” also has low number of hits, but observing Table 27, the “pink” cluster in Table 27, and “pink” cluster in Table 28 show exactly same value for the minimum and maximum as well as approximately same value for the average and number of hits. Due to this observation, it is possible to state that the “yellow” cluster in Table 27 and “pink” cluster in Table 28 are same clusters. Therefore, 6 cluster in the data set is not an optimal number of clusters since the “blue” cluster in Figure 27 was forced to separate into different cluster. By observing 3-6 cluster outputs, the Kohonen SOM was successful in clustering the data in to 5 clusters for the September 7th data set.

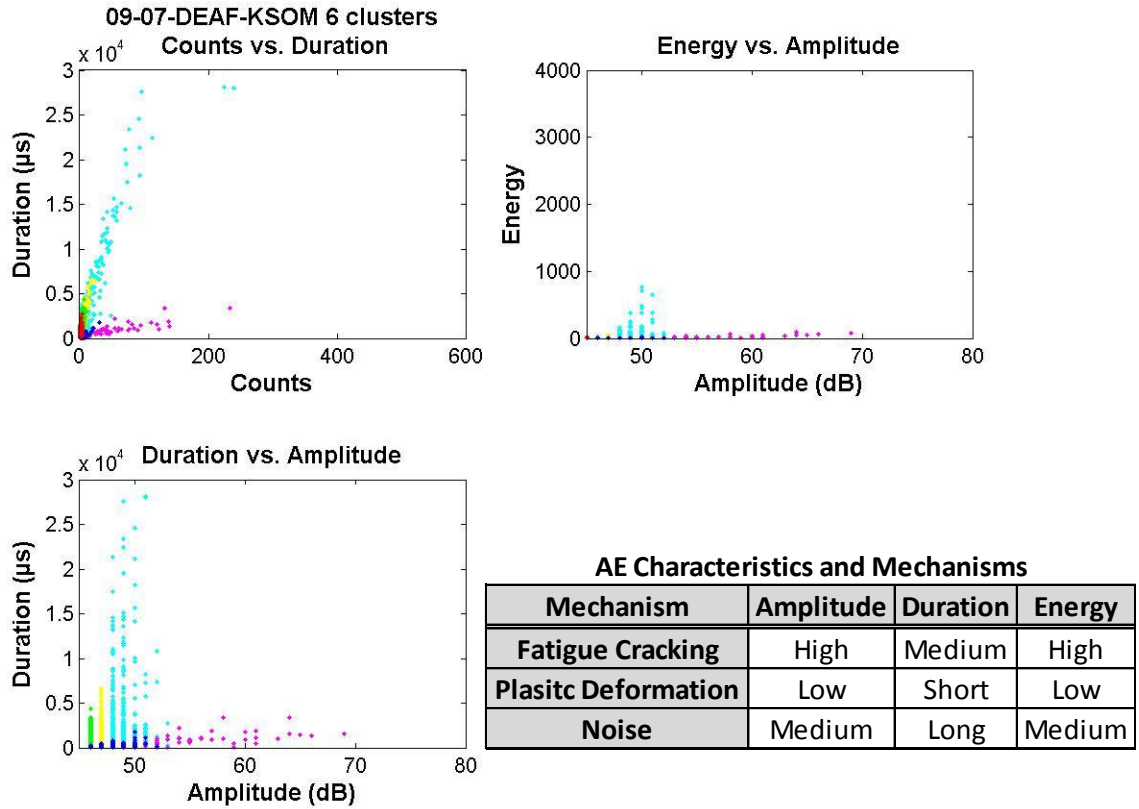


Figure 47 September 7th – 6 Clusters Kohonen SOM Clustering AE Plots

Table 28 September 7th – 6 Clusters KSOM Statistical Analysis (a)Amplitude (b)Duration (c)Energy

Amplitude							Duration					
	Min	Max	Average	SD	# of Hits	Rank	Min	Max	Average	SD	# of Hits	Rank
	45	45	45.00	0.00	3102	6	88	2724	517.30	286.03	3102	5
	46	46	46.00	0.00	8123	5	81	4352	656.42	427.33	8123	4
	45	52	47.21	1.15	408	3	61	1735	185.15	149.80	408	6
	47	47	47.00	0.00	4406	4	97	6617	797.71	592.58	4406	2
	48	53	48.52	0.76	2415	2	88	97406	1455.38	4328.71	2415	1
	47	69	54.12	5.09	69	1	45	3417	718.64	717.16	69	3

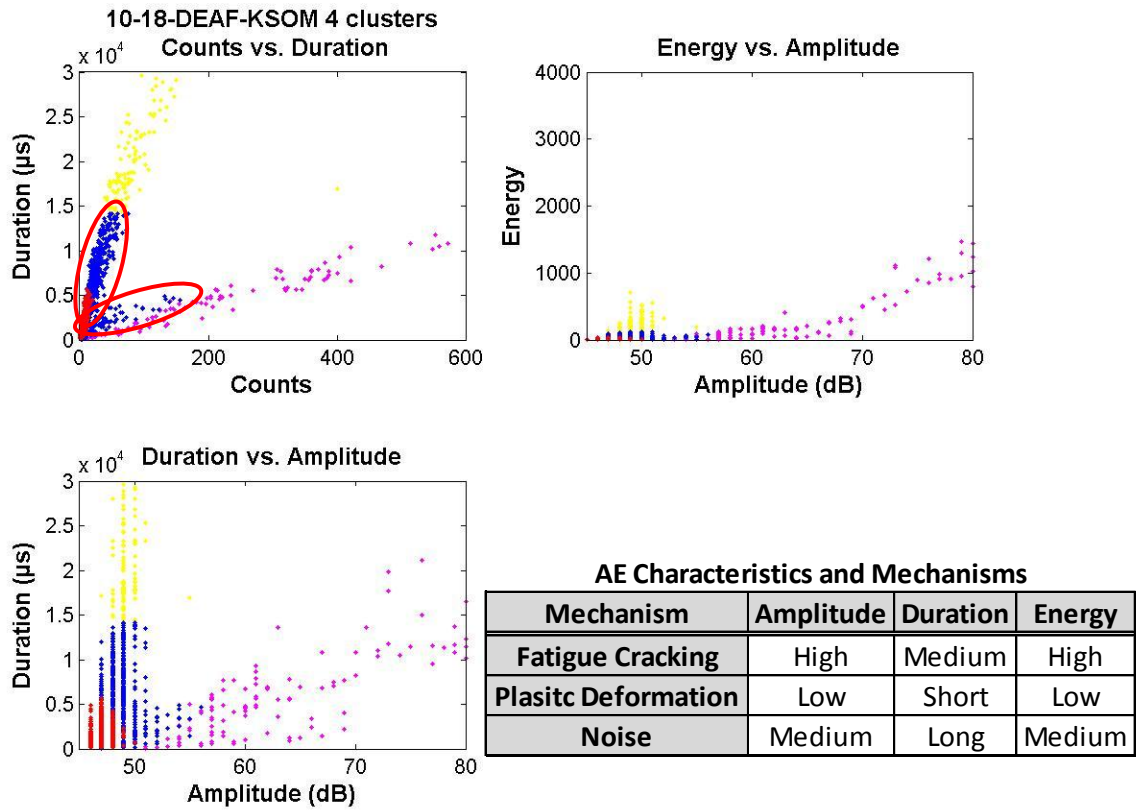
  

Energy						
	Min	Max	Average	SD	# of Hits	Rank
	1	16	2.85	1.77	3102	5
	1	28	3.83	2.76	8123	4
	1	19	1.49	1.54	408	6
	1	52	4.84	4.12	4406	3
	1	772	10.21	37.07	2415	2
	1	87	15.23	18.88	69	1

**9.2.3.2. October 18th Kohonen SOM**

Figures 48-50 and Tables 29-31 show the AE plots and the statistical analysis performed on October 18th using the Kohonen Self-organizing Map clustering algorithm. From the cluster criteria, the analysis has suggested that 5 clusters in the September 7th data is the optimal number of clusters.

From Figure 48, it is possible to detect that one of the clusters is extending in two directions from the duration versus counts plot circled in red. This indicates that there are two different average frequencies in one cluster. Therefore, the 4 clustering for October 18th is not an optimal number of clusters in the data set.



**Figure 48 October 18th – 4 Cluster Kohonen SOM Clustering AE Plots**

**Table 29 October 18th – 4 Clusters KSOM Statistical Analysis (a)Amplitude (b)Duration (c)Energy**

Amplitude							Duration						
	Min	Max	Average	SD	# of Hits	Rank		Min	Max	Average	SD	# of Hits	Rank
	49	94	66.17	11.437	135	1		78	48491	7501.06	6721.04	135	2
	47	55	49.39	0.899	142	2		14367	82363	26357.57	12055.00	142	1
	46	56	48.74	1.253	571	3		88	14165	5798.16	3612.56	571	3
	45	50	46.97	0.795	697	4		116	5634	1544.86	1158.55	697	4

Energy						
	Min	Max	Average	SD	# of Hits	Rank
	1	5528	652.06	1071.24	135	1
	119	713	225.91	106.22	142	2
	1	122	48.48	29.71	571	3
	1	41	11.93	9.24	697	4

Figure 49 and Table 30 show the AE plot and statistical analysis of the 5 clustering of October 18th data. By performing the same analysis of observing the mechanism characteristics with each cluster, it shows a distinct difference in each. Since the “red” cluster has the lowest amplitude, duration, and energy, it is plastic deformation cluster. The “blue” and “yellow” clusters have similar characteristics of the noise. The “pink” and “green” clusters have characteristics of fatigue cracking, so it is considered as two different types within the fatigue crack.

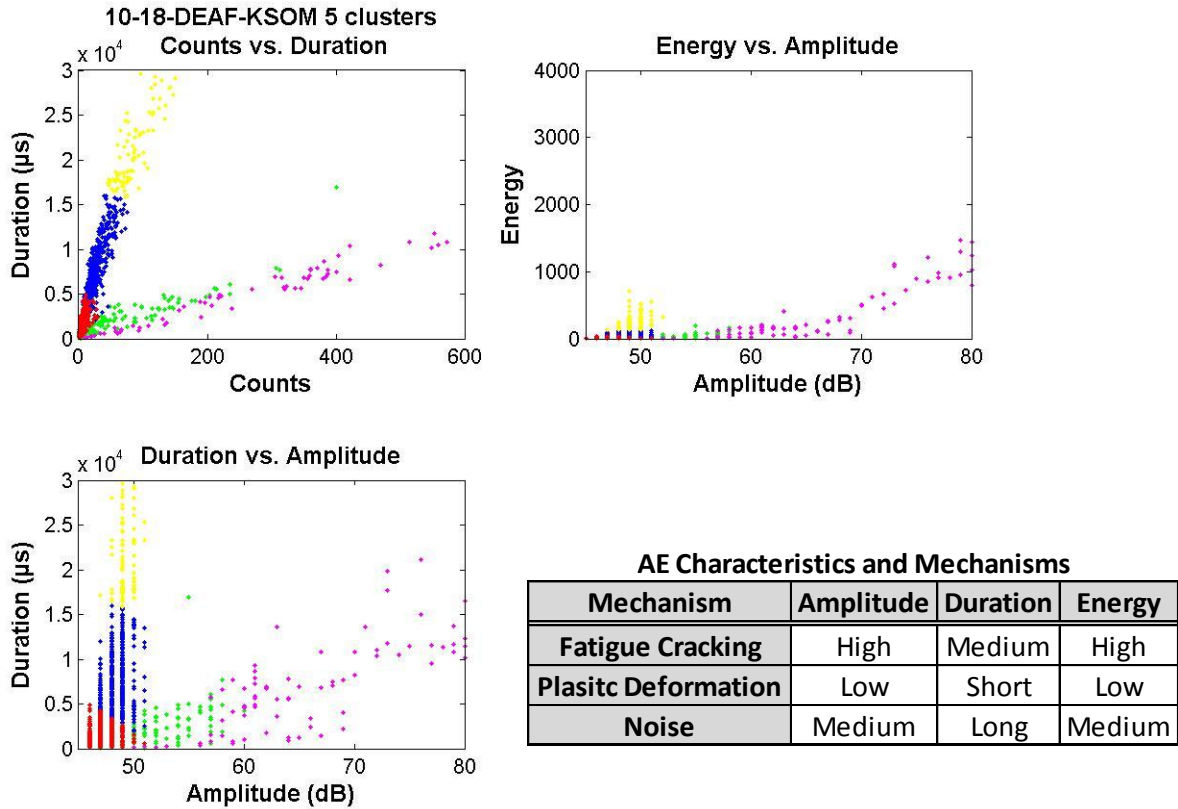


Figure 49 October 18th – 5 Cluster Kohonen SOM Clustering AE Plots

Table 30 October 18th – 5 Clusters KSOM Statistical Analysis (a)Amplitude (b)Duration (c)Energy

Amplitude							Duration						
	Min	Max	Average	SD	# of Hits	Rank		Min	Max	Average	SD	# of Hits	Rank
	45	51	47.09	0.924	724	5		116	4862	1409.52	966.04	724	5
	46	61	51.61	3.397	121	2		78	16931	2120.88	2173.13	121	4
	47	51	48.44	0.841	474	4		1970	15951	7296.68	3238.09	474	3
	47	52	49.40	0.771	126	3		15899	82363	27770.47	12083.20	126	1
	50	94	69.93	10.944	100	1		126	48491	9081.14	7059.88	100	2
Energy													
	Min	Max	Average	SD	# of Hits	Rank							
	1	39	10.96	7.87	724	5							
	1	196	29.72	35.39	121	4							
	16	132	59.96	27.30	474	3							
	131	713	238.08	106.59	126	2							
	1	5528	861.50	1175.45	100	1							

Figure 50 and Table 31 show the AE plot and statistical analysis of the 6 clustering of October 18th data. The “red” cluster in Figure 50 shows lowest amplitude, duration, and energy, indicating it is a plastic deformation. The “yellow” and “light blue” clusters can be identified as noise since it has the AE characteristics of noise. Two types of fatigue crack can be observed from the “green” and “pink” clusters having high amplitude and energy as well as medium duration. Comparing with Figure 49, it can be observed that the “light blue” cluster was introduced by separating the “blue” and “yellow” cluster to Figure 50. Also, the size of the “red” and “blue” clusters between the two figures is different. Observing Figure 50, the “blue” cluster can be concluded to be noise since it has low average frequency compared to other failure mechanisms. However, the average duration of the “blue” cluster does not support this conclusion since it has a value very similar to the “red” and “green” clusters. Moreover, the amplitude and energy characteristic of the “blue” cluster are very similar to the “red” cluster. From this, it can be concluded that the “red” and “blue” are the same cluster showing that 5 clusters in this data set is the optimal number of clusters by using Kohonen SOM.



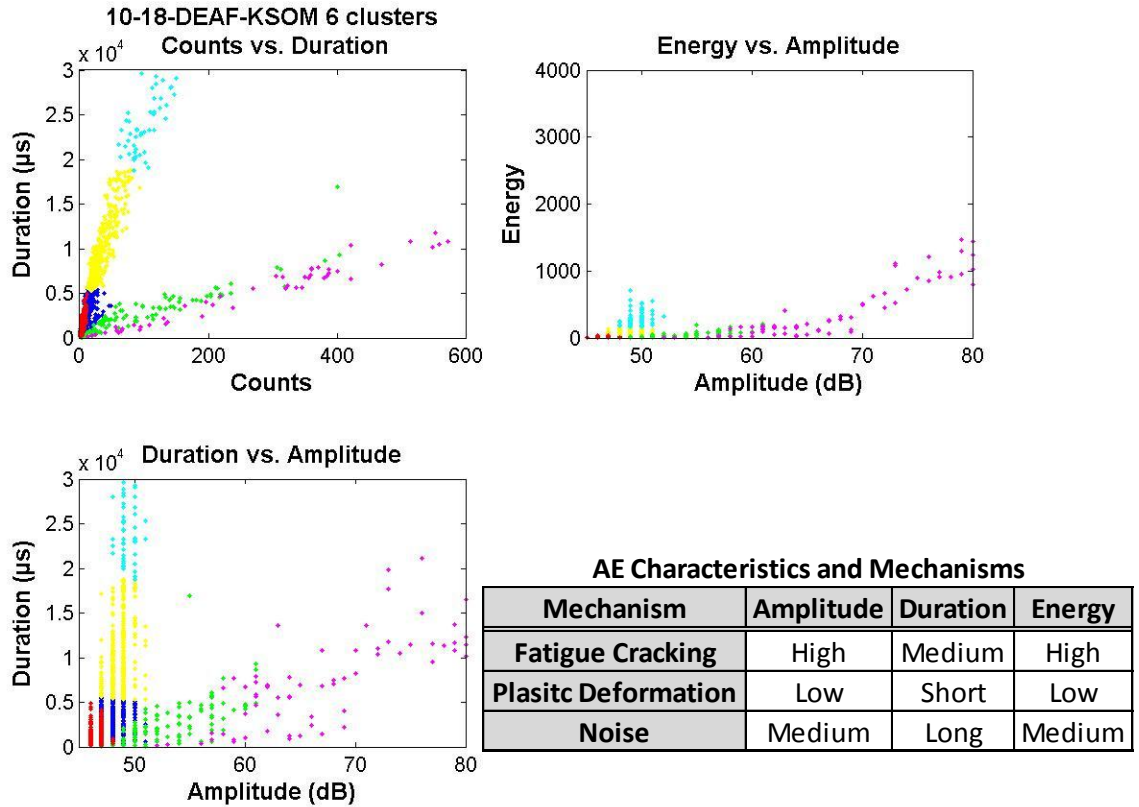


Figure 50 October 18th – 6 Cluster Kohonen SOM Clustering AE Plots

Table 31 October 18th – 6 Clusters KSOM Statistical Analysis (a)Amplitude (b)Duration (c)Energy

Amplitude							Duration						
	Min	Max	Average	SD	# of Hits	Rank		Min	Max	Average	SD	# of Hits	Rank
	46	61	52.97	3.767	109	2		78	16931	2589.67	2446.23	109	5
	45	48	46.58	0.572	460	6		117	4862	1229.94	895.32	460	6
	46	52	48.14	0.826	470	5		116	5676	2681.65	1576.10	470	4
	47	51	48.59	0.830	328	4		5323	18764	10084.30	3675.11	328	2
	48	52	49.54	0.767	90	3		18757	82363	31885.32	12037.06	90	1
	52	94	71.58	10.561	88	1		145	48491	9740.75	7207.31	88	3

Energy						
	Min	Max	Average	SD	# of Hits	Rank
	1	210	40.81	45.07	109	4
	1	39	9.27	7.10	460	6
	1	48	21.84	12.89	470	5
	41	163	83.33	31.46	328	3
	163	713	274.76	105.70	90	2
	2	5528	966.64	1215.95	88	1

### 9.3. Cluster Number Possibility from Criteria and Visual Inspection of 2010 Data

After performing the visual inspection from the suggested optimal number of cluster by the criteria, it is possible to conclude on the number of the cluster that is in the data set. Table 32 shows the result from the visual inspection. Some clustering algorithms could not produce a distinct clustering for September 7th data, but all of the clustering algorithms have indicated that 5 clusters is the optimal number of the cluster in the data set. Also, the Kohonen SOM algorithm was capable of producing a distinct clustering for both of the dates. Therefore, it is possible to conclude that the Kohonen SOM is one of the best clustering algorithms that are capable of producing 5 distinct clustering outputs in the data set of this rotating component.

**Table 32 Number of Clusters by Criterion and Visual Inspection**

Clustering Algorithm	Date	Number of Clusters by Criteria	Number of Clusters by Visual Inspection
K-mean	9/7	6	---
	10/18	6	5
GMM	9/7	6	---
	10/18	5	5
KSOM	9/7	4	5
	10/18	5	5

### 9.4. Fatigue Crack Growth Observation of 2010 Data

Since 5 classifications using Kohonen SOM has shown the best output classification from the criterion verification and visual verification, it is possible to use the plots from the Kohonen SOM 5 cluster to perform the analysis of the bevel gear. Figure 51 and 52 shows the AE plots of September 7th and October 18th, respectively. From Section 8.2.2, it can be recalled that each mechanism has its own distinct AE characteristics, and mechanisms can be determined by looking at the characteristics of each cluster. The legend in each plots indicate the mechanisms for the each cluster.

In Figure 51, it is possible to observe the small amount of the plane stress circled in red. As the time progresses to October 18th, it can be observed that the plane strain and plane stress area increases from

Figure 52. This is an indication of fatigue crack growth by the increase of the activities of each mechanism.

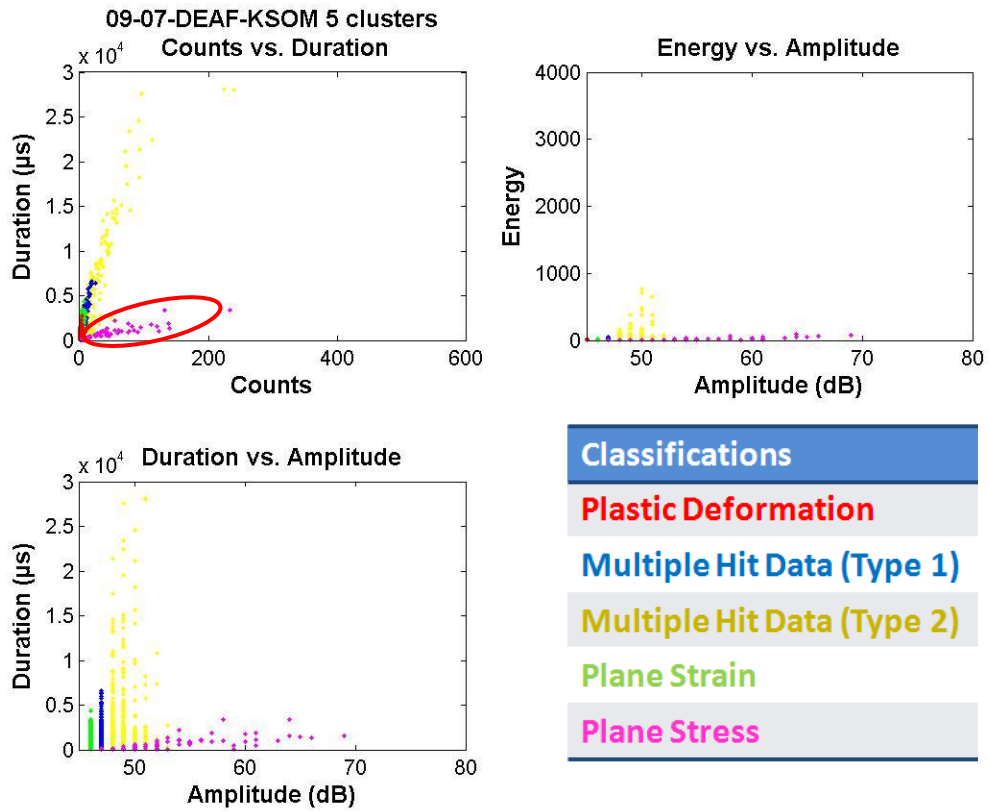


Figure 51 September 7th – Kohonen SOM 5 Cluster AE plots

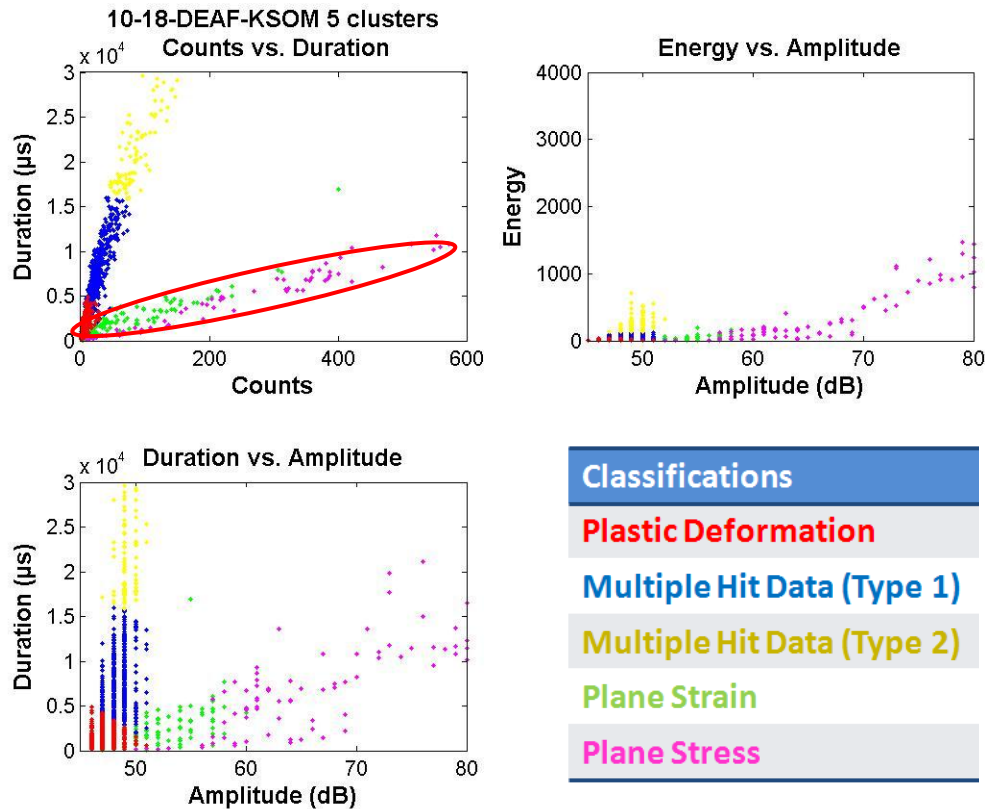
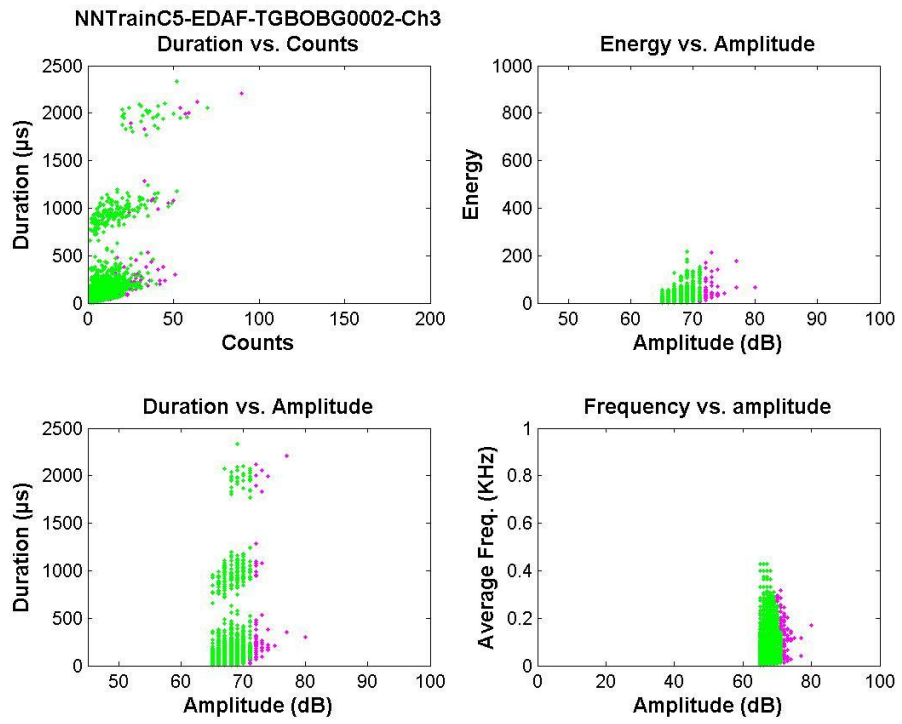


Figure 52 October 18th – Kohonen SOM 5 Cluster AE plots

### 9.5. Kohonen SOM Clustering Analysis of 2011 Data

From the 2010 results, we have concluded that the 5 cluster output using Kohonen SOM will generate the best clustering result. From this understanding, the 2011 acquisition data was analyzed by training the SOM using the October 18, 2010 with 5 clusters for the clustering analysis. The input parameters were the same as iteration 4 of the 2010 data by using duration, energy, amplitude, and average frequency. Figure 53 shows the output AE plots of one of the test dates. The output was set as 5 clusters but the result showed SOM only being able to classify the data set into 2 clusters. One reasons for this result was the setting of the amplitude threshold. Since the 2011 data's amplitude threshold was at 65 dB while the 2010 data threshold was at 45dB, the signals below were eliminated. Due to this elimination, signals of mechanisms that had low amplitude were not acquired resulting in AE plot that has not enough information about the fatigue crack growth of the test specimen.



**Figure 53 TGBOBG0002 Kohonen SOM Clustering**

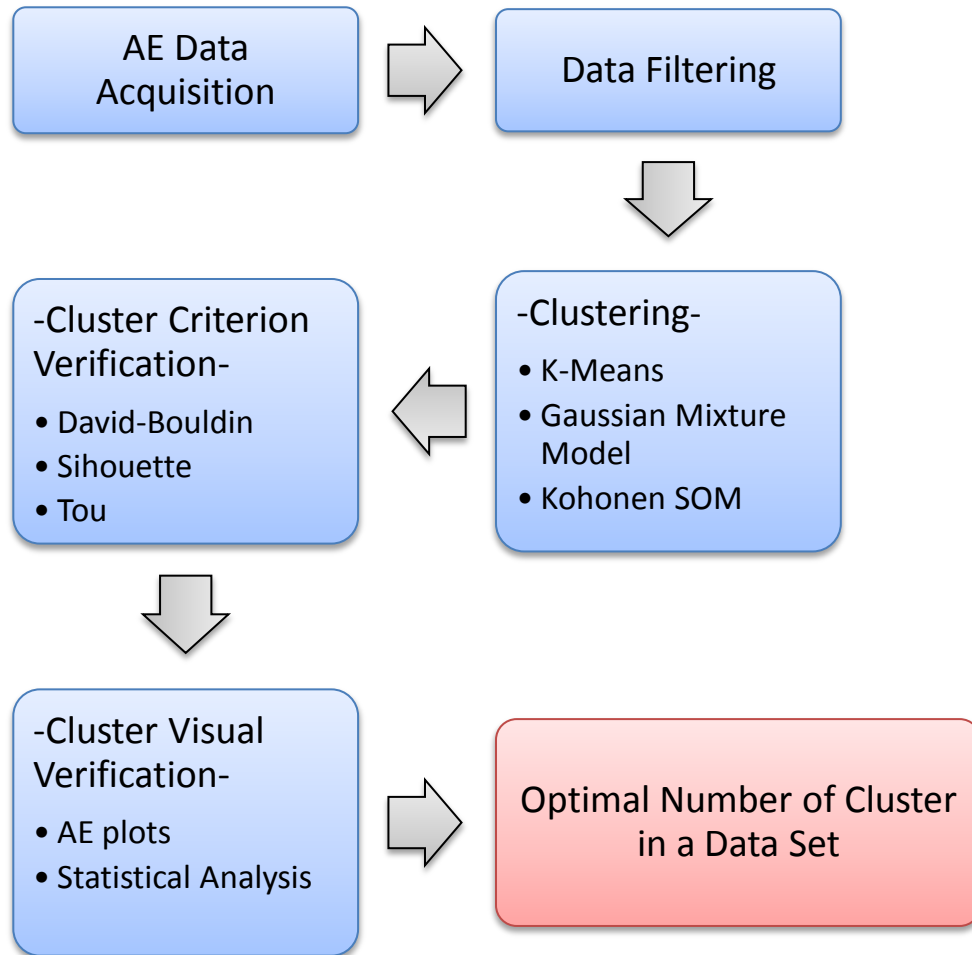
## **Chapter 10. Conclusions and Recommendation**

### **10.1. Conclusions**

Acoustic Emission nondestructive evaluation is a powerful technique for passive volumetric NDE techniques that are capable of performing pressure testing, diagnostics, condition monitoring, and leak detection through detection and analysis of AE signals. It is powerful technique that requires knowledge on the analysis of the AE signals acquired.

In this research, the H-60 Seahawk tail gear box output bevel gear was analyzed for fatigue crack growth inspection as well as developing a methodology and techniques for the analysis of the AE signals from rotating component for early crack detection. The acquired data was filtered to eliminate duration that were greater than 100,000  $\mu$ s and 0 energy hits which are not the area of interest for the failure mechanisms. From the filtered data, three different clustering criteria were performed on three different clustering techniques with different AE input parameters to find out the optimal number of clusters in the data set. The three clustering criterion algorithms are the Davies-Bouldin, Silhouette and Tou Criteria. The three clustering algorithms utilized are the k-mean, Gaussian Mixture Model, and Kohonen Self-organizing Map. From the clustering criteria, it has been suggested that the optimal number of clusters is between 4 to 6 clusters in the data set using the input of energy, duration, amplitude, and average frequency. After the clustering criteria, the visual verification was performed by observing the acoustic emission plots and statistical analysis of each cluster for the performance of the clustering. From the visual inspection, it was observed that all three clustering algorithms were able to very distinctly classify the data set into 5 clusters. The Kohonen SOM especially showed the best distinct classification compared to the other two algorithms. Therefore, it was concluded that using Kohonen SOM with an input of energy, duration, amplitude, average frequency, and output of 5 clusters will show distinct clustering from the data acquired from this test stand with rotating component. Figure 54 shows the complete flow of the analysis to obtain the optimal number of clusters in the data set.

This methodology and technique can be utilized on any kind of structure, not only for rotating components, to find out the optimal number of the output from the AE signal parameters. In the real life situations with hundreds and thousands of data sets, the analyzer may take few data sets to apply this methodology and technique to determine the optimal number of clusters for those AE data sets.



**Figure 54 Analysis Flow to Obtain Optimal Number of Cluster in a Data Set**

The Kohonen SOM was used as a clustering technique and output of 5 clusters, fatigue growth analysis was performed. By comparing the analyzed data sets at midpoint and end point of duration of the test from which the data was acquired, it was possible to observe the increase in fatigue crack activities. This increase in activity indicates the fatigue crack growth.

The AE signals acquired from the tail gearbox output bevel gear of the ground test stand was able to develop a methodology and technique to determine the optimal number of the cluster in the data set. Additionally, the clustered acoustic emission plots from optimal cluster number found and effective clustering algorithm has indicated the increase in fatigue crack activity as a signal for fatigue crack growth. From this, the research steps toward the real-time monitoring system of the H-60 helicopter community have advanced to achieve the goal of safely landing the aircraft before a critical failure.

## **10.2. Recommendations**

Development of a MATLAB<sup>®</sup> code with MATLAB<sup>®</sup> neural toolbox that can perform Kohonen SOM analysis is recommended. If the MATLAB<sup>®</sup> code can be written for Kohonen SOM and shows similar results with the NeuralWorks<sup>®</sup> Professional/II, the entire process of obtaining the optimal number of clusters in a data set can be done with a single command. While using MATLAB<sup>®</sup> to perform Kohonen SOM is recommended, supervised clustering techniques such as learning vector quantization algorithm, support vector machine algorithm, and kth nearest neighbor algorithm should be investigated. These can use training file artificially made or use data set that has distinct clustering such as the October 18<sup>th</sup>, 2010 data.

It is further recommended to refine the methodology and techniques by using more clean AE signal data. The data set used here was very unclean data due to the nature of rotating component creating much noise. If the methodology and techniques can be refined by using AE signal data obtained from controlled environment, it is possible to state that the results will be improved with more understanding of methodology.

This research is one small step towards analyzing the acoustic emission data from a noisy environment like the gearbox. With the methodology and technique established to understand the optimal number of the clusters in the data, it has taken one of the steps necessary to accomplish the future goal of fatigue crack diagnosis and prognosis of rotating gearbox machinery. It is true that further research is still



necessary to accomplish the goal, but it is possible to conclude that the direction of the future research has the potential to accomplish it.

## References

- [1] Sikorsky Aircraft Corporation, "Corporate Brochure," 6 2007. [Online]. Available: [http://www.sikorsky.com/StaticFiles/Sikorsky/Assets/Attachments/Mission%20Downloads/SAC\\_Corporate\\_Brochure.pdf](http://www.sikorsky.com/StaticFiles/Sikorsky/Assets/Attachments/Mission%20Downloads/SAC_Corporate_Brochure.pdf). [Accessed 7 7 2012].
- [2] Naval Air Systems Command, "Aircraft and Weapons H-60 Helicopters," [Online]. Available: <http://www.navair.navy.mil/>. [Accessed 7 7 2012].
- [3] Naval Air Warfare Center Aircraft Division, "H-60 Tail Gearbox Output Bevel Gear Spline Crack Growth Test," Patuxent River, 2008.
- [4] Jet Airlines, "Sikorsky SH-60 Seahawk," [Online]. Available: <http://jet-airlinezz.blogspot.com/2011/10/sikorsky-sh-60-seahawk.html>. [Accessed 8 12 2011].
- [5] United States Navy, "Materials Engineering Analysis Report No. 115416MER1," Materials Engineering Division (AIR-4.3.4) In-Service Support Center , Cherry Point, 2007.
- [6] Sikorsky A United Technologies Company, "Materials Engineering Report MER-SU1101241-Rev1," 2011.
- [7] Y. Liu and S. Nayak, "Structural Health Monitoring: State of the Art and Perspectives," *JOM*, vol. 64, no. 7, pp. 789-792, 2012.
- [8] D. Montalvão, N. M. M. Maia and A. M. R. Riberio, "A Review of Vibration-based Structural Health Monitoring with Special Emphasis on Composite Materials," *The Shock and Vibration Digest*, vol. 38, no. 4, pp. 295-324, 2006.
- [9] A. Belsak and J. Flasker, "Detecting cracks in the tooth root of gears," *Engineering Failure Analysis*, vol. 14, pp. 1466-1475, 2007.
- [10] A. K. Mahamad, T. Hiyama and M. I. Ghazali, "Diagnosis and Prognosis of Bearing Failure in Rotating Machinery Using Acoustic Emission and Artificial Neural Network," The Institute of Electrical Engineers of Japan, 2010.
- [11] S. Yella, N. K. Gupta and M. S. Dougherty, "Comparison of pattern recognition techniques for the classification of impact acoustic emissions," *Transportation Research*, vol. C, no. 15, pp. 345-360, 2007.
- [12] S. G. Vaughn, "In-Flight Fatigue Crack Monitoring of an Aircraft Engine Cowling," Embry Riddle Aeronautical University, Daytona Beach, 1998.
- [13] C. L. Rovik, "Classification of In-Flight Fatigue Cracks in Aircraft Structures using Acoustic Emission and Neural Networks," Embry-Riddle Aeronautical University, Daytona Beach, 1998.

- [14] J. Meriaux, M. Boinet, S. Fouvry and J. C. Lenain, "Identification of fretting fatigue crack propagation mechanisms using acoustic emission," Elsevier, 2010.
- [15] M. G. Sause, T. Müller, A. Horoschenkoff and S. Horn, "Quantification of failure mechanisms in mode-I loading of fiber reinforced plastics utilizing acoustic emission analysis," *Composites Science and Technology*, 2012.
- [16] S. Glodež, M. Šraml and J. Kramberger, "A computational model for determination of service life of gears," *International Journal of Fatigue*, no. 24, pp. 1013-1020, 2002.
- [17] J. Flasker, S. Goldez and S. Pehan, "Influence of contact area on service life of gears with crack in tooth root," *Communications in Numerical Methods Engineering* 11, 1995.
- [18] R. K. Miller, E. v. Hill and P. O. Moore, *Nondestructive Testing Handbook Volume 6 Acoustic Emission Testing*, 3rd ed., Columbus, OH: American Society for Nondestructive Testing, 2005.
- [19] S. L. Knoeller and B. J. Ingold, "the AMMTIAC Quarterly, techsolutions 16," [Online]. Available: <http://ammtiac.alionscience.com/pdf/TechSolutions16.pdf>. [Accessed 30 May 2012].
- [20] "Equipment," NDT Resource Center, [Online]. Available: [http://www.ndt-ed.org/EducationResources/CommunityCollege/Other%20Methods/AE/AE\\_Equipment.htm](http://www.ndt-ed.org/EducationResources/CommunityCollege/Other%20Methods/AE/AE_Equipment.htm). [Accessed 4 7 2012].
- [21] "AE Signal Features," NDT Resource Center, [Online]. Available: [http://www.ndt-ed.org/EducationResources/CommunityCollege/Other%20Methods/AE/AE\\_Signal%20Features.htm](http://www.ndt-ed.org/EducationResources/CommunityCollege/Other%20Methods/AE/AE_Signal%20Features.htm). [Accessed 4 7 2012].
- [22] Physical Acoustics Corporation, *PocketAE & Pocket AEwin User's Manual*, 2 ed., Princeton Junction, NJ: Physical Acoustics Corporation, 2007, p. 16.
- [23] M. A. Okur, "Neural Network Fatigue Life Prediction in Aluminum From Acoustic Emission Data," Embry-Riddle Aeronautical University, Daytona Beach, 2010.
- [24] J. A. Collins, *Failure of Materials in Mechanical Design*, John Wilery & Sons, Inc, 1981.
- [25] NDT Resource Center, "Fracture Toughness," [Online]. Available: <http://www.ndt-ed.org/EducationResources/CommunityCollege/Materials/Mechanical/FractureToughness.htm>. [Accessed 5 July 2012].
- [26] *AEwin SOFTWARE Installation, Operation and User's Reference Manual*, Physical Acoustics Corporation, 2004.
- [27] M. G. Sause, A. Gribov, A. R. Unwin and S. Horn, "Pattern recognition approach to identify natural clusters of acoustic emission signals," Elsevier, 2011.

- [28] F. F. Barsoum, J. Suleman, A. Korcak and E. V. K. Hill, "Acoustic Emission Monitoring and Fatigue Life Prediction in Axially Loaded Notched Steel Specimens," Acoustic Emission Group, 2009.
- [29] J. Iizuka, "Neural Network Burst Pressure Prediction in Composite Overwrapped Pressure Vessels Using Mathematically Modeled Acoustic Emission Data," Embry-Riddle Aeronautical University, Daytona Beach, 2010.
- [30] J. MacQueen, "Some Methods for Classification and Analysis of Multivariate Observations," in *Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, 1967.
- [31] S. V. Vishwanathan, "K-Means and Gaussian Mixture Models," 2012.
- [32] "A Tutorial on Clustering Algorithms-Clustering as a Mixture of Gaussians," [Online]. Available: [http://home.dei.polimi.it/matteucc/Clustering/tutorial\\_html/mixture.html](http://home.dei.polimi.it/matteucc/Clustering/tutorial_html/mixture.html). [Accessed 30 10 2012].
- [33] "Gaussian Mixture Models and the K-means Algorithm," [Online]. Available: [http://www.cse.psu.edu/~rcollins/CSE586Spring2010/lectures/cse586gmmemPart1\\_6pp.pdf](http://www.cse.psu.edu/~rcollins/CSE586Spring2010/lectures/cse586gmmemPart1_6pp.pdf). [Accessed 30 10 2012].
- [34] D. L. Davies and W. D. Bouldin, "A Cluster Separation Measure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vols. PAMI-1, no. 2, pp. 224-227, 1979.
- [35] L. Vendramin, R. J. Campello and E. R. Hruschka, "On the Comparison of Relative Clustering Validity Criteria," *SIAM*, pp. 733-744.
- [36] J. T. Tou, "DYNOC - A Dynamic Optimal Cluster-seeking Technique," *International Journal of Computer and Information Science*, vol. 8, no. 6, 1979.
- [37] "NeuralWorks Professional II/PLUS," Neural Ware, 2009. [Online]. Available: [http://www.neuralware.com/products\\_pro2.jsp](http://www.neuralware.com/products_pro2.jsp). [Accessed 15 10 2012].
- [38] Mistras Group, "R15I-AST Sensor Product Data Sheet," 2011. [Online]. Available: [http://www.mistrasgroup.com/products/company/publications/2\\$Acoustic\\_Emission/Model\\_R15I-AST.pdf](http://www.mistrasgroup.com/products/company/publications/2$Acoustic_Emission/Model_R15I-AST.pdf). [Accessed 15 6 2012].
- [39] Mistras Group, "WD Sensor Product Data Sheet," 2011. [Online]. Available: [http://www.mistrasgroup.com/products/company/publications/2\\$Acoustic\\_Emission/Model\\_WD.pdf](http://www.mistrasgroup.com/products/company/publications/2$Acoustic_Emission/Model_WD.pdf). [Accessed 15 6 2012].
- [40] Mistras Group, "Micro30S Sensor Product Data Sheet," 2011. [Online]. Available: [http://www.mistrasgroup.com/products/company/publications/2\\$Acoustic\\_Emission/Model\\_Micro30S.pdf](http://www.mistrasgroup.com/products/company/publications/2$Acoustic_Emission/Model_Micro30S.pdf). [Accessed 15 6 2012].
- [41] H Series Helicopter Maintenance and Training Manuals, "Figure 3-10 Tail Gearbox," [Online].

Available: [http://hserieshelicopters.tpub.com/TM-1-1520-265-23/css/TM-1-1520-265-23\\_168.htm](http://hserieshelicopters.tpub.com/TM-1-1520-265-23/css/TM-1-1520-265-23_168.htm).  
[Accessed 11 12 2011].

[42] "Acoustic Emission Technology," Mistras, [Online]. Available: <http://www.mistrasgroup.com/products/technologies/acousticemission.aspx>. [Accessed 12 9 2012].

[43] A. Blarasin, M. Guaglzano and L. Vergani, "Fatigue & Fracture of Engineering Materials & Structure 20," 1997.

Hill, Eric v.K., Fady F. Barsoum, Jun Shishino, Ning Leung, Prathikshen Salvadorai, Alan Timmons, and William Hardman. Fatigue Crack Growth Analysis from Acoustic Emission Data on the Navy H-60 Seahawk Helicopter Tail Gearbox. Proc. of TMS 2012, Walt Disney World Dolphin Resort, Orlando, Florida. N.p.: n.p., n.d. 249. Print.

Barsoum, Fady F., Jun Shishino, Ning Y. Leung, Prathikshen N. Salvadorai, Alan B. Timmons, and William J. Hardman. Fatigue Crack Growth Analysis from Acoustic Emission Data on the Navy H-60 Seahawk Helicopter Tail Gearbox. Proc. of AEWG-54, Physical Acoustics Corp., Princeton, NJ. N.p.: n.p., n.d. 13. Print.

Barsoum, F., J. Shishino, and E. Hill. Fatigue Crack Growth Analysis of Aircraft Rotary Components Using Acoustic Emission Data. Proc. of ASNT Fall Conference, Rosen Shingle Creek REsort, Orlando, Florida. N.p.: n.p., n.d. 12. Print.

## Appendices

### Appendix A. MATLAB Code for Filter and Create Input File: InputSeparation.m

```
%%% File Name: Thesis01_InputSeparation.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This InputSeparation.m mfile is use to perform two task in analysis of AE
% signal parameters.
% The first task is to output the TXT file for the analysis using NeuralWorks.
% User input is (1).
% The second task is to output the XLSX file that has all of the AE
parameters.
% User input is (2).
% % % % % % % % % % % % % % % % IMPORTANT !! % % % % % % % % % % % % % % %
% It is important to change the output parameters depending on the
performance
% you are going to take.
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
clc;clear all;close all;

ask1 = input('Enter number for choice of performance. For NN(txt):[1], For
all AE(xlsx);[3] ');
while (ask1 ~= 1 & ask1 ~=3)
    clc
    ask1 = input('Enter number for choice of performance. For NN(txt):[1],
For all AE(xlsx);[3] ');
end
if ask1 == 1
    fprintf('\n>>> Output is for NN analysis (.txt) <<<\n');
elseif ask1 == 3
    fprintf('\n>>> Output is for all AE data (.xlsx) <<<\n');
end
fprintf('\nNN Input Separation Program Running\n\n')
%% Initializer
mkdir('Channel_allCh_EDAF');
mkdir('Channel_1_EDAF');
mkdir('Channel_2_EDAF');
mkdir('Channel_3_EDAF');
mkdir('Channel_4_EDAF');
checkCh1 = 0;
checkCh2 = 0;
checkCh3 = 0;
checkCh4 = 0;
% read the name of excel file that has all AE signals
fileNameInput = importdata('listRawDataName.xls');
[listS1,listS2]=size(fileNameInput.textdata.Sheet1);
%%
%%% 1ID | 2D | 3H:M:S.mun | 4ss | 5CH | 6COUN | 7ENER | 8DUR | 9AMP |10
FEQ %%%
%%% 1T | 2CH | 3COUNTS |4NRG | 5DUR | 6AMP %%%
for fileInc = 1:listS1
    fprintf('NN Input Separation Program Running\n')
    clear allChData ch1Data ch2Data ch3Data ch4Data
    clear checkCh1 checkCh2 checkCh3 checkCh4
```

```

rawInput = fileNameInput.textdata.Sheet1(fileInc,1);
fileNameforSave = fileNameInput.textdata.Sheet1(fileInc,4);
filename = strcat(rawInput)

counterAllCh = 1;
counterCh1 = 1;
counterCh2 = 1;
counterCh3 = 1;
counterCh4 = 1;
checkCh1 = 0;
checkCh2 = 0;
checkCh3 = 0;
checkCh4 = 0;

file = importdata(str2mat(filename));
[m,n] = size(file.Sheet1);

%% Filtering Unwanted Data
fprintf('\nFiltering...')
for incl = 1:m
    if file.Sheet1(incl,4) > 0 %Energy less than 0
        %Duration greater than 0 - 100000 micro sec
        if (file.Sheet1(incl,5) < 100000) && (file.Sheet1(incl,5) > 0)
            freqAll = (file.Sheet1(incl,3))/(file.Sheet1(incl,5));
            % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
            % need to change the parameters to be save depending on the
            % output (allChData)
            allChData(counterAllCh,:) = [file.Sheet1(incl,4),...
                file.Sheet1(incl,5), file.Sheet1(incl,6), freqAll];
            counterAllCh = counterAllCh+1;
            if file.Sheet1(incl,2) == 1 %Channel 1 data
                freqCh1 = (file.Sheet1(incl,3))/(file.Sheet1(incl,5));
                % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
                % need to change the parameters to be save depending on
                the
                % output (ch1Data)
                ch1Data(counterCh1,:) = [file.Sheet1(incl,4),...
                    file.Sheet1(incl,5), file.Sheet1(incl,6), freqCh1];
                counterCh1 = counterCh1 +1;
                checkCh1 = 1;
            elseif file.Sheet1(incl,2) == 2 %Channel 2 data
                freqCh2 = (file.Sheet1(incl,3))/(file.Sheet1(incl,5));
                % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
                % need to change the parameters to be save depending on
                the
                % output (ch2Data)
                ch2Data(counterCh2,:) = [file.Sheet1(incl,4),...
                    file.Sheet1(incl,5), file.Sheet1(incl,6), freqCh2];
                counterCh2 = counterCh2 +1;
                checkCh2 = 1;
            elseif file.Sheet1(incl,2) == 3 %Channel 3 data
                freqCh3 = (file.Sheet1(incl,6))/(file.Sheet1(incl,8));
                % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
                % need to change the parameters to be save depending on
                the

```

```

        % output (ch3Data)
        ch3Data(counterCh3,:) = [file.Sheet1(inc1,4),...
            file.Sheet1(inc1,5), file.Sheet1(inc1,6), freqCh3];
        counterCh3 = counterCh3 +1;
        checkCh3 = 1;
    elseif file.Sheet1(inc1,2) == 4 %Channel 4 data
        freqCh4 = (file.Sheet1(inc1,6))/(file.Sheet1(inc1,8));
        % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
        % need to change the parameters to be save depending on
the
        % output (ch4Data)
        ch4Data(counterCh4,:) = [file.Sheet1(inc1,4),...
            file.Sheet1(inc1,5), file.Sheet1(inc1,6), freqCh4];
        counterCh4 = counterCh4 +1;
        checkCh4 = 1;
    end
end
end
end
end
end
end
    %% Export to .xlsx
    fprintf('\nExporting...')

    % for all channel
    if ask1 == 3
        excelName = strcat('EDAF_',fileNameforSave,'_allCh.xlsx');
        xlswrite(str2mat(excelName), allChData);

movefile(str2mat(excelName),'../InputSeparation_CODE/Channel_allCh_EDAF');
    end
    if ask1 == 1
        txtName = strcat('EDAF_',fileNameforSave,'_allCh.txt');
        dlmwrite(str2mat(txtName), allChData, 'delimiter', '\t');

movefile(str2mat(txtName),'../InputSeparation_CODE/Channel_allCh_EDAF');
    end
    % for channel 1 data
    if checkCh1 == 1
        if ask1 == 3
            excelNameCh1 = strcat('EDAF_',fileNameforSave,'_Ch1.xlsx');
            xlswrite(str2mat(excelNameCh1), ch1Data);

movefile(str2mat(excelNameCh1),'../InputSeparation_CODE/Channel_1_EDAF');
            end
            if ask1 == 1
                txtNameCh1 = strcat('EDAF_',fileNameforSave,'_Ch1.txt');
                dlmwrite(str2mat(txtNameCh1), ch1Data, 'delimiter', '\t');

movefile(str2mat(txtNameCh1),'../InputSeparation_CODE/Channel_1_EDAF');
            end
            end
            % for channel 2 data
            if checkCh2 == 1
                if ask1 == 3
                    excelNameCh2 = strcat('EDAF_',fileNameforSave,'_Ch2.xlsx');
                    xlswrite(str2mat(excelNameCh2), ch2Data);

movefile(str2mat(excelNameCh2),'../InputSeparation_CODE/Channel_2_EDAF');
                end
            end
        end
    end

```



```

        end
        if ask1 == 1
            txtNameCh2 = strcat('EDAF_',fileNameforSave,'_Ch2.txt');
            dlmwrite(str2mat(txtNameCh2), ch2Data, 'delimiter', '\t');

movefile(str2mat(txtNameCh2), '../InputSeparation_CODE/Channel_2_EDAF');
        end
    end
    % for channel 3 data
    if checkCh3 == 1
        if ask1 == 3
            excelNameCh3 = strcat('EDAF_',fileNameforSave,'_Ch3.xlsx');
            xlswrite(str2mat(excelNameCh3), ch3Data);

movefile(str2mat(excelNameCh3), '../rawData_separation_work/Channel_3_EDAF');
        end
        if ask1 ==1
            txtNameCh3 = strcat('EDAF_',fileNameforSave,'_Ch3.txt');
            dlmwrite(str2mat(txtNameCh3), ch3Data, 'delimiter', '\t');

movefile(str2mat(txtNameCh3), '../rawData_separation_work/Channel_3_EDAF');
        end
    end
    % for channel 4 data
    if checkCh4 == 1
        if ask1 == 3
            excelNameCh4 = strcat('EDAF_',fileNameforSave,'_Ch4.xlsx');
            xlswrite(str2mat(excelNameCh4), ch4Data);

movefile(str2mat(excelNameCh4), '../rawData_separation_work/Channel_4_EDAF');
        end
        if ask1 == 1
            txtNameCh4 = strcat('EDAF_',fileNameforSave,'_Ch4.txt');
            dlmwrite(str2mat(txtNameCh4), ch4Data, 'delimiter', '\t');

movefile(str2mat(txtNameCh4), '../rawData_separation_work/Channel_4_EDAF');
        end
    end
end
%% END PROGRAM
fprintf('\nEND OF PROGRAM')

```

## Appendix B. MATLAB Code for Kohonen SOM: nnr2excel.m

```
%%% File Name: Thesis02_nnr2excel.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% *****
% %%%%%%%%%% READ BEFORE RUNNING %%%%%%%%%%
% READ the README.docx before running this mfile!!!!
% *****
% This code can convert the nnr file created by the
% Neural Works Professional II Plus to Excel files
% after adding the file extention of ".txt" to end
% of each nnr files.
% As well as converting, it will add the original
% AE signal parameters to be used for further analysis
clc;clear all;close all

nnrI = 4;
exI = 2;
nameI = 1;
fileNameInput = importdata('listFileName.xls');
[listS1,listS2]=size(fileNameInput.textdata.Sheet1);
%%
for inc = 1:listS1/4
    fprintf('Transferring .nnr to excel\n')
    nameFileRead = fileNameInput.textdata.Sheet1(nnrI,4);
    nameFileName = strcat(nameFileRead)

    nnrFileRead = fileNameInput.textdata.Sheet1(nnrI,1);
    nnrFileName = strcat(nnrFileRead);
    nnrFile = importdata(str2mat(nnrFileName));

    exFileRead = fileNameInput.textdata.Sheet1(exI,1);
    exFileName = strcat(exFileRead);
    exFile = importdata(str2mat(exFileName));

    [exR,exC] = size(exFile.Sheet1);

    %Change Here for different classification Current 4%
    hitAndNN = [exFile.Sheet1(:,2),exFile.Sheet1(:,3),exFile.Sheet1(:,4),...
exFile.Sheet1(:,5),exFile.Sheet1(:,6),nnrFile(:,1),nnrFile(:,2),nnrFile(:,3),
...
                nnrFile(:,4),nnrFile(:,5)];

    excelName = strcat('NNC5_',nameFileName, '.xlsx');
    xlswrite(str2mat(excelName), hitAndNN);

    nnrI = nnrI+4;
    exI = exI +4;
    nameI = nameI +4;
end
fprintf('ENDPROGRAM\n\n')
```

## Appendix C. MATLAB Code for Kohonen SOM: Rij\_FromKSOMnnpro.m

```
%%% File Name: Thesis021_Rij_FromKSOMnnpro.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This Thesis021_Rij_FromKSOMnnpro.m converts the NNpro result
% of the binary numbers into index numbers.
% Than, from the index number of the each cluster,
% it performs the Rij Criterion.
% Input is the (1)AE input parameters and (2)different excel files with
% NNpro results.
% Output is (1)Rij Criterion plot and (2)excel file of Rij criterion
% values.
clc; close all; clear all;
%% Read KSOM NNPro Result Excel
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% read excel file that has name of all excel file from 2 cluster to 7
readFileName = importdata('listFileNameForKSOMVerification.xls');
% Files that has AE input for analysis
trainFile = xlsread ('DEAF_10_18_all.xlsx');
% Name of the output file
fileName = strcat('10-18-DEAF-KSOM-Rij');
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %

%% Normalization of input data
% trainFile = tansig(trainFile);
% for i=1:4
%     trainFile(:,i)=tanh(trainFile(:,i));
%     trainFile(:,i)=log(trainFile(:,i));
% end
for i=1:4
    trainFile(:,i)=(trainFile(:,i)-mean(trainFile(:,i)))/std(trainFile(:,i));
end
train = [trainFile(:,1),trainFile(:,2), trainFile(:,3), trainFile(:,4)];

%% Change NNPro output to idx
q=2;p=2;
[r0 c0] = size(readFileName.textdata.Sheet1);
for fileInc = 1:r0
    fileNN = strcat(readFileName.textdata.Sheet1(fileInc,1));
    readNNOutputFile = xlsread (str2mat(fileNN));
    [r1 c1] = size(readNNOutputFile);
    numClust = c1 - 5;
    idxR = 1;
    for incl = 1:r1
        if numClust >= 1
            if readNNOutputFile(incl, 6) == 1
                idx(idxR,1) = 1;
                idxR = idxR +1;
            end
        end
        if numClust >= 2
            if readNNOutputFile(incl, 7) == 1
                idx(idxR,1) = 2;
            end
        end
    end
end
```

```

        idxR = idxR +1;
    end
end
if numClust >= 3
    if readNNOutputFile(incl, 8) == 1
        idx(idxR,1) = 3;
        idxR = idxR +1;
    end
end
if numClust >= 4
    if readNNOutputFile(incl, 9) == 1
        idx(idxR,1) = 4;
        idxR = idxR +1;
    end
end
if numClust >= 5
    if readNNOutputFile(incl, 10) == 1
        idx(idxR,1) = 5;
        idxR = idxR +1;
    end
end
if numClust >= 6
    if readNNOutputFile(incl, 11) == 1
        idx(idxR,1) = 6;
        idxR = idxR +1;
    end
end
if numClust >= 7
    if readNNOutputFile(incl, 12) == 1
        idx(idxR,1) = 7;
        idxR = idxR +1;
    end
end
end
fileInc;
idxSave(:,fileInc) = idx;
end
%% Rij Criterion
k =7;
for num=2:k
    idx = idxSave(:,num-1);
    for i=1:num
        N0=0;
        A=size(train(idx==i,:));
        T(i,1)=A(1,1);
        D=train(idx==i,:);
        ctrs(i,:)=mean(D);
        fprintf('p2')
        pause
        for l=1:T(i,1)
            N=(norm(D(l,:)-ctrs(i,:)))^q;
            N0=N0+N;
        end
        N1(i,1)=N0;
        S(i,1)=(N0/T(i,1))^(1/q);
    end
end
S;

```

```

for i=1:num
    for j=1:num
        M(i,j)=(sum((((ctrs(i,:)-ctrs(j,:)).^2).^ (1/2)).^p))^(1/p);
    end
end
M;
for i=1:num
    for j=1:num
        R1(i,j)=(S(i,1)+S(j,1))/M(i,j);
    end
    R1(i,i)=0;
end
R1;
for i=1:num
    C(i,1)=max(R1(i,:));
end
C;
R(num-1,1)=(sum(C(:,1))./num);
cluster_num(num-1,1)=num;
end
%% Plot Rij verification
R;
cluster_num;
figure(1)
plot(cluster_num,R.^-1)
set(gca,'xtick',2:7);
title([fileName, ' ', 'Verification Plot'], 'fontsize',12, 'fontweight', 'b');
xlabel('Number of Clusters', 'fontsize',12, 'fontweight', 'b');
ylabel('Rij Value', 'fontsize',12, 'fontweight', 'b');
filename1 = strcat(fileName, '_', '_Verification.jpg');
saveas(1,str2mat(filename1));
%% excel file output Rij value
fileNameRvalue = strcat(fileName, '_RVALUE.xlsx');
xlswrite(fileNameRvalue,R.^-1)
%% excel file output of KSOM idx
fileNameExIDX = strcat(fileName, '_IDXKSOM.xlsx');
xlswrite(fileNameExIDX,idxSave)

fprintf('\n\nEND of PROGRAM')
beep
pause(0.25)
beep

```

## Appendix D. MATLAB Code for Kohonen SOM: Silhouette\_FromKSOMnnpro.m

```
%%% File Name: Thesis022_Silhouette_FromKSOMnnpro.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This Thesis022_Silhouette_FromKSOMnnpro.m converts the NNpro result
% of the binary numbers into index numbers.
% Than, from the index number of the each cluster,
% it performs the Silhouette Criterion.
% Input is the (1)AE input parameters and (2)different excel files with
% NNpro results.
% Output is (1)Silhouette Criterion plot and (2)excel file of Silhouette
criterion
% values.
clc; close all; clear all;
%%% Read KSOM NNPro Result Excel
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % CHANGE DATE % % % % % % % % % % %
% read excel file that has name of all excel file from 2 cluster to 7
readFileName = importdata('listFileNameForKSOMVerification.xls');
% Files that has AE input for analysis
trainFile = xlsread ('DEAF_10_18_all.xlsx');
% Name of the output file
fileName = strcat('10-18-DEAF-KSOM-Sihouette');
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % %

%% Normalization of input data
% trainFile = tansig(trainFile);
% for i=1:4
%     trainFile(:,i)=tanh(trainFile(:,i));
% %     trainFile(:,i)=log(trainFile(:,i));
% end
for i=1:4
    trainFile(:,i)=(trainFile(:,i)-mean(trainFile(:,i)))/std(trainFile(:,i));
end
train = [trainFile(:,1),trainFile(:,2), trainFile(:,3), trainFile(:,4)];

%% Change NNPro output to idx
% q=2;p=2;
[r0 c0] = size(readFileName.textdata.Sheet1);
for fileInc = 1:r0
    fileNN = strcat(readFileName.textdata.Sheet1(fileInc,1))
    readNNOutputFile = xlsread (str2mat(fileNN));
    [r1 c1] = size(readNNOutputFile);
    numClust = c1 - 5;
    idxR = 1;
    for incl = 1:r1
        if numClust >= 1
            if readNNOutputFile(incl, 6) == 1
                idx(idxR,1) = 1;
                idxR = idxR + 1;
            end
        end
        if numClust >= 2
            if readNNOutputFile(incl, 7) == 1
```

```

        idx(idxR,1) = 2;
        idxR = idxR +1;
    end
end
if numClust >= 3
    if readNNOOutputFile(incl, 8) == 1
        idx(idxR,1) = 3;
        idxR = idxR +1;
    end
end
if numClust >= 4
    if readNNOOutputFile(incl, 9) == 1
        idx(idxR,1) = 4;
        idxR = idxR +1;
    end
end
if numClust >= 5
    if readNNOOutputFile(incl, 10) == 1
        idx(idxR,1) = 5;
        idxR = idxR +1;
    end
end
if numClust >= 6
    if readNNOOutputFile(incl, 11) == 1
        idx(idxR,1) = 6;
        idxR = idxR +1;
    end
end
if numClust >= 7
    if readNNOOutputFile(incl, 12) == 1
        idx(idxR,1) = 7;
        idxR = idxR +1;
    end
end
end
fileInc;
idxSave(:,fileInc) = idx;
end
%% Silhouette Criterion
k =7;
for num=2:k
    Sx1=0;
    idx = idxSave(:,num-1);
    for i=1:num
        D=train(idx==i,:);
        ctrs(i,:)=mean(D)
    end
    for i=1:num
        D=train(idx==i,:);
        ctrs(i,:)=mean(D);
        A=size(train(idx==i,:));
        T(i,1)=A(1,1);
        for g=1:T(i,1)
            N1=0;
            N=(norm(D(g,:)-ctrs(i,:)))/(num-1);
            Nsum(i,g)=N;
            Nsum2(i,g)=inf;
        end
    end
end

```

```

        for h=1:num
            if (h~=i)
                N3=0;
                N2=(norm(D(g,:)-ctrs(h,:)))/(num-1);
                N3=N2+N3;
                Nsum2(i,g)=min(Nsum2(i,g),N3);
            end
        end
        Sx=(Nsum2(i,g)-Nsum(i,g))/max(Nsum2(i,g),Nsum(i,g));
        Sx1=Sx1+Sx;
    end
end
SWc(num-1,1)=Sx1/sum(T(:,1))
cluster_num(num-1,1)=num;
end
%% Plot Silhouette verification
SWc;
cluster_num;
figure(1)
plot(cluster_num,SWc)
set(gca,'xtick',2:7)
title([fileName, ' ', 'Verification Plot'],'fontsize',12,'fontweight','b');
xlabel('Number of Clusters','fontsize',12,'fontweight','b');
ylabel('Silhouette Value','fontsize',12,'fontweight','b');
filename1 = strcat(fileName, '_', '_Verification.jpg');
saveas(1,str2mat(filename1))
%% excel file output Silhouette value
fileNameRvalue = strcat(fileName, '_SWcVALUE.xlsx');
xlswrite(fileNameRvalue,SWc)
%% excel file output of KSOM idx
fileNameExIDX = strcat(fileName, '_IDXKSOM.xlsx');
xlswrite(fileNameExIDX,idxSave)

fprintf('\n\nEND of PROGRAM')
beep
pause(0.25)
beep

```



## Appendix E. MATLAB Code for Kohonen SOM: Tou\_FromKSOMnnpro.m

```
%%% File Name: Thesis023_Tou_FromKSOMnnpro.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This Thesis023_Tou_FromKSOMnnpro.m converts the NNpro result
% of the binary numbers into index numbers.
% Than, from the index number of the each cluster,
% it performs the Tou Criterion.
% Input is the (1)AE input parameters and (2)different excel files with
% NNpro results.
% Output is (1)Tou Criterion plot and (2)excel file of Tou criterion
% values.
clc; close all; clear all;

%% Read KSOM NNPro Result Excel
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % CHANGE DATE % % % % % % % % % % % %
% read excel file that has name of all excel file from 2 cluster to 7
readFileName = importdata('listFileNameForKSOMVerification.xls');
% Files that has AE input for analysis
trainFile = xlsread ('DEAF_10_18_all.xlsx');
% Name of the output file
fileName = strcat('10-18-DEAF-KSOM-Tou');
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %

%% Normalization of input data
% trainFile = tansig(trainFile);
% for i=1:4
%     trainFile(:,i)=tanh(trainFile(:,i));
%     trainFile(:,i)=log(trainFile(:,i));
% end
for i=1:4
    trainFile(:,i)=(trainFile(:,i)-mean(trainFile(:,i)))/std(trainFile(:,i));
end
train = [trainFile(:,1),trainFile(:,2), trainFile(:,3), trainFile(:,4)];

%% Change NNPro output to idx
% q=2;p=2;
[r0 c0] = size(readFileName.textdata.Sheet1);
for fileInc = 1:r0
    fileNN = strcat(readFileName.textdata.Sheet1(fileInc,1))
    readNNOutputFile = xlsread (str2mat(fileNN));
    [r1 c1] = size(readNNOutputFile);
    numClust = c1 - 5;
    idxR = 1;
    for incl = 1:r1
        if numClust >= 1
            if readNNOutputFile(incl, 6) == 1
                idx(idxR,1) = 1;
                idxR = idxR +1;
            end
        end
        if numClust >= 2
            if readNNOutputFile(incl, 7) == 1
```

```

        idx(idxR,1) = 2;
        idxR = idxR +1;
    end
end
if numClust >= 3
    if readNNOOutputFile(incl, 8) == 1
        idx(idxR,1) = 3;
        idxR = idxR +1;
    end
end
if numClust >= 4
    if readNNOOutputFile(incl, 9) == 1
        idx(idxR,1) = 4;
        idxR = idxR +1;
    end
end
if numClust >= 5
    if readNNOOutputFile(incl, 10) == 1
        idx(idxR,1) = 5;
        idxR = idxR +1;
    end
end
if numClust >= 6
    if readNNOOutputFile(incl, 11) == 1
        idx(idxR,1) = 6;
        idxR = idxR +1;
    end
end
if numClust >= 7
    if readNNOOutputFile(incl, 12) == 1
        idx(idxR,1) = 7;
        idxR = idxR +1;
    end
end
end
fileInc;
idxSave(:,fileInc) = idx;
end
%% Tou Criterion
k=7;
for num=2:k
    idx = idxSave(:,num-1);
    for i=1:num
        N0=0;
        A=size(train(idx==i,:));
        T(i,1)=A(1,1); %number of elements in cluster i, Ti
        D=train(idx==i,:);
        ctrs(i,:)=mean(D)
        for l=1:T(i,1)
            N=(norm(D(l,:)-ctrs(i,:)).^2);
            N0=N0+N;
        end
        N1(i,1)=N0;
        S(i,1)=(2*(N0/T(i,1))).^0.5; %Si
    end
end
S;
S1=max(S);

```

```

for i=1:num
    for j=1:num
        M(i,j)=norm(ctrs(i,:)-ctrs(j,:));

        end
    M(i,i)=inf;
end
M;
min(M);
N=min(min(M));
To(num-1,1)=N/S1;
cluster_num(num-1,1)=num;
end
%% Plot Tou verification
To;
cluster_num;
figure(1);
plot(cluster_num,To)
set(gca,'xtick',2:7)
title([fileName, ' ', 'Verification Plot'],'fontsize',12,'fontweight','b');
xlabel('Number of Clusters','fontsize',12,'fontweight','b');
ylabel('Tou Value','fontsize',12,'fontweight','b');
filename1 = strcat(fileName, '_', '_Verification.jpg');
saveas(1,str2mat(filename1));
%% excel file output Tou value
fileNameRvalue = strcat(fileName, '_TouVALUE.xlsx');
xlswrite(fileNameRvalue,To)
%% excel file output of KSOM idx
fileNameExIDX = strcat(fileName, '_IDXKSOM.xlsx');
xlswrite(fileNameExIDX,idxSave)

fprintf('\n\nEND of PROGRAM')
beep
pause(0.25)
beep

```

## Appendix F. MATLAB Code for Kohonen SOM: vPlotKSOM.m

```
%%% File Name: Thesis024_vPlotKSOM.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This vPlotKSOM.m file input the all three criterion values
% and then plots them while normalizing them between -1 to 1.
% Then the rank of each point is found and assigned a points
% in k*2 scale from the highest peak. Where k is number of clusters
% analyzed.
% This rank plot is plotted as well as showing the 1st and 2nd
% peak in the voting plot.
% Plot is saved as JPEG.
% THIS IS FOR KSOM FOR FILE NAME SIMPLICITY
%-----
clc; clear all, close all;
% % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % %
fileName = ('10-18-DEAF-KSOM')
RijValueFile = xlsread ('10-18-DEAF-KSOM-Rij_RVALUE.xlsx');
SihouetteValueFile = xlsread ('10-18-DEAF-KSOM-Sihouette_SWcVALUE.xlsx');
TouValueFile = xlsread ('10-18-DEAF-KSOM-Tou_TouVALUE.xlsx');
% % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % %
x =[2;3;4;5;6;7];
%% Rank the three criterion results
[vsR, viR] = sort(RijValueFile);
[xR, vrR] = sort(viR);
[vsS, viS] = sort(SihouetteValueFile);
[xS, vrS] = sort(viS);
[vsT, viT] = sort(TouValueFile);
[xT, vrT] = sort(viT);
VotingValue = 2*vrR + 2*vrS + 2*vrT
VVmax = max(VotingValue);
VVplot = VotingValue / VVmax;
[aa, bb]=sort(VVplot);
[cc,dd] = sort(bb);
vRank(:,1)= VVplot;
vRank(:,2)= dd;
%% have criterion results between -1 to 1
[Rmax, Ri] = max(abs(RijValueFile));
Smax = max(abs(SihouetteValueFile));
Tmax = max(abs(TouValueFile));
Rvalue(:,1) = RijValueFile / Rmax;
Svalue(:,1) = SihouetteValueFile / Smax;
Tvalue(:,1) = TouValueFile / Tmax;
%% plot verification plot
figure(1)
plot(x,Rvalue,'r', 'LineWidth', 3)
hold on
plot(x,Svalue,'g', 'LineWidth', 3)
hold on
plot(x,Tvalue,'b', 'LineWidth', 3)
hold on
plot(x,VVplot,':k', 'LineWidth', 3)
hold on
```

```

%% plot the dot on 1st and 2nd highest peak
for inc = 1:6
    if vRank(inc,2) == 6
        hold on
        plot(inc+1,vRank(inc,1), '*m','LineWidth', 8)
    elseif vRank(inc,2) == 5
        hold on
        plot(inc+1,vRank(inc,1), '*c','LineWidth', 8)

    end
end
ylim([-1.25 1.25])
set(gca,'xtick',2:7)
set(gca,'linewidth',2)
set(gca, 'FontWeight', 'bold')
legend('Rij','Sihouette','Tou','Voting Value', 'Location','SouthWest')
title([fileName, ' ', 'Verification Criterion'], 'FontWeight', 'bold')
xlabel('Number of Clusters', 'FontWeight', 'bold')
ylabel('Index', 'FontWeight', 'bold')
filename = strcat(fileName, '_vPlot.jpg');
saveas(1, str2mat(filename));

fprintf('END of PROGRAM\n\n')
beep
pause(0.25)
beep

```

## Appendix G. MATLAB Code for K-mean: Kmeans.m

```

%%% File Name: Thesis03_Kmeans.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This Thesis03_Kmeans.m performs k-mean analysis of the data set.
% Input is (1)excel file with AE input for analysis
% and (2)excel file with all AE parameters.
% Output is (1)excel file that contains centroids for each cluster (2)AE Plot
% (3) index number cluster by kmean.
clc; clear all; close all;

% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% Files that has input for analysis
trainFile = xlsread ('10_18_DEAF_KNN.xls');
% File that has original AE parameters
readOFile = xlsread ('10_18_DEAF_KNN_original.xls');
% Name of the output file
fileName = strcat('10-18-DEAF-Kmean')
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %

% Normalization of input data
for i=1:4
trainFile(:,i)=log(trainFile(:,i))
end
train = [trainFile(:,1),trainFile(:,2), trainFile(:,3), trainFile(:,4)];

q=2;p=2;k=7;
for num=2:k
    %% k-mean and save centroids
    [idx,ctrs,sumd] = kmeans
(train,num,'replicates',100,'display','final','maxiter',7500);
    fileNameCtrsIDX = strcat(fileName,'_',num2str(num),'_CtrsKmean.xlsx');
    xlswrite(fileNameCtrsIDX,ctrs);
    idxSave(:,num-1) = idx;
    %% Separate each AE parameter into class
    clusterNum = num2str(num);
    c1count = 1;c2count = 1;c3count = 1;c4count = 1;c5count = 1;c6count =
1;c7count = 1;
    [r c] = size(idx);
    % A      |B      |C      |D      |E      |F
    % Counts  |Energy |Duration |Amp    |Avg Freq |idx
    % 1      |2      |3      |4      |5      |6
    for inc = 1:r
        if num >=2
            if idx(inc,1) == 1
                class1(c1count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
                c1count = c1count+1;
            elseif idx(inc,1) == 2
                class2(c2count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
                c2count = c2count+1;
            end
        end
    end
end

```

```

end
if num>=3
    if idx(inc,1) == 3
        class3(c3count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
        c3count = c3count+1;
    end
end
if num>=4
    if idx(inc,1) == 4
        class4(c4count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
        c4count = c4count+1;
    end
end
if num>=5
    if idx(inc,1) == 5
        class5(c5count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
        c5count = c5count+1;
    end
end
if num>=6
    if idx(inc,1) == 6
        class6(c6count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
        c6count = c6count+1;
    end
end
if num==7
    if idx(inc,1) == 7
        class7(c7count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
        c7count = c7count+1;
    end
end
end
end
%% Plot AE Plot
figure(2)
%% Duration3 vs. Counts2
subplot(2,2,1)
if num>=2
    plot(class1(:,1),class1(:,3),'.r')
    hold on
    plot(class2(:,1),class2(:,3),'.g')
    hold on
end
if num>=3
    plot(class3(:,1),class3(:,3),'.b')
    hold on
end
if num>=4
    plot(class4(:,1),class4(:,3),'.y')
    hold on
end
if num>=5
    plot(class5(:,1),class5(:,3),'.m')

```

```

        hold on
    end
    if num>=6
        plot(class6(:,1),class6(:,3),'.c')
        hold on
    end
    if num>=7
        plot(class7(:,1),class7(:,3),'.k')
        hold on
    end
    title([fileName, ' ',clusterNum, ' clusters'];[' Counts vs.
Duration'],'fontsize',12,'fontweight','b')
    xlabel('Counts','fontsize',12,'fontweight','b')
    ylabel('Duration ( $\mu$ s)','fontsize',12,'fontweight','b')
    axis([0 600 0 30000])
    %% Energy2 vs. Amplitude4
    subplot(2,2,2)
    if num >=2
        plot(class1(:,4),class1(:,2),'.r')
        hold on
        plot(class2(:,4),class2(:,2),'.g')
        hold on
    end
    if num>=3
        plot(class3(:,4),class3(:,2),'.b')
        hold on
    end
    if num>=4
        plot(class4(:,4),class4(:,2),'.y')
        hold on
    end
    if num>=5
        plot(class5(:,4),class5(:,2),'.m')
        hold on
    end
    if num>=6
        plot(class6(:,4),class6(:,2),'.c')
        hold on
    end
    if num==7
        plot(class7(:,4),class7(:,2),'.k')
        hold on
    end
    title(['Energy vs. Amplitude'],'fontsize',12,'fontweight','b')
    xlabel('Amplitude (dB)','fontsize',12,'fontweight','b')
    ylabel('Energy','fontsize',12,'fontweight','b')
    axis([45 80 0 4000])
    %% Duration3 vs. Amplitude4
    subplot(2,2,3)
    if num >=2
        plot(class1(:,4),class1(:,3),'.r')
        hold on
        plot(class2(:,4),class2(:,3),'.g')
        hold on
    end
    if num>=3
        plot(class3(:,4),class3(:,3),'.b')

```



```

        hold on
    end
    if num>=4
        plot(class4(:,4),class4(:,3),'.y')
        hold on
    end
    if num>=5
        plot(class5(:,4),class5(:,3),'.m')
        hold on
    end
    if num>=6
        plot(class6(:,4),class6(:,3),'.c')
        hold on
    end
    if num==7
        plot(class7(:,4),class7(:,3),'.k')
        hold on
    end
    title(['Duration vs. Amplitude'],'fontsize',12,'fontweight','b')
    xlabel('Amplitude (dB)','fontsize',12,'fontweight','b')
    ylabel('Duration ( $\mu$ s)','fontsize',12,'fontweight','b')
    axis([45 80 0 30000])
    % Save AE plot to JPEG
    filename = strcat(fileName, '_',clusterNum,'_Subplot.jpg');
    saveas(2,str2mat(filename));
    clearvars inc class1 class2 class3 class4 class5 class6 class7 figure(2)
    close all
    c1count = 1;c2count = 1;c3count = 1;c4count = 1;c5count = 1;c6count =
1;c7count = 1;
end

% Save k-mean index number to JPEG
fileNameExIDX = strcat(fileName,'_IDXKmean.xlsx');
xlswrite(fileNameExIDX,idxSave);

%% END PROGRAM
fprintf('\nEND OF PROGRAM')

```

## Appendix H. MATLAB Code for K-mean: Rij\_FromKmeanIDX

```
%%% File Name: Thesis031_Rij_FromKmeanIDX.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This Thesis031_Rij_FromKmeanIDX.m code is used to obtain the Rij critrion
result.
% The input is (1)excel file with AE input for analysis, (2)excel file with
% all AE parameters, (3)excel file that has cluster index numbers.
% The output is (1)Rij Criterion plot and (2)excel file of Rij criterion
% values.
clc; clear all; close all;

% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % CHANGE DATE % % % % % % % % % % % % % % % %
% Files that has input for analysis
trainFile = xlsread ('10_18_DEAF_KNN.xls');
% File that has original AE parameters
readOFile = xlsread ('10_18_DEAF_KNN_original.xls');
% File that has cluster index number
readIDX = xlsread('10-18-DEAF-Kmean_IDXKmean.xlsx');
% Name of the output file
fileName = strcat('10-18-DEAF-Kmean-Rij')
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %

%% Normalization of input data
for i=1:4
trainFile(:,i)=log(trainFile(:,i));
end
train = [trainFile(:,1),trainFile(:,2), trainFile(:,3), trainFile(:,4)];

%% Rij Analysis
q=2;p=2;k=7;
for num=2:k
    idx = readIDX(:,num-1);
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % CHANGE DATE % % % % % % % % % % % % % % % %
ctrsExcel = strcat('10-18-DEAF-Kmean_', num2str(num), '_CtrsKmean.xlsx');
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
    readCtrs = xlsread(ctrsExcel);
    for i=1:num
        N0=0;
        A=size(train(idx==i,:));
        T(i,1)=A(1,1); %number of elements in cluster i, Ti
        D=train(idx==i,:);
        readCtrs(i,:);
        for l=1:T(i,1)
            N=(norm(D(l,:)-readCtrs(i,:))q);
            N0=N0+N;
        end
        N1(i,1)=N0;
        S(i,1)=(N0/T(i,1))(1/q); %Si
    end
end
S;
```

```

    for i=1:num
        for j=1:num
            M(i,j)=(sum((((readCtrs(i,:) -
readCtrs(j,:)).^2).^ (1/2)).^p))^(1/p);
        end
    end
    M;
    for i=1:num
        for j=1:num
            R1(i,j)=(S(i,1)+S(j,1))/M(i,j);
        end
        R1(i,i)=0;
    end
    R1;
    for i=1:num
        C(i,1)=max(R1(i,:));
    end
    C;
    R(num-1,1)=(sum(C(:,1))./num);
    cluster_num(num-1,1)=num;
end
%% Plot Rij verification
newdata = R';
cluster_num;
figure(1)
plot(cluster_num,R.^-1);
set(gca,'xtick',2:7);
title([fileName, ' ', 'Verification Plot'], 'fontsize',12, 'fontweight', 'b');
xlabel('Number of Clusters', 'fontsize',12, 'fontweight', 'b');
ylabel('Rij Value', 'fontsize',12, 'fontweight', 'b');
filename1 = strcat(fileName, '_','_Verification.jpg');
saveas(1,str2mat(filename1));

%% excel file output out Rij value
fileNameRvalue = strcat(fileName, '_RVALUE.xlsx');
xlswrite(fileNameRvalue,R.^-1)

fprintf('\n\nEND of PROGRAM')
beep
pause(0.25)
beep

```

## Appendix I. MATLAB Code for K-mean: Silhouette\_FromKmeanIDX.m

```
%%% File Name: Thesis032_Silhouette_FromKmeanIDX.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This Thesis032_Silhouette_FromKmeanIDX.m code is used to obtain
% the Silhouette critrion result.
% The input is (1)excel file with AE input for analysis, (2)excel file with
% all AE parameters, (3)excel file that has cluster index numbers.
% The output is (1)Silhouette Criterion plot and (2)excel file of
% Silhouette criterion values.
clc; clear all; close all;

% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % CHANGE DATE % % % % % % % % % % % % % % % % % % % % % % % % % %
% Files that has input for analysis
trainFile = xlsread('10_18_DEAF_KNN.xls');
% File that has original AE parameters
readOFile = xlsread('10_18_DEAF_KNN_original.xls');
% File that has cluster index number
readIDX = xlsread('10-18-DEAF-Kmean_IDXKmean.xlsx');
% Name of the output file
fileName = strcat('10-18-DEAF-Kmean-Silhouette')
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %

%% Normalization of input data
for i=1:4
    trainFile(:,i)=log(trainFile(:,i));
end
train = [trainFile(:,1),trainFile(:,2), trainFile(:,3), trainFile(:,4)];

%% Silhouette Analysis
k=7;
for num=2:k
    Sx1=0;
    idx = readIDX(:,num-1);
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % CHANGE % % % % % % % % % % % % % % % % % % % % % % % % % %
% ctrsExcel = strcat('09-07-DEAF-Kmean_',num2str(num),'_CtrsKmean.xlsx');
ctrsExcel = strcat('10-18-DEAF-Kmean_',num2str(num),'_CtrsKmean.xlsx');
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
    readCtrs = xlsread(ctrsExcel);
    for i=1:num
        D=train(idx==i,:);
        A=size(train(idx==i,:));
        T(i,1)=A(1,1);
        for g=1:T(i,1)
            N1=0;
            N=(norm(D(g,:)-readCtrs(i,:)))/(num-1);
            Nsum(i,g)=N;
            Nsum2(i,g)=inf;
            for h=1:num
                if (h~=i)
                    N3=0;
```

```

        N2=(norm(D(g,:)-readCtrs(h,:))/(num-1);
        N3=N2+N3;
        Nsum2(i,g)=min(Nsum2(i,g),N3);
    end
    end
    Sx=(Nsum2(i,g)-Nsum(i,g))/max(Nsum2(i,g),Nsum(i,g));
    Sx1=Sx1+Sx;
end
end
SWc(num-1,1)=(Sx1/sum(T(:,1)));
cluster_num(num-1,1)=num;
end
%% Plot Silhouette verification
newdata = SWc';
cluster_num;
figure(1)
plot(cluster_num,SWc)
set(gca,'xtick',2:7)
title([fileName, ' ', 'Verification Plot'],'fontsize',12,'fontweight','b');
xlabel('Number of Clusters','fontsize',12,'fontweight','b');
ylabel('Silhouette Value','fontsize',12,'fontweight','b');
filename1 = strcat(fileName, '_', '_Verification.jpg');
saveas(1,str2mat(filename1))

%% excel file output out Silhouette value
fileNameRvalue = strcat(fileName, '_SWcVALUE.xlsx');
xlswrite(fileNameRvalue,SWc)

fprintf('\n\nEND of PROGRAM')
beep
pause(0.25)
beep

```

## Appendix J. MATLAB Code for K-mean: Tou\_FromKmeanIDX.m

```

%%% File Name: Thesis033_Tou_FromKmeanIDX.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This Thesis033_Tou_FromKmeanIDX.m code is used to obtain the Tou critrion
result.
% The input is (1)excel file with AE input for analysis, (2)excel file with
% all AE parameters, (3)excel file that has cluster index numbers.
% The output is (1)Tou Criterion plot and (2)excel file of Tou criterion
% values.
clc; clear all; close all;

% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % CHANGE DATE % % % % % % % % % % % % % % % %
% Files that has input for analysis
trainFile = xlsread ('10_18_DEAF_KNN.xls');
% File that has original AE parameters
readOFile = xlsread ('10_18_DEAF_KNN_original.xls');
% File that has cluster index number
readIDX = xlsread('10-18-DEAF-Kmean_IDXKmean.xlsx');
% Name of the output file
fileName = strcat('10-18-DEAF-Kmean-Tou')
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %

%% Normalization of input data
for i=1:4
    trainFile(:,i)=log(trainFile(:,i));
end
train = [trainFile(:,1),trainFile(:,2), trainFile(:,3), trainFile(:,4)];

%% Tou Analysis
k=7;
for num=2:k
    idx = readIDX(:,num-1);
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % CHANGE DATE % % % % % % % % % % % % % % % %
ctrExcel = strcat('10-18-DEAF-Kmean_', num2str(num), '_CtrsKmean.xlsx');
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
    readCtrs = xlsread(ctrExcel);
    for i=1:num
        N0=0;
        A=size(train(idx==i,:));
        T(i,1)=A(1,1); %number of elements in cluster i, Ti
        D=train(idx==i,:);
        for l=1:T(i,1)
            N=(norm(D(l,:)-readCtrs(i,:)).^2);
            N0=N0+N;
        end
        N1(i,1)=N0;
        S(i,1)=(2*(N0/T(i,1))).^0.5; %Si
    end
    S1=max(S);
    for i=1:num

```

```

        for j=1:num
            M(i,j)=norm(readCtrs(i,:)-readCtrs(j,:));
        end
        M(i,i)=inf;
    end
    M;
    min(M);
    N=min(min(M));
    To(num-1,1)=N/S1;
    cluster_num(num-1,1)=num;
end
%% Plot Tou verification
newdata = (To)';
cluster_num;
figure(1);
plot(cluster_num,To)
set(gca,'xtick',2:7)
title([fileName, ' ', 'Verification Plot'],'fontsize',12,'fontweight','b');
xlabel('Number of Clusters','fontsize',12,'fontweight','b');
ylabel('Tou Value','fontsize',12,'fontweight','b');
filename1 = strcat(fileName, '_', '_Verification.jpg');
saveas(1,str2mat(filename1));
%% excel file output out Tou value
fileNameRvalue = strcat(fileName, '_TouVALUE.xlsx');
xlswrite(fileNameRvalue,To)

fprintf('\n\nEND of PROGRAM')
beep
pause(0.25)
beep

```

## Appendix K. MATLAB Code for K-mean: vPlotKmean.m

```

%%% File Name: Thesis034_vPlotKmean.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This Thesis034_vPlotKmean.m file input the all three criterion values
% and then plots them while normalizing them between -1 to 1.
% Then the rank of each point is found and assigned a points
% in 10 scale from the highest peak.
% This rank plot is plotted as well as showing the 1st and 2nd
% peak in the voting plot.
% Plot is saved as JPEG.
% THIS IS FOR K-MEAN FOR FILE NAME SIMPLICITY
%-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*
clc; clear all, close all;clear
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
fileName = ('10-18-DEAF-Kmean')
RijValueFile = xlsread ('10-18-DEAF-Kmean-Rij_RVALUE.xlsx');
SihouetteValueFile = xlsread ('10-18-DEAF-Kmean-Silhouette_SWcVALUE.xlsx');
TouValueFile = xlsread ('10-18-DEAF-Kmean-Tou_TouVALUE.xlsx');
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
x =[2;3;4;5;6;7];
%% Rank the three criterion results
[vsR, viR] = sort(RijValueFile);
[xR, vrR] = sort(viR);
[vsS, viS] = sort(SihouetteValueFile);
[xS, vrS] = sort(viS);
[vsT, viT] = sort(TouValueFile);
[xT, vrT] = sort(viT);
VotingValue = 2*vrR + 2*vrS + 2*vrT
VVmax = max(VotingValue);
VVplot = VotingValue / VVmax;
[aa, bb]=sort(VVplot);
[cc,dd] = sort(bb);
vRank(:,1)= VVplot;
vRank(:,2)= dd;
%% have criterion results between -1 to 1
[Rmax, Ri] = max(abs(RijValueFile));
Smax = max(abs(SihouetteValueFile));
Tmax = max(abs(TouValueFile));
Rvalue(:,1) = RijValueFile / Rmax;
Svalue(:,1) = SihouetteValueFile / Smax;
Tvalue(:,1) = TouValueFile / Tmax;
%% plot verification plot
figure(1)
plot(x,Rvalue,'r', 'LineWidth', 3)
hold on
plot(x,Svalue,'g', 'LineWidth', 3)
hold on
plot(x,Tvalue,'b', 'LineWidth', 3)
hold on
plot(x,VVplot,':k', 'LineWidth', 3)
hold on
%% plot the dot on 1st and 2nd highest peak

```



```

for inc = 1:6
    if vRank(inc,2) == 6
        hold on
        plot(inc+1,vRank(inc,1), '*m','LineWidth', 8)
    elseif vRank(inc,2) == 5
        hold on
        plot(inc+1,vRank(inc,1), '*c','LineWidth', 8)

        end
end
ylim([-1.25 1.25])
set(gca,'xtick',2:7)
set(gca,'linewidth',2)
set(gca, 'FontWeight', 'bold')
legend('Rij','Sihouette','Tou','Voting Value', 'Location','SouthWest')
title([fileName, ' ', 'Verification Criterion'], 'FontWeight', 'bold')
xlabel('Number of Clusters', 'FontWeight', 'bold')
ylabel('Index', 'FontWeight', 'bold')
filename = strcat(fileName, '_vPlot.jpg');
saveas(1, str2mat(filename));

fprintf('END of PROGRAM\n\n')
beep
pause(0.25)
beep

```

## Appendix L. MATLAB Code for GMM: GMM.m

```
%%% File Name: Thesis040_GMM.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This Silhouette_GMM.m does GMM analysis and Silhouette Criterion
% of the data set.
% Shows the AE Plot and later will compute the Silhouette Criterion
% All of the centroid, cluster index and Silhouette value will be saved
% in excel.
clc; clear all; close all;
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% Files that has input for analysis
trainFile = xlsread('10_18_DEAF_KNN.xls');
% File that has original AE parameters
readOFile = xlsread('10_18_DEAF_KNN_original.xls');
% Name of the output file
fileName = strcat('10-18-DEAF-GMM')
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %

% Normalization of input data
for i=1:4
    trainFile(:,i)=log(trainFile(:,i));
end
train = [trainFile(:,1),trainFile(:,2), trainFile(:,3), trainFile(:,4)];

k=7;
options = statset('Display','final','maxiter',7500);
for num=2:k
    %% GMM
    gm =
gmdistribution.fit(train,num,'Options',options,'Regularize',0.01,'Replicates'
,100);

    Sx1=0;

    idx = cluster(gm,train);
    idxSave(:,num-1) = idx;
    %% Separate each AE parameter into class
    clusterNum = num2str(num);
    c1count = 1;c2count = 1;c3count = 1;c4count = 1;c5count = 1;c6count =
1;c7count = 1;
    [r c] = size(idx);
    % A      |B      |C      |D      |E      |F
    % Counts  |Energy |Duration |Amp    |Avg Freq |idx
    % 1      |2      |3      |4      |5      |6
    for inc = 1:r
        if num >=2
            if idx(inc,1) == 1
                class1(c1count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
                c1count = c1count+1;
            elseif idx(inc,1) == 2
```

```

        class2(c2count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
        c2count = c2count+1;
    end
end
if num>=3
    if idx(inc,1) == 3
        class3(c3count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
        c3count = c3count+1;
    end
end
if num>=4
    if idx(inc,1) == 4
        class4(c4count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
        c4count = c4count+1;
    end
end
if num>=5
    if idx(inc,1) == 5
        class5(c5count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
        c5count = c5count+1;
    end
end
if num>=6
    if idx(inc,1) == 6
        class6(c6count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
        c6count = c6count+1;
    end
end
if num==7
    if idx(inc,1) == 7
        class7(c7count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
        c7count = c7count+1;
    end
end
end
end
%% Plot AE Plot
figure(2)
%% Duration3 vs. Counts2
subplot(2,2,1)
if num >=2
    plot(class1(:,1),class1(:,3),'.r')
    hold on
    plot(class2(:,1),class2(:,3),'.g')
    hold on
end
if num>=3
    plot(class3(:,1),class3(:,3),'.b')
    hold on
end
if num>=4
    plot(class4(:,1),class4(:,3),'.y')

```

```

        hold on
    end
    if num>=5
        plot(class5(:,1),class5(:,3),'.m')
        hold on
    end
    if num>=6
        plot(class6(:,1),class6(:,3),'.c')
        hold on
    end
    if num==7
        plot(class7(:,1),class7(:,3),'.k')
        hold on
    end
    title([fileName, ' ',clusterNum, ' Counts vs.
Duration'], 'fontsize',12, 'fontweight', 'b')
    xlabel('Counts', 'fontsize',12, 'fontweight', 'b')
    ylabel('Duration ( $\mu$ s)', 'fontsize',12, 'fontweight', 'b')
    axis([0 600 0 30000])
    %% Energy2 vs. Amplitude4
    subplot(2,2,2)
    if num >=2
        plot(class1(:,4),class1(:,2),'.r')
        hold on
        plot(class2(:,4),class2(:,2),'.g')
        hold on
    end
    if num>=3
        plot(class3(:,4),class3(:,2),'.b')
        hold on
    end
    if num>=4
        plot(class4(:,4),class4(:,2),'.y')
        hold on
    end
    if num>=5
        plot(class5(:,4),class5(:,2),'.m')
        hold on
    end
    if num>=6
        plot(class6(:,4),class6(:,2),'.c')
        hold on
    end
    if num==7
        plot(class7(:,4),class7(:,2),'.k')
        hold on
    end
    title(['Energy vs. Amplitude'], 'fontsize',12, 'fontweight', 'b')
    xlabel('Amplitude (dB)', 'fontsize',12, 'fontweight', 'b')
    ylabel('Energy', 'fontsize',12, 'fontweight', 'b')
    axis([45 80 0 4000])
    %% Duration3 vs. Amplitude4
    subplot(2,2,3)
    if num >=2
        plot(class1(:,4),class1(:,3),'.r')
        hold on
        plot(class2(:,4),class2(:,3),'.g')

```

```

        hold on
    end
    if num>=3
        plot(class3(:,4),class3(:,3),'.b')
        hold on
    end
    if num>=4
        plot(class4(:,4),class4(:,3),'.y')
        hold on
    end
    if num>=5
        plot(class5(:,4),class5(:,3),'.m')
        hold on
    end
    if num>=6
        plot(class6(:,4),class6(:,3),'.c')
        hold on
    end
    if num==7
        plot(class7(:,4),class7(:,3),'.k')
        hold on
    end
    title(['Duration vs. Amplitude'],'fontsize',12,'fontweight','b')
    xlabel('Amplitude (dB)','fontsize',12,'fontweight','b')
    ylabel('Duration (μs)','fontsize',12,'fontweight','b')
    axis([45 80 0 30000])
    % Save AE plot to JPEG
    filename = strcat(fileName, '_',clusterNum,'_Subplot.jpg');
    saveas(2,str2mat(filename))
    clearvars inc class1 class2 class3 class4 class5 class6 class7 figure(2)
    close all
    clcount = 1;c2count = 1;c3count = 1;c4count = 1;c5count = 1;c6count =
1;c7count = 1;
end

% Save Swc index number to JPEG
fileNameExIDX = strcat(fileName, '_IDXGMM.xlsx');
xlswrite(fileNameExIDX,idxSave)

beep
pause(0.25)
beep

```

## Appendix M. MATLAB Code for GMM: Rij\_FromGMMIDX.m

```

%%% File Name: Thesis041_Rij_FromGMMIDX.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This Thesis041_Rij_FromGMMIDX.m code is used to obtain the Rij critrion
result.
% The input is (1)excel file with AE input for analysis and (2)excel file
with
% all AE parameters.
% The output is (1)Rij Criterion plot and (2)excel file of Rij criterion
% values.
clc; clear all; close all;
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % CHANGE DATE % % % % % % % % % % %
% Files that has input for analysis
trainFile = xlsread ('10_18_DEAF_KNN.xls');
% File that has original AE parameters
readOFile = xlsread ('10_18_DEAF_KNN_original.xls');
% File that has cluster index number
readIDX = xlsread('10-18-DEAF-GMM_IDXSilhouette.xlsx');
% Name of the output file
fileName = strcat('10-18-DEAF-GMM-Rij')
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %

%% Normalization of input data
for i=1:4
    trainFile(:,i)=log(trainFile(:,i));
end
train = [trainFile(:,1),trainFile(:,2), trainFile(:,3), trainFile(:,4)];

%% Rij Analysis
q=2;p=2;k=7;
for num=2:k
    idx = readIDX(:,num-1);
    for i=1:num
        N0=0;
        A=size(train(idx==i,:));
        T(i,1)=A(1,1);
        D=train(idx==i,:);
        ctrs(i,:)=mean(D);
        for l=1:T(i,1)
            N=(norm(D(l,:)-ctrs(i,:))q);
            N0=N0+N;
        end
        N1(i,1)=N0;
        S(i,1)=(N0/T(i,1))(1/q);
    end
    S;
    for i=1:num
        for j=1:num
            M(i,j)=(sum((((ctrs(i,:)-ctrs(j,:)).2).(1/2)).p))(1/p);
        end
    end
end
M;

```

```

for i=1:num
    for j=1:num
        R1(i,j)=(S(i,1)+S(j,1))/M(i,j);
    end
    R1(i,i)=0;
end
R1;
for i=1:num
    C(i,1)=max(R1(i,:));
end
C;
R(num-1,1)=(sum(C(:,1))./num)
cluster_num(num-1,1)=num;
end
%% Plot Rij verification
newdata = R';
cluster_num;
figure(1)
plot(cluster_num,R.^-1);
set(gca,'xtick',2:7);
title([fileName, ' ', 'Verification Plot'],'fontsize',12,'fontweight','b');
xlabel('Number of Clusters','fontsize',12,'fontweight','b');
ylabel('Rij Value','fontsize',12,'fontweight','b');
filename1 = strcat(fileName, '_', '_Verification.jpg');
saveas(1,str2mat(filename1));

%% excel file output out Rij value
fileNameRvalue = strcat(fileName, '_RVALUE.xlsx');
xlswrite(fileNameRvalue,R.^-1)

fprintf('\n\nEND of PROGRAM')
beep
pause(0.25)
beep

```

## Appendix N. MATLAB Code for GMM: Silhouette\_FromGMMIDX.m

```
%%% File Name: Thesis042_Silhouette_FromGMMIDX.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This Thesis042_Silhouette_FromGMMIDX.m code is used to obtain
% the Silhouette critrion result.
% The input is (1)excel file with AE input for analysis and (2)excel file
with
% all AE parameters.
% The output is (1)Silhouette Criterion plot and (2)excel file of
% Silhouette criterion values.

clc; clear all; close all;
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % CHANGE DATE % % % % % % % % % % % % % % % %
% Files that has input for analysis
trainFile = xlsread ('10_18_DEAF_KNN.xls');
% File that has original AE parameters
readOFile = xlsread ('10_18_DEAF_KNN_original.xls');
% File that has cluster index number
readIDX = xlsread('10-18-DEAF-GMM_IDXSilhouette.xlsx');
% Name of the output file
fileName = strcat('10-18-DEAF-GMM-Silhouette')
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %

%% Normalization of input data
for i=1:4
    trainFile(:,i)=log(trainFile(:,i));%10%9invers
end
train = [trainFile(:,1),trainFile(:,2), trainFile(:,3), trainFile(:,4)]; %10

%% Silhouette Analysis
k=7;
for num=2:k
    Sx1=0;
    idx = readIDX(:,num-1);
    for i=1:num
        D=train(idx==i,:);
        ctrs(i,:)=mean(D);
    end
    for i=1:num
        D=train(idx==i,:);
        ctrs(i,:)=mean(D);
        A=size(train(idx==i,:));
        T(i,1)=A(1,1);
        for g=1:T(i,1)
            N1=0;
            N=(norm(D(g,:)-ctrs(i,:)))/(num-1);
            Nsum(i,g)=N;
            Nsum2(i,g)=inf;
            for h=1:num
                if (h~=i)
                    N3=0;
                    N2=(norm(D(g,:)-ctrs(h,:)))/(num-1);
```



```

        N3=N2+N3;
        Nsum2(i,g)=min(Nsum2(i,g),N3);
    end
    end
    Sx=(Nsum2(i,g)-Nsum(i,g))/max(Nsum2(i,g),Nsum(i,g));
    Sx1=Sx1+Sx;
end
end
SWc(num-1,1)=Sx1/sum(T(:,1));
cluster_num(num-1,1)=num;
end
%% Plot Silhouette verification
newdata = SWc';
cluster_num;
figure(1)
plot(cluster_num,SWc)
set(gca,'xtick',2:7)
title([fileName, ' ', 'Verification Plot'],'fontsize',12,'fontweight','b');
xlabel('Number of Clusters','fontsize',12,'fontweight','b');
ylabel('Silhouette Value','fontsize',12,'fontweight','b');
filename1 = strcat(fileName, '_', '_Verification.jpg');
saveas(1,str2mat(filename1))

%% excel file output out Silhouette value
fileNameRvalue = strcat(fileName, '_SWcVALUE.xlsx');
xlswrite(fileNameRvalue,SWc)

fprintf('\n\nEND of PROGRAM')
beep
pause(0.25)
beep

```

## Appendix O. MATLAB Code for GMM: Tou\_FromGMMIDX.m

```
%%% File Name: Thesis043_Tou_FromGMMIDX.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This Thesis043_Tou_FromGMMIDX.m code is used to obtain the Tou critrion
result
% from GMM idx result.
% The input is (1)excel file with AE input for analysis and (2)excel file
with
% all AE parameters.
% The output is (1)Tou Criterion plot and (2)excel file of Tou criterion
% values.
clc; clear all; close all;

% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % CHANGE DATE % % % % % % % % % % % % % % % % % %
% Files that has input for analysis
trainFile = xlsread ('10_18_DEAF_KNN.xls');
% File that has original AE parameters
readOFile = xlsread ('10_18_DEAF_KNN_original.xls');
% File that has cluster index number
readIDX = xlsread('10-18-DEAF-GMM_IDXSilhouette.xlsx');
% Name of the output file
fileName = strcat('10-18-DEAF-GMM-Tou')
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %

%% Normalization of input data
for i=1:4
    trainFile(:,i)=log(trainFile(:,i));
end
train = [trainFile(:,1),trainFile(:,2), trainFile(:,3), trainFile(:,4)];

%% Tou Analysis
k=7;
for num=2:k
    idx = readIDX(:,num-1);
    for i=1:num
        N0=0;
        A=size(train(idx==i,:));
        T(i,1)=A(1,1); %number of elements in cluster i, Ti
        D=train(idx==i,:);
        ctrs(i,:)=mean(D)
        for l=1:T(i,1)
            N=(norm(D(l,:)-ctrs(i,:)).^2);
            N0=N0+N;
        end
        N1(i,1)=N0;
        S(i,1)=(2*(N0/T(i,1))).^0.5; %Si
    end
    S;
    S1=max(S);
    for i=1:num
        for j=1:num
            M(i,j)=norm(ctrs(i,:)-ctrs(j,:));
        end
    end
end
```

```

        end
        M(i,i)=inf;
    end
    M;
    min(M);
    N=min(min(M));
    To(num-1,1)=N/S1;
    cluster_num(num-1,1)=num;
end
%% Plot Tou verification
newdata = (To)';
cluster_num;
figure(1);
plot(cluster_num,To)
set(gca,'xtick',2:7)
title([fileName, ' ', 'Verification Plot'],'fontsize',12,'fontweight','b');
xlabel('Number of Clusters','fontsize',12,'fontweight','b');
ylabel('Tou Value','fontsize',12,'fontweight','b');
filename1 = strcat(fileName, '_', '_Verification.jpg');
saveas(1,str2mat(filename1));
%% excel file output out Tou value
fileNameRvalue = strcat(fileName, '_TouVALUE.xlsx');
xlswrite(fileNameRvalue,To)

fprintf('\n\nEND of PROGRAM')
beep
pause(0.25)
beep

```

## Appendix P. MATLAB Code for GMM: vPlotGMM.m

```

%%% File Name: Thesis044_vPlotGMM.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This Thesis044_vPlotGMM.m file input the all three criterion values
% and then plots them while normalizing them between -1 to 1.
% Then the rank of each point is found and assigned a points
% in 10 scale from the highest peak.
% This rank plot is plotted as well as showing the 1st and 2nd
% peak in the voting plot.
% Plot is saved as JPEG.
% THIS IS FOR GMM FOR FILE NAME SIMPLICITY
%*****
clc; clear all, close all;
% % % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % CHANGE DATE % % % % % % % %
fileName = ('10-18-DEAF-GMM')
RijValueFile = xlsread ('10-18-DEAF-GMM-Rij_RVALUE.xlsx');
SilhouetteValueFile = xlsread ('10-18-DEAF-GMM-Silhouette_SWcVALUE.xlsx');
TouValueFile = xlsread ('10-18-DEAF-GMM-Tou_TouVALUE.xlsx');
% % % % % % % % % % % % % % % % % % % % % % %
% % % % % % % % % % % % % % % % % % % % % % %
x =[2;3;4;5;6;7];
%% Rank the three criterion results
[vsR, viR] = sort(RijValueFile);
[xR, vrR] = sort(viR);
[vsS, viS] = sort(SilhouetteValueFile);
[xS, vrS] = sort(viS);
[vsT, viT] = sort(TouValueFile);
[xT, vrT] = sort(viT);
VotingValue = 2*vrR + 2*vrS + 2*vrT
VVmax = max(VotingValue);
VVplot = VotingValue / VVmax;
[aa, bb]=sort(VVplot);
[cc,dd] = sort(bb);
vRank(:,1)= VVplot;
vRank(:,2)= dd;
%% have criterion results between -1 to 1
[Rmax, Ri] = max(abs(RijValueFile));
Smax = max(abs(SilhouetteValueFile));
Tmax = max(abs(TouValueFile));
Rvalue(:,1) = RijValueFile / Rmax;
Svalue(:,1) = SilhouetteValueFile / Smax;
Tvalue(:,1) = TouValueFile / Tmax;
%% plot verification plot
figure(1)
plot(x,Rvalue,'r', 'LineWidth', 3)
hold on
plot(x,Svalue,'g', 'LineWidth', 3)
hold on
plot(x,Tvalue,'b', 'LineWidth', 3)
hold on
plot(x,VVplot,':k', 'LineWidth', 3)
hold on
%% plot the dot on 1st and 2nd highest peak

```

```

for inc = 1:6
    if vRank(inc,2) == 6
        hold on
        plot(inc+1,vRank(inc,1), '*m','LineWidth', 8)
    elseif vRank(inc,2) == 5
        hold on
        plot(inc+1,vRank(inc,1), '*c','LineWidth', 8)

        end
end
ylim([-1.25 1.25])
set(gca,'xtick',2:7)
set(gca,'linewidth',2)
set(gca, 'FontWeight', 'bold')
legend('Rij','Sihouette','Tou','Voting Value', 'Location','SouthWest')
title([fileName, ' ', 'Verification Criterion'], 'FontWeight', 'bold')
xlabel('Number of Clusters', 'FontWeight', 'bold')
ylabel('Index', 'FontWeight', 'bold')
filename = strcat(fileName, '_vPlot.jpg');
saveas(1, str2mat(filename));

fprintf('END of PROGRAM\n\n')
beep
pause(0.25)
beep

```

## Appendix Q. MATLAB Code for Visual Verification: AEPlot.m

```
%%% File Name: Thesis050_AEPlot.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This Thesis050_AEPlot.m is capable of plotting the AE plots with three
% different plots in subplot. This code has NO capability of plotting
% multiple sample at the same time. This can be fixed by coding a loop
% inside.
% Input for this code are the following. (1) the excle file that has
% input AE parameters (2) excle file that has cluster index.
% Output for this code is AE plots with 3 subplots.
clc; close all, clear all;

ask1 = input('What is cluster number? ');
while (ask1 < 0 || ask1 > 7)
    ask1 = input('What is cluster number? ');
end
ask2 = input('Enter number for clustering Algorithms. KSOM:[1], Others:[2]');
while (ask2 ~= 1 & ask2 ~=3)
    ask2 = input('Enter number for clustering Algorithms. KSOM:[1],
Others:[3]');
end
if ask2 == 3
    ask3 = input('Enter number for clustering Algorithms. KMEAN:[7],
GMM:[9]');
    while (ask3 ~= 7 & ask3 ~=9)
        ask3 = input('Enter number for clustering Algorithms. KMEAN:[7],
GMM:[9]');
    end
end
if ask2 == 1
    ClusterType = ('KSOM');
end
if ask2 == 3
    ClusterType = ('Kmean/GMM');
end

NumberOfCluster = ask1
ClusterType
num = ask1;

%% Input
if ask2 == 1
    % For KSOM
    % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
    % % % % % % % % % % CHANGE DATE % % % % % % % % % % %
    readOfFile = xlsread (strcat(num, '_DEAF_10_18.xlsx')); % Change Cluster
Number
    indexFile = xlsread('10-18-DEAF-KSOM-Rij_IDX.xlsx');
    fileName = strcat('10-18-DEAF-KSOM')
elseif ask3 == 7
    % For KMEAN
    % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
    % % % % % % % % % % CHANGE DATE % % % % % % % % % % %
    readOfFile = xlsread ('10_18_DEAF_KNN_original.xls');
```

```

    indexFile = xlsread('10-18-DEAF-GMM_IDXKmean.xlsx');
    fileName = strcat('10-18-DEAF-Kmean')
elseif ask3 ==9
    % For GMM
    %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %%
    %% %% %% %% %% %% %% %% %% CHANGE DATE %% %% %% %% %% %% %% %% %% %%
    readOFile = xlsread ('10_18_DEAF_KNN_original.xls');
    indexFile = xlsread('10-18-DEAF-GMM_IDXGMM.xlsx');
    fileName = strcat('10-18-DEAF-GMM')
end
%% Colors
% Original
c1 = '.r';c2 = '.g';c3 = '.b';c4 = '.y';c5 = '.m';c6 = '.c';c7 = '.k';
% IF Change needed
% c1 = '.b';
% c2 = '.m';
% c3 = '.c';
% c4 = '.r';
% c5 = '.g';
% c6 = '.y';
% c7 = '.k';
%% Separate AE parameters in different clusters
idx = indexFile(:,num-1);
c1count = 1;c2count = 1;c3count = 1;c4count = 1;c5count = 1;c6count =
1;c7count = 1;
[r c] = size(idx);
% A      |B      |C      |D      |E      |F
% Counts |Energy |Duration |Amp |Avg Freq |idx
(xx_xx_DEAF_KNN_original.xls)
% Counts |Duration | Energy | Amp | Avg Freq | NN results
(x_DEAF_10_18.xls)
% 1      |2      |3      |4      |5      |6
for inc = 1:r
    if num >=2
        if idx(inc,1) == 1
            class1(c1count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
            c1count = c1count+1;
        elseif idx(inc,1) == 2
            class2(c2count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
            c2count = c2count+1;
        end
    end
    if num>=3
        if idx(inc,1) == 3
            class3(c3count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
            c3count = c3count+1;
        end
    end
    if num>=4
        if idx(inc,1) == 4
            class4(c4count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
            c4count = c4count+1;
        end
    end
end

```

```

end
if num>=5
    if idx(inc,1) == 5
        class5(c5count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
        c5count = c5count+1;
    end
end
if num>=6
    if idx(inc,1) == 6
        class6(c6count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
        c6count = c6count+1;
    end
end
if num==7
    if idx(inc,1) == 7
        class7(c7count,:) = [readOFile(inc,1), readOFile(inc,2),
readOFile(inc,3), readOFile(inc,4), readOFile(inc,5), idx(inc,1)];
        c7count = c7count+1;
    end
end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% For KSOM
% 1      |2          |3          |4          |5          |6
% Counts  |Duration | Energy  | Amp      | Avg Freq  | NN results
(x_DEAF_10_18.xls)
if ask2 == 1
    %% Duration2 vs. Counts1 (KSOM)
    figure(2)
    subplot(2,2,1)
    if num >=2
        plot(class1(:,1),class1(:,2),c1)
        hold on
        plot(class2(:,1),class2(:,2),c2)
        hold on
    end
    if num>=3
        plot(class3(:,1),class3(:,2),c3)
        hold on
    end
    if num>=4
        plot(class4(:,1),class4(:,2),c4)
        hold on
    end
    if num>=5
        plot(class5(:,1),class5(:,2),c5)
        hold on

        plot(class4(:,1),class4(:,2),c4)
        hold on
    end
    if num>=6
        plot(class6(:,1),class6(:,2),c6)
        hold on
    end
end

```



```

if num==7
    plot(class7(:,1),class7(:,2),c7)
    hold on
end
title([fileName, ' ', num2str(num), ' clusters']; [' Counts vs.
Duration']), 'fontsize',12, 'fontweight', 'b')
xlabel('Counts', 'fontsize',12, 'fontweight', 'b')
ylabel('Duration (µs)', 'fontsize',12, 'fontweight', 'b')
axis([0 600 0 30000])
%% Energy3 vs. Amplitude4 (KSOM)
subplot(2,2,2)
if num >=2
    plot(class1(:,4),class1(:,3),c1)
    hold on
    plot(class2(:,4),class2(:,3),c2)
    hold on
end
if num>=3
    plot(class3(:,4),class3(:,3),c3)
    hold on
end
if num>=4
    plot(class4(:,4),class4(:,3),c4)
    hold on
end
if num>=5
    plot(class5(:,4),class5(:,3),c5)
    hold on
end
if num>=6
    plot(class6(:,4),class6(:,3),c6)
    hold on
end
if num==7
    plot(class7(:,4),class7(:,3),c7)
    hold on
end
title(['Energy vs. Amplitude'], 'fontsize',12, 'fontweight', 'b')
xlabel('Amplitude (dB)', 'fontsize',12, 'fontweight', 'b')
ylabel('Energy', 'fontsize',12, 'fontweight', 'b')
axis([45 80 0 4000])
%% Duration2 vs. Amplitude4 (KSOM)
subplot(2,2,3)
if num >=2
    plot(class1(:,4),class1(:,2),c1)
    hold on
    plot(class2(:,4),class2(:,2),c2)
    hold on
end
if num>=3
    plot(class3(:,4),class3(:,2),c3)
    hold on
end
if num>=4
    plot(class4(:,4),class4(:,2),c4)
    hold on
end
end

```

```

if num>=5
    plot(class5(:,4),class5(:,2),c5)
    hold on
end
if num>=6
    plot(class6(:,4),class6(:,2),c6)
    hold on
end
if num==7
    plot(class7(:,4),class7(:,2),c7)
    hold on
end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% For Kmean & GMM
if ask2 == 3
    %% Duration3 vs. Counts2 (Kmean & GMM)
    figure(2)
    subplot(2,2,1)
    if num >=2
        plot(class1(:,1),class1(:,3),c1)
        hold on
        plot(class2(:,1),class2(:,3),c2)
        hold on
    end
    if num>=3
        plot(class3(:,1),class3(:,3),c3)
        hold on
    end
    if num>=4

        plot(class4(:,1),class4(:,3),c4)
        hold on
    end
    if num>=5
        plot(class5(:,1),class5(:,3),c5)
        hold on
    end
    if num>=6
        plot(class6(:,1),class6(:,3),c6)
        hold on
    end
    if num==7
        plot(class7(:,1),class7(:,3),c7)
        hold on
    end
    end
    title({'[fileName, ' ', num2str(num), ' clusters']; [' Counts vs.
Duration']}, 'fontsize',12, 'fontweight', 'b')
    xlabel('Counts', 'fontsize',12, 'fontweight', 'b')
    ylabel('Duration ( $\mu$ s)', 'fontsize',12, 'fontweight', 'b')
    axis([0 600 0 30000])
    %% Energy2 vs. Amplitude4 (Kmean & GMM)
    subplot(2,2,2)
    if num >=2
        plot(class1(:,4),class1(:,2),c1)
        hold on
        plot(class2(:,4),class2(:,2),c2)

```

```

        hold on
    end
    if num>=3
        plot(class3(:,4),class3(:,2),c3)
        hold on
    end
    if num>=4
        plot(class4(:,4),class4(:,2),c4)
        hold on
    end
    if num>=5
        plot(class5(:,4),class5(:,2),c5)
        hold on
    end
    if num>=6
        plot(class6(:,4),class6(:,2),c6)
        hold on
    end
    if num==7
        plot(class7(:,4),class7(:,2),c7)
        hold on
    end
    title(['Energy vs. Amplitude'],'fontsize',12,'fontweight','b')
    xlabel('Amplitude (dB)','fontsize',12,'fontweight','b')
    ylabel('Energy','fontsize',12,'fontweight','b')
    axis([45 80 0 4000])
    %% Duration3 vs. Amplitude4 (Kmean & GMM)
    subplot(2,2,3)
    if num >=2
        plot(class1(:,4),class1(:,3),c1)
        hold on
        plot(class2(:,4),class2(:,3),c2)
        hold on
    end
    if num>=3
        plot(class3(:,4),class3(:,3),c3)
        hold on
    end
    if num>=4
        plot(class4(:,4),class4(:,3),c4)
        hold on
    end
    if num>=5
        plot(class5(:,4),class5(:,3),c5)
        hold on
    end
    if num>=6
        plot(class6(:,4),class6(:,3),c6)
        hold on
    end
    if num==7
        plot(class7(:,4),class7(:,3),c7)
        hold on
    end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Plot

```

```
title(['Duration vs. Amplitude'],'fontsize',12,'fontweight','b')
xlabel('Amplitude (dB)','fontsize',12,'fontweight','b')
ylabel('Duration ( $\mu$ s)','fontsize',12,'fontweight','b')
axis([45 80 0 30000])
filename = strcat(fileName, '_',num2str(num),'_Subplot.jpg');
saveas(2,str2mat(filename))
clearvars inc class1 class2 class3 class4 class5 class6 class7 figure(2)
c1count = 1;c2count = 1;c3count = 1;c4count = 1;c5count = 1;c6count =
1;c7count = 1;

fprintf('\n\nEND of PROGRAM')
beep
pause(0.25)
beep
```

## Appendix R. MATLAB Code for Visual Verification: StatisticalAnalysisIDX.m

```
%%% File Name: Thesis051_StatisticalAnalysisIDX.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This Thesis051_StatisticalAnalysisIDX.m is capable of computing the
% statistical analysis of the classified data from KMEAN and GMM. This is
% because it uses the index number generated by the both functions in
% MATLAB.

clc; clear all; close all;

ask1 = input('What is cluster number? ');
while (ask1 < 0 || ask1 > 7)
    ask1 = input('What is cluster number? ');
end
ask3 = input('Enter number for clustering Algorithms. KMEAN:[7], GMM:[9]');
while (ask3 ~= 7 & ask3 ~=9)
    ask3 = input('Enter number for clustering Algorithms. KMEAN:[7],
GMM:[9]');
end

clustNumIn = ask1; %User Input

if ask3 == 7 % for Kmean
    fileName = strcat('10-18-DEAF-Kmean-Rij_',clustNumIn)
    AEdatfilename = '10_18_DEAF_KNN_original.xls'; % User AE input value
    clusterIDX = xlsread('10-18-DEAF-GMM_IDXKmean.xlsx'); % User input value
elseif ask3 == 9 % for GMM
    fileName = strcat('10-18-DEAF-GMM-Rij_',clustNumIn)
    AEdatfilename = '10_18_DEAF_KNN_original.xls'; % User input value
    clusterIDX = xlsread('10-18-DEAF-GMM_IDXGMM.xlsx'); % User input value
end

net_output = xlsread(AEdatfilename);
clustNum = clustNumIn - 1;
rowC1 = 1; rowC2 = 1; rowC3 = 1; rowC4 = 1; rowC5 = 1;rowC6 = 1;

% % % % % % % % % CHECK YOUR VALUES % % % % % % % % % %
%CEDAF
%12345
for i = 1:length(net_output)
    if clusterIDX(i,clustNum) == 1
        E1(rowC1,:) = net_output(i,2);
        D1(rowC1,:) = net_output(i,3);
        A1(rowC1,:) = net_output(i,4);
        rowC1 = rowC1+1;
    elseif clusterIDX(i,clustNum) == 2
        E2(rowC2,:) = net_output(i,2);
        D2(rowC2,:) = net_output(i,3);
        A2(rowC2,:) = net_output(i,4);
        rowC2 = rowC2+1;
    elseif clusterIDX(i,clustNum) == 3
```

```

        E3(rowC3,:) = net_output(i,2);
        D3(rowC3,:) = net_output(i,3);
        A3(rowC3,:) = net_output(i,4);
        rowC3 = rowC3+1;
    elseif clusterIDX(i,clustNum) == 4
        E4(rowC4,:) = net_output(i,2);
        D4(rowC4,:) = net_output(i,3);
        A4(rowC4,:) = net_output(i,4);
        rowC4 = rowC4+1;
    elseif clusterIDX(i,clustNum) == 5
        E5(rowC5,:) = net_output(i,2);
        D5(rowC5,:) = net_output(i,3);
        A5(rowC5,:) = net_output(i,4);
        rowC5 = rowC5+1;
    elseif clusterIDX(i,clustNum) == 6
        E6(rowC6,:) = net_output(i,2);
        D6(rowC6,:) = net_output(i,3);
        A6(rowC6,:) = net_output(i,4);
        rowC6 = rowC6+1;
    end
end
end
%% Statistical Analysis of Energy
StatE1 = [min(E1); max(E1); mean(E1); std2(E1); length(E1)];
StatE2 = [min(E2); max(E2); mean(E2); std2(E2); length(E2)];
StatE3 = [min(E3); max(E3); mean(E3); std2(E3); length(E3)];
StatE4 = [min(E4); max(E4); mean(E4); std2(E4); length(E4)];
if clustNumIn >= 5
    StatE5 = [min(E5); max(E5); mean(E5); std2(E5); length(E5)];
end
if clustNumIn >= 6
    StatE6 = [min(E6); max(E6); mean(E6); std2(E6); length(E6)];
end
%% Statistical Analysis of Duration
StatD1 = [min(D1); max(D1); mean(D1); std2(D1); length(D1)];
StatD2 = [min(D2); max(D2); mean(D2); std2(D2); length(D2)];
StatD3 = [min(D3); max(D3); mean(D3); std2(D3); length(D3)];
StatD4 = [min(D4); max(D4); mean(D4); std2(D4); length(D4)];
if clustNumIn >= 5
    StatD5 = [min(D5); max(D5); mean(D5); std2(D5); length(D5)];
end
if clustNumIn >= 6
    StatD6 = [min(D6); max(D6); mean(D6); std2(D6); length(D6)];
end
%% Statistical Analysis of Amplitude
StatA1 = [min(A1); max(A1); mean(A1); std2(A1); length(A1)];
StatA2 = [min(A2); max(A2); mean(A2); std2(A2); length(A2)];
StatA3 = [min(A3); max(A3); mean(A3); std2(A3); length(A3)];
StatA4 = [min(A4); max(A4); mean(A4); std2(A4); length(A4)];
if clustNumIn >= 5
    StatA5 = [min(A5); max(A5); mean(A5); std2(A5); length(A5)];
end
if clustNumIn >= 6
    StatA6 = [min(A6); max(A6); mean(A6); std2(A6); length(A6)];
end
end

```

```

%% clustNumIN is between 4 through 5 since it was the # of cluster
% wanted to verify.
if clustNumIn == 4
    State = [State1'; State2'; State3'; State4'];
    StatD = [StatD1'; StatD2'; StatD3'; StatD4'];
    StatA = [StatA1'; StatA2'; StatA3'; StatA4'];
elseif clustNumIn == 5
    State = [State1'; State2'; State3'; State4'; State5'];
    StatD = [StatD1'; StatD2'; StatD3'; StatD4'; StatD5'];
    StatA = [StatA1'; StatA2'; StatA3'; StatA4'; StatA5'];
elseif clustNumIn == 6
    State = [State1'; State2'; State3'; State4'; State5'; State6'];
    StatD = [StatD1'; StatD2'; StatD3'; StatD4'; StatD5'; StatD6'];
    StatA = [StatA1'; StatA2'; StatA3'; StatA4'; StatA5'; StatA6'];
end
%% Export to excel
head = {'Min', 'Max', 'Average', 'SD', '# of Hits'};
fileNameExStat = strcat(fileName, '_Stat.xlsx');
xlswrite(fileNameExStat, head, 'Energy', 'A1')
xlswrite(fileNameExStat, State, 'Energy', 'A2')
xlswrite(fileNameExStat, head, 'Duration', 'A1')
xlswrite(fileNameExStat, StatD, 'Duration', 'A2')
xlswrite(fileNameExStat, head, 'Amplitude', 'A1')
xlswrite(fileNameExStat, StatA, 'Amplitude', 'A2')
close all;

fprintf('END if PROGRAM\n\n')
beep
pause(0.25)
beep

```

## Appendix S. MATLAB Code for Visual Verification: StatisticalAnalysisNNpro.m

```
%%% File Name: Thesis052_StatisticalAnalysisNNpro.m
%%% Date: 11/7/2012
%%% Programmer: JUN SHISHINO
%%% SHM&NDE Lab ERAU LB184
% This Thesis051_StatisticalAnalysisNNpro.m is capable of computing the
% statistical analysis of the classified data from KSOM using NuralWorks
% Professional II/Plus.
clc; clear all; close all;

ask1 = input('What is cluster number? ');
while (ask1 < 0 || ask1 > 7)
    ask1 = input('What is cluster number? ');
end

clustNumIn = ask1; %User Input

fileName = strcat('10-18-DEAF-KSOM-',clustNumIn);
filename1 = strcat(clustNumIn,'_DEAF_10_18.xlsx'); % User input value
net_output = xlsread(filename1);

rowC1 = 1; rowC2 = 1; rowC3 = 1; rowC4 = 1; rowC5 = 1;rowC6 = 1;

% % % % % % % % % CHECK YOUR VALUES % % % % % % % % % %
% 1 = Count
% 2 = Duration
% 3 = Energy
% 4 = Amplitude
% 5 = Avg Freq

for i = 1:length(net_output)
    hold on;
    if net_output(i,6) == 1
        E1(rowC1,:) = net_output(i,3);
        D1(rowC1,:) = net_output(i,2);
        A1(rowC1,:) = net_output(i,4);
        rowC1 = rowC1+1;
    elseif net_output(i,7) == 1
        E2(rowC2,:) = net_output(i,3);
        D2(rowC2,:) = net_output(i,2);
        A2(rowC2,:) = net_output(i,4);
        rowC2 = rowC2+1;
    elseif net_output(i,8) == 1
        E3(rowC3,:) = net_output(i,3);
        D3(rowC3,:) = net_output(i,2);
        A3(rowC3,:) = net_output(i,4);
        rowC3 = rowC3+1;
    elseif net_output(i,9) == 1
        E4(rowC4,:) = net_output(i,3);
        D4(rowC4,:) = net_output(i,2);
        A4(rowC4,:) = net_output(i,4);
        rowC4 = rowC4+1;
    elseif net_output(i,10) == 1
        E5(rowC5,:) = net_output(i,3);
        D5(rowC5,:) = net_output(i,2);
```



```

        A5(rowC5,:) = net_output(i,4);
        rowC5 = rowC5+1;
    elseif net_output(i,11) == 1
        E6(rowC6,:) = net_output(i,3);
        D6(rowC6,:) = net_output(i,2);
        A6(rowC6,:) = net_output(i,4);
        rowC6 = rowC6+1;
    end
end
%% Statistical Analysis of Energy
StatE1 = [min(E1); max(E1); mean(E1); std2(E1); length(E1)];
StatE2 = [min(E2); max(E2); mean(E2); std2(E2); length(E2)];
StatE3 = [min(E3); max(E3); mean(E3); std2(E3); length(E3)];
StatE4 = [min(E4); max(E4); mean(E4); std2(E4); length(E4)];
if clustNumIn >= 5
    StatE5 = [min(E5); max(E5); mean(E5); std2(E5); length(E5)];
end
if clustNumIn >= 6
    StatE6 = [min(E6); max(E6); mean(E6); std2(E6); length(E6)];
end
%% Statistical Analysis of Duration
StatD1 = [min(D1); max(D1); mean(D1); std2(D1); length(D1)];
StatD2 = [min(D2); max(D2); mean(D2); std2(D2); length(D2)];
StatD3 = [min(D3); max(D3); mean(D3); std2(D3); length(D3)];
StatD4 = [min(D4); max(D4); mean(D4); std2(D4); length(D4)];
if clustNumIn >= 5
    StatD5 = [min(D5); max(D5); mean(D5); std2(D5); length(D5)];
end
if clustNumIn >= 6
    StatD6 = [min(D6); max(D6); mean(D6); std2(D6); length(D6)];
end
%% Statistical Analysis of Amplitude
StatA1 = [min(A1); max(A1); mean(A1); std2(A1); length(A1)];
StatA2 = [min(A2); max(A2); mean(A2); std2(A2); length(A2)];
StatA3 = [min(A3); max(A3); mean(A3); std2(A3); length(A3)];
StatA4 = [min(A4); max(A4); mean(A4); std2(A4); length(A4)];
if clustNumIn >= 5
    StatA5 = [min(A5); max(A5); mean(A5); std2(A5); length(A5)];
end
if clustNumIn >= 6
    StatA6 = [min(A6); max(A6); mean(A6); std2(A6); length(A6)];
end
%% clustNumIN is between 4 through 5 since it was the # of cluster
% wanted to verify.
if clustNumIn == 4
    StatE = [StatE1'; StatE2'; StatE3'; StatE4'];
    StatD = [StatD1'; StatD2'; StatD3'; StatD4'];
    StatA = [StatA1'; StatA2'; StatA3'; StatA4'];
elseif clustNumIn == 5
    StatE = [StatE1'; StatE2'; StatE3'; StatE4'; StatE5'];
    StatD = [StatD1'; StatD2'; StatD3'; StatD4'; StatD5'];
    StatA = [StatA1'; StatA2'; StatA3'; StatA4'; StatA5'];
elseif clustNumIn == 6
    StatE = [StatE1'; StatE2'; StatE3'; StatE4'; StatE5'; StatE6'];

```

```

        StatD = [StatD1'; StatD2'; StatD3'; StatD4'; StatD5'; StatD6'];
        StatA = [StatA1'; StatA2'; StatA3'; StatA4'; StatA5'; StatA6'];
end
%% Export to excel
head = {'Min', 'Max', 'Average', 'SD', '# of Hits'};
fileNameExStat = strcat(fileName, '_Stat.xlsx');
xlswrite(fileNameExStat, head, 'Energy', 'A1')
xlswrite(fileNameExStat, StatE, 'Energy', 'A2')
xlswrite(fileNameExStat, head, 'Duration', 'A1')
xlswrite(fileNameExStat, StatD, 'Duration', 'A2')
xlswrite(fileNameExStat, head, 'Amplitude', 'A1')
xlswrite(fileNameExStat, StatA, 'Amplitude', 'A2')
close all;

fprintf('END if PROGRAM\n\n')
beep
pause(0.25)
beep

```