Damage Analysis and Prediction in Glass Fiber Reinforced Polyester Composite Using Acoustic Emission and Machine Learning

S. Gholizadeh

Department of Mechanical Engineering, University of Cape Town, Rondebosch, Cape Town, 7701, South Africa

Abstract

One of the most pervasive types of structural problems in aircraft industries is fatigue cracking that can potentially occur without anticipation with catastrophic failures and unexpected downtime. Acoustic emission (AE) is a passive structural health monitoring (SHM) technique, since it offers real time damage detection based on stress waves generated by cracking in the structure. Machine learning techniques have presented great success over the past few years with a large number of applications. This study assesses the progression of damage occurring on glass fiber reinforced polyester composite specimens using two approaches of machine learning, namely, Supervised and Unsupervised learning. A methodology for damage detection and characterization of composite is presented. The result shows that machine learning can predict damages in composite materials. All predictive models and their performance as well as AE parameters had a direct relationship with the applied stress values, suggesting that these correlations are reliable means of predicting fatigue life in a composite material.

Keywords: Machine learning, Composite, failure mode, fatigue, Acoustic Emission

Introduction

Over the last few decades the use of composite materials has been constantly diversified and composites have been playing crucial rules in different industries; from helicopters and aircrafts to automobile and military services. Material characterization is the critical aspect of the discovery process as well as material designs. Composite offers distinct features such as strength characteristics and stiffness, the absence of corrosion which leads to reduce the cost of the maintenance, low weight, simple design and lower energy consumption. These superior specific features make the composite materials distinctive compared with metals [1].

Composite structures is made from two or more dissimilar materials that are combined together in order to make a new set of characteristics and properties that each component could achieve on their own. Machine learning (ML) is a part of artificial intelligence and subfield of computer science that its foundation is set of statistical tools which focus on the development of models by learning and training data that can be used to predict new data in the future. Machine learning models have acquired lots of attention due to their ability to handle multiple variables through self-improvement without any explicit instructions [2-3].

Generally, Machine learning is a set of algorithms and respective process that create relationship between set of data. These algorithms are broken into four broad categories: 1) Supervised learning, 2) Unsupervised learning, 3) Semi-supervised learning, and 4) Reinforcement learning [4-5]. Supervised learning algorithms are the most widely used algorithms for prediction, provide a learning scheme with “labeled data”, which are developed to classify new data set [6]. Unsupervised learning algorithms are most frequently used for anomaly detection and are related to the pattern detection within the data sets which consists of “unlabeled data”, i.e. data sets can be grouped together based on general rules, natural relationship or natural affinity for each other with unspecified outputs [6].

SHM techniques are usually divided into two groups; damage detection and characterization and Impact detection and identification [7-8]. The goal of SHM system is to perform structural prognosis and diagnosis as well as provide required action for the maintenance engineers and the remaining useful life of the structure. AE can be defined as the energy emitted as a result of changes in the microstructure of a material, which then generates stress waves with transient elasticity [9]. SHM systems are the constant monitoring of structural systems that according to their functionality are categorized to different levels and tasks such as to localize, detect, and assess irregularities and defects. Figure 1 illustrates the level of the complexity of the structures compared with the different types of intelligent level.
From an SHM viewpoint, supervised learning where machines receive inputs and the expected outputs can be employed to detect the severity and types of damage and unsupervised learning where the machine is trained to find similarities in data and is used for detecting the existence of damage by way of clustering of structural response data. The combination of two of the above mentioned schemes represent Semi-supervised learning. This technique typically aims at obtaining a classification of data using both labeled and unlabeled data [10-11]. Machine learning in SHM is using to aim to build models or representations for mapping input patterns in measured sensor data to output targets for damage assessment at different levels [12].

![Figure 1: 6 Levels of SHM [13]](image)

Machine learning algorithms and data-driven techniques have been used widely in cyber-physical systems such as SHM. Das et al investigated classification of the cracks in cementitious components based on the AE signals. In this study RA values (Rise time to maximum amplitude) and Average frequency (AF) are clustered using density dictated unsupervised clustering algorithm and the prediction of the damage state in the structure was obtained. Ince et al reported the locating of the microcracks using multiple-sensor measurements of the acoustic emissions (AEs) by generated by crack inception and propagation and implementing Support Vector Machine (SVM) classifier for recognizing the P-wave arrivals in the presence of noise [14-15].

Supervised and Unsupervised learning has been widely used in different research problems. Vernon et al used multiple linear regression for predicting the density and hardness of the red mud polymer composite. The effects of pH, composition and aggregate particle size of red mud was determined on the properties of recycled polymer red mud composite. Used supervised and unsupervised learning as methods to reduce the computational requirements of reliability analyses [16-17]. The unsupervised learning method, pattern recognition, Fuzzy C-Means clustering (FCM) and Gustafson-Kessel (GK) algorithms were used and tested for evaluation of UAV flight data performance through Receiver Operating Characteristic (ROC) analysis by [18]. Implemented unsupervised pattern recognition analyses (fuzzy C-means clustering) associated with a principal component analysis as tools for the classification of the acoustic emission data [19].

Other machine learning algorithms that have been deployed in different applications include; Artificial Neural Network (ANN), Convolutional Neural Network (CNN), Object recognition, speech recognition, Probabilistic neural network (PNN), Least square support vector machines (LSSVM) [20-24]. Have suggested and formulated seven axioms for SHM that stated the general aspects and principles that have merged in several decades of experience on SHM [25]. The following are particular parts of their states that relevant to this paper:

**Axiom III:** Identifying the existence and location of damage can be done in an unsupervised learning mode, but identifying the type of damage present and the damage severity can generally only be done in a supervised learning mode.

**Axiom IVa:** Sensors cannot measure damage. Feature extraction through signal processing and statistical classification is necessary to convert sensor data into damage information.

**Axiom IVb:** Without intelligent feature extraction, the more sensitive a measurement is to damage, the more sensitive it is to changing operational and environmental conditions.

**Axiom V:** The length- and time-scales associated with damage initiation and evolutions dictate the required properties of the SHM sensing system.

In this paper, AE basic parameters namely amplitude, energy, counts, duration, rise time as well as signal strength were used to develop linear regression as well as multiple linear regression model, and AE clustering using K-Means method to predict and assess the onset damage and damage propagation in glass fiber reinforced polyester composite material under cyclic fatigue test. The result shows that supervised and unsupervised learning can identify and predicts damages in composite materials. The clusters that were contributed through the AE data were taken from the cyclic fatigue tests. The predictive models and their performance as well as AE parameters had a direct relationship with the applied stress values, suggesting that these correlations are reliable means of predicting fatigue life in a composite material. The data that has been used were the same data that was used in previous published paper [26].

**Experimental**

Woven reinforced glass fibers were cut into 30×30 cm sizes to make glass fiber reinforced composites consisting of 40wt% (40 percent by weight) glass fiber, 60wt% matrix (GP 268 BQT-W) and 2wt% hardener. A total of 19 specimens, each 5 mm thick,
250 mm length and 25 mm wide (Figure 2), were cut according to ASTM D3039. Three samples were subjected to a tensile test (see below) before carrying out the cyclic fatigue test, in accordance with ASTM D3479.

**Figure 2: Specimen Geometry**

**Tensile testing techniques**

The goal of the prior tensile tests in this study was to determine the ultimate tensile strength (UTS) of the materials. The three specimens were tested on a universal machine testing system type INSTRON 3382, with a 100 kN capacity. Average UTS of 135.5 MPa was obtained from the tensile test, and this UTS was used as the basis for subsequent experiments.

**Fatigue testing with acoustic emission sensor attachment**

The specimens were installed into the test rig under a one-point test setup (AE sensor), as shown in Figure 3. They were loaded using a 100 kN hydraulic MTS test machine. The load was applied in a sinusoidal waveform. An AE sensor was attached on the center of the surface of each specimen, as the position shown in Figure 4.

The MTS 647 Hydraulic Wedge Grip testing machine with a maximum load of 100 kN and maximum pressure of 21 MPa or 3000 psi were used to apply the load to the specimens until the specimens separated. A set of 16 specimens were loaded under tension-tension cyclic loading, at a frequency of 8 Hz with 45% to 60% of UTS, a maximum load of 10.07 kN and minimum load of 7.33 kN, and at a stress ratio of R = 0.1 with 3.5 MPa of pressure. Load data from the testing machine was fed to the parametric channel of the AE data acquisition system. This load data was recorded simultaneously with the transient AE signals detected during the test. Both sets of data were then used to characterize the AE source mechanism by correlating the AE parameters with the load values.
Acoustic Emission (AE)
A MISTRAS AE system from the Physical Acoustic Corporation Two-Channel was used to acquire the AE signals released by fatigue crack growth during the tests. One wideband (WSa) AE transducer with a frequency range of 100 to 1000 kHz was used to detect the AE signals from the fatigue test at the center of the specimen. This sensor was attached to the specimens and connected to the AE data acquisition system through a coaxial cable. A 40dB threshold level for AE data acquisition was set to avoid interference from any background environmental noise below this level.

The detected events were amplified by a 26dB pre-amplifier and a 40dB amplifier. All the recorded signals were stored on the computer for further analysis. AE WinTM software for data acquisition and signal processing was used throughout this study to capture, replay and display stored AE data. To ensure proper AE monitoring, certain parameters of the data acquisition systems needed to be adjusted to the specific testing materials and existing noise levels: in particular, Peak Definition Time (PDT), Hit Definition Time (HDT) and Hit Lockout Time (HLT). The specific values used for these timing parameters of the signal acquisition process are shown in Table 1.

Table 1: AE control parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Set Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak Definition time (PDT)</td>
<td>50 μs</td>
</tr>
<tr>
<td>Hit Definition time (HDT)</td>
<td>150 μs</td>
</tr>
<tr>
<td>Hit Lockout time (HLT)</td>
<td>300 μs</td>
</tr>
<tr>
<td>Sample rate</td>
<td>5 M sample/sec</td>
</tr>
</tbody>
</table>

Methodology and Analysis
Proposed Methodology
In order to analyze behavior of composite and damage mechanisms in targeted materials, machine learning techniques comprises three sequential techniques: (a) Data preparation, (b) Machine learning modeling, (c) Model evaluation, and each of these steps has their own stages. The schematic of these steps is shown in Figure 5.

Data Preparation
At this first step, all relevant data to the targeted damage mechanisms and composite material’s behavior were collected at University Putra Malaysia (UPM) laboratory in order to gather a comprehensive database to construct the ML model. All values have been assessed to ensure data quality, the accuracy of the model, and verified that no erroneous value was included in the database mistakenly. Relevant materials and variables have been classified into different groups for learning algorithms.

As mentioned in previous study [26] some AE parameters such as energy, amplitude and root mean square (RMS) increase along with the load increment, consequently crack propagation increased as the load was increased, therefore the AE signal patterns were recorded in the early stages of testing and as the fracture zone approached, at applied stress levels of 45%, 50%, 55%, and 60% of...
UTS with a loading of 60.97 MPa, 67.75 MPa, 74.52 MPa, and 81.30 MPa respectively (Table 2).

Table 2: Loading and percentage of UTS

<table>
<thead>
<tr>
<th>Percentage of UTS</th>
<th>Applied load (MPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>45%</td>
<td>60.97</td>
</tr>
<tr>
<td>50%</td>
<td>67.75</td>
</tr>
<tr>
<td>55%</td>
<td>74.52</td>
</tr>
<tr>
<td>60%</td>
<td>81.30</td>
</tr>
</tbody>
</table>

Table 2: Loading and percentage of UTS

<table>
<thead>
<tr>
<th>Percentage of UTS</th>
<th>Applied load (MPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>45%</td>
<td>60.97</td>
</tr>
<tr>
<td>50%</td>
<td>67.75</td>
</tr>
<tr>
<td>55%</td>
<td>74.52</td>
</tr>
<tr>
<td>60%</td>
<td>81.30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Percentage of UTS</th>
<th>Applied load (MPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>45%</td>
<td>60.97</td>
</tr>
<tr>
<td>50%</td>
<td>67.75</td>
</tr>
<tr>
<td>55%</td>
<td>74.52</td>
</tr>
<tr>
<td>60%</td>
<td>81.30</td>
</tr>
</tbody>
</table>

ML Model Building

Technique selection
There are many ML algorithms and techniques out there that have been developed for different types of learning purposes such as unsupervised learning, supervised learning, and deep learning. In this research the following ML modes has been used in order to identify damage severity and types of damage based on AE parameters.

Supervised Learning
Supervised learning is the most simple and basic method that has been used widely in scientific and engineering fields [27-29]. A linear regression is the approximation of a linear model with high interpretability used to describe the relationship between two variables; input and output. Multiple-linear regression is the extension of the simple linear regression, there is more than one independent variable in the process. In this study four predictor variables have been used.

Unsupervised Learning
Unsupervised learning is the algorithm in which the model works on its own to discover data. One of the popular partitioning methods that have been used vastly is K-means clustering. K-means has been implemented in many data science applications, especially when you need to quickly discover insights from unlabeled data. In this study to determine damage severity, K-means clustering was implemented to group and cluster signals. The numbers of 3 clusters with random centroids were implemented in to the formula in order to classify damages.

The objective functions for K-means are shown in the following formula:

$$ J = \sum_{j=1}^{K} \sum_{i=1}^{n} ||x^{(i)} - \mu_j || $$

Where we are looping over all centroids, ||x^{(i)} - \mu_j || is the Euclidian distance, which is the distance between x^{(i)} data point belonging to cluster j and \mu_j centroid. K is number of clusters and n is the number of data point. Clustering is a rich algorithmic framework and conceptual that remarkably is used as a general methodology for data analysis and interpretation [19-31]. Some real-world applications of K-means are customer segmentation, pattern recognition, data compression and etc.

Results and Discussion

Models Building and Evaluation
Based on previous paper [26] the range of 40 dB to 55 dB of amplitude is associated to matrix cracking, 55 dB to 60 dB is AE indication of interface failure, 65 dB to 85 dB related to fiber debonding, 65 dB to 85 dB is linked to fiber pull out and 85 dB to 100 dB is related to fiber fracture cracks. Figure 6 illustrates the scatter plot of the number of amplitude and strength of the signal in the early stage of testing and near fracture zone at different applied stress using multiple linear regressions. Before the main crack occurred, some micro-cracks appeared in the matrix resin and there were also signs of fiber debonding between 40 dB to 60 dB of amplitude. The composite also showed increased delamination between 50 to 70 dB. Some larger cracks appeared between 60 dB to 80 dB, due to fiber pull-out. Once the specimens started to divide into two main pieces, the strength of signal reached the peak level until the fiber broke since it relates to the vibration signal in a time series, which could be expected to increase sharply upon breakage. As the level of applied stress was increased, the fracture shows the high level of amplitude.

From Seaborn data visualization (Python data visualization library), the direct correlation between signal strength and amplitude based on the amount of energy is apparent. As can be seen in Figure 6, the cracks were more scattered in the lower level of applied loads compared to the higher level at the early stage. As the loads were increasing, the amount of energy became more intense, and the signal strength was increased. At the end of the testing, the signal strength increment shows that specimen is close to rupture while loading was increasing. However the amount of energy and signal strength was more scattered towards higher amplitude which indicates the fiber fracture.

Considering above damage details regarding matrix cracking and fiber debonding and pull out in different range of Amplitude from 40 dB to 85dB, from Figure 6 the level of energy and single strength (pVs) can be a good indicator in order to predict and assess damages in the specimens. As the signal strength varied between 20000 at the early stage of testing to 80000 near fracture zone and amount of energy became more intense. The different colours show different levels of energy.
45% of UTS

50% of UTS

55% of UTS

60% of UTS

**Figure 6:** Signal strength versus Amplitude at different applied stress level based on amount of energy a) Early stage of testing, b) Near fracture zone
Creating Train and Test Dataset
An important step in testing the model is to split data into training and testing data. The dataset has been split into training and testing sets with 70% of the entire data for training and 30% for testing. This will provide a more accurate evaluation on out-of-sample accuracy because the testing dataset is not part of the dataset that have been used to train the data. It is more realistic for real world problems. Figure 7 shows the regression plot of the training data and actual data near fracture zone.

**Figure 7:** Signal strength vs Amplitude at different applied stress level near fracture zone (a) before training (actual data) (b) after training (70% of the entire data)
Figure 8 shows the correlation between different AE parameters at the minimum and maximum applied load. As it can be seen, the least correlation is related to RMS to other parameters and energy and signal strength have the strong correlation with each other. Therefore RMS didn’t count as the strong predictor parameter and was dropped as a predictor parameter in the following models.

The sklearn package, one of the machine learning libraries, was used to analyze multi-linear regression and identify the damage severity as well as types of damage based on the amplitude. Figure 9 shows the prediction and actual value of amplitude at different applied stress level with 30% of testing data near fracture zone.
The predictor variables that have been used to predict the damage type and severity based on the amplitude level were based on the AE basic parameters: rise time, count, energy, and duration. Based on the Figure 9, damages were classified based on amplitude level. The illustration shows the training and testing prediction of amplitude with comparison between actual and predicted values at different applied stress.

The existence of the damage in the materials were detected through unsupervised learning where machine is trained and is used by way of clustering of structural response data. K-Means clustering has many parameters that can be used. For the modeling of this datasets, the number of 3 cluster indicating matrix cracks, fiber debonding, and fiber pull out to form the K-means algorithm has been run with random centroid seeds. Figure 10 depicts the K-means class of amplitude versus duration and energy which is contributed to damage of the composite and partitioned the signals into exclusive groups near fracture zone at different applied stress. The clustering shows the types of damages.

Figure 10(a) shows the scatter plot of datasets based on amount of energy and count that is scattered based on amplitude. The scatter plot depicts the different level of amplitude near fracture zone under minimum and maximum applied load. Figure 10(b) illustrates the K-means clustering prediction, the number of 3 clusters was applied with random centroids in order to predict the amplitude and classify severity of damages into 3 groups. It can be seen that the K-means clustering algorithm has produced 3 clusters fairly similar to plot (a). We can now make predictions based on these clusters and centroids.

In 45% of UTS applied stress:
- Cluster 0 is most likely refers to matrix cracking and interface failure.
- Cluster 1 is most likely refers to fiber pull out.
- And Cluster 2 is most likely refers to fiber debonding.

In 60% of UTS applied stress:
- Cluster 0 is most likely refers to fiber pull out.
- Cluster 1 is most likely refers to matrix cracking and interface failure.
- And Cluster 2 is most likely refers to fiber debonding.

Figure 10: K-means clustering of AE strength of signal vs Amplitude vs Duration at different applied stress: (a) scatter plot (b) K-means clustering
Conclusion
The present study looked at the capability of the AE technique and machine learning algorithms to assess and predict the onset of damage in glass fiber reinforced polyester composite material. AE basic parameters namely amplitude, energy, counts and duration were used to develop multiple linear regression model, as well as AE clustering using K-Means method to predict and assess the onset damage and damage propagation in glass fiber reinforced polyester composite material under cyclic fatigue test. According to the results, most fiber breakage occurred at between 85 dB and 100 dB. Higher signal strength values were observed for both fiber breakage and matrix cracks as the same result obtained before. There was a satisfactory level of agreement between the signal strength, amount of energy and amplitude suggests that these parameters are a reliable way of identifying different types of damage in composite materials. All predictive models and their performance as well as AE parameters had a direct relationship with the applied stress values, suggesting that these correlation coefficients are reliable means of predicting fatigue life in a composite material.

Data Availability Statement
The data that support the findings of this study were generated at UPM laboratory. The data that supporting this study is the same data that is available within the published author’s articles: DOI: 10.1080/14484846.2016.1264284, and ISBN: 978-93-85973-63-5.

References
23. De Oliveira, M. A., Araujo, N. V., Da Silva, R. N., Da Silva, T.