Fatigue Feature Classification for Automotive Strain Data

To cite this article: M F M Yunoh *et al* 2012 *IOP Conf. Ser.: Mater. Sci. Eng.* **36** 012031

Fatigue Feature Classification for Automotive Strain Data

M.F.M. Yunoh[1,3]**, S. Abdullah**[1]**, Z.M. Nopiah**[1]**, M.Z. Nuawi**[1]**and N. Ismail**[2]

[1]Dept. of Mechanical and Material, Faculty of Engineering and Built Enviroment, UniversitiKebangsaan Malaysia, 43600 UKM, Bangi Selangor, Malaysia.
[2]Faculty of Mechanical Engineering, Universiti Malaysia Pahang, 2660 Pekan, Pahang, Malaysia.

E-mail: shahrum@eng.ukm.my

Abstract. Fatigue strain signal were analysed using data segmentation and data clustering. For data segmentation, value of fatigue damage and global statistical signal analysis such as kurtosis was obtained using specific software. Data clustering were carried out using K-Mean clustering approaches. The objective function was calculated in order to determine the best numbers of groups. This method is used to calculate the average distance of each data in the group from its centroid. Finally, the fatigue failure indexes of metallic components were generated from the best number of group that has been acquired. Based on four data collect from two different roads which are D1, D2, the index value generated is not the same for all of data because due to K-Mean clustering, the best group is different for each of the data used. The maximum indexes generated are different for two types of road and namely the index 4 for D1 and index 5 for D2. Due to the road surface condition, higher distributions of the best groups give higher values of index and reflect to higher fatigue damage experienced by the system.

1. Introduction

Many engineering failures especially in mechanical engineering field involve a great loss in term of life and property. Phenomenon of fatigue failure is major factors contribute to the failure of products, component and structures in this field. Basically fatigue failure often occurs when repeated stresses imposed on the structural components within a certain time. More than 50% of mechanical failures lead by fatigue failure and mostly it unpredictable [1].

Material fatigue is a one of the most safety issues for structures subject to the cyclic loads and the cause of failure in a majority cases. In the automotive industry, most of the components are subjected in service to fatigue loading which may result in failures [2]. Most industries are susceptible to fatigue failure due to repetitive loading used every day. Nowadays, there are many techniques and tools developed to detect and prevent fatigue failure. Those techniques need to do the laboratory test, data collection, specific tools and software in order to do fatigue assessment [3].

Fatigue feature classification is a new analysis in fatigue related field and not been studied widely. Fatigue feature classification should be studied in order to produce a new standard and enable a design to the same reference level. When referring to some related studies. Mao et al. [4] had introduced a safety index of fatigue failure for ship structure details. In their research, the expected fatigue damage and its coefficient are estimated using Gussian and Bayesian process in order to

[1] Corresponding author. E-mail: shahrum@eng.ukm.my.

develop the index. In addition, fatigue life under Gaussian load was also performed in order to compute a fatigue risk of a component and the safety index [5].

In this paper, the main focus is developing a new fatigue feature classification using the data editing, segmentation and clustering approach to predict fatigue life of suspension car system names as fatigue failure indexes. Fatigue failure indexes are very important in component design in order to predict components life and to prevent sudden breakdown or failure [6]. Thus the developing a new fatigue failure indexes for automotive component will give a great advantage to automotive industry. Besides that, it will reduce maintenance cost, replacement a time saving.

2. Literature Background

2.1 Global Signal Statistical Parameters

In fatigue research area, the signals consist of a measurement of cyclic loads, i.e. force, strain, and stress against time. Time series typically consists of a set of observations of a variable were taken at equally spaced intervals of time. Global signal statistical parameters are frequently used to classify random signals and monitor the pattern of analyzed signals. For a signal with a numbers of data point n in a sampled sequence, the mean is given by: x

$$\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j \tag{1}$$

In the fatigue signals, the calculation of the root-mean-square (r.m.s.) and the kurtosis are very important in order to retain a certain amount of the signal amplitude range characteristics. Both values are defined respectively as [7]:

$$r.m.s = \left\{\frac{1}{n}\sum_{j=1}^{n} x_j^{\,2}\right\}^{1/2} \tag{2}$$

$$K = \frac{1}{n(r.m.s)^4}\sum_{j-1}^{n}\left(x_j - \bar{x}\right)^4 \tag{3}$$

wherexj is the amplitude of signal.

2.2 Total Fatigue Damage

Three major approaches have widely been used toanalyse fatigue damage or fatigue life, namely thestress-life approach (*S-N*), the strain-life approach(ε-*N*) and the linear elastic fracture mechanics(*LEFM*) [8]. However, the strain-life approach (ε-*N*)is used for the analysis as the case study was relatedto low cycle fatigue, which is a suitable approach toanalyse random data collected from automotive components. ThePalmgren-Miner linear cumulative damaging rule normallyassociated with the established strain-life fatigue damaging models, i.e. the Coffin-Manson, the Morrow, and the Smith-Watson-Topper (SWT). Nevertheless, the Coffin-Manson relationship only considers the damaging calculation at zero mean stress. However, in real situation, some of the realistic service situations involve nonzero mean stresses. For example, in a case of the loading being predominantly compressive, particularly for wholly compressive cycles, the Morrow mean stress correction effect provides more realistic life estimates and seems to work reasonably well for steels. The strain-life model is mathematically defined as the following expression [9].

$$\varepsilon_a = \frac{\sigma'_f}{E}(2N_f)^b + \varepsilon'_f(2N_f)^c \tag{4}$$

where εa is the true strain amplitude, σ'f is the fatigue strength coefficient, E is the material modulus of elasticity, σm is the mean stress, Nf is the numbers of cycle to failure for a particular stress range and mean, b is the fatigue strength exponent, ε'f is the fatigue ductility coefficient, and c is the fatigue ductility exponent. The fatigue damage caused by each cycle of repeated loading is calculated by reference to material life curves, such as S-N or ε-N curves. The fatigue damage D for one cycle and the total fatigue damage ΣD caused by cycles are expressed respectively as [10]:

$$D = \left(\frac{1}{N_f}\right) \quad (5)$$

$$\Sigma D = \sum\left(\frac{N_i}{N_f}\right) \quad (6)$$

where*Ni* is the numbers of cycle within a particular stress range and mean.

2.3 Feature Classification

To extracting the information from the data columns, it is important for researcher to find the methods that can divide the data into specific parts and analysed each part in optimum [11]. Segmentation divided the non stationary times series segments with the constant statistics value. Therefore, segmentation technique can identify accurately and quickly the location changes in statistical signal [12]. Segmentation enables to divide the time series data and organised into several sections that have the same criteria [13]. Recently, many methods have been designed to segment the data as accurately as possible, but most of the methods have an own name and application. Thus, the segmentation algorithms in time series can be divided into three main groups which is sliding windows, top-down and bottom-up [14].

Clustering is a process a division of data into the same groups whereas the data in the groups will have the similar character and behaviour. There are two category of the clustering which is partitioning and hierarchical. For partitioning clustering, generally result in a set of M clusters, each object belonging to one cluster and each cluster may be represented by a centroid or a cluster.K-means is a fast and popular method to perform clustering because it easier to manage the huge of data [15]. The basic intuition behind K-means is the continuous reassignment of objects into different clusters so that the within-cluster distance is minimized. It uses an iterative algorithm divided in two phases to minimize the sum of point-to-centroid distances, over all K clusters[14]. Objective functions for K-means define as following expression:

$$J_h = \sum_{j=1}^{k} \sum_{i}^{c} \left\| x_i^{(j)} - c_j \right\|^2 \quad (7)$$

Where $\|x_i^{(j)} - c_j\|^2$is distances between point $x_i^{(j)}$ and mean of point c_j.

3. Methodology

Four data obtained from the suspension and lower arm is collected during the test at highway and country road. The highway and country road is then referred as D1 and D2 in this paper. The suspension and lower arm is attached with strain gauge in order to collect the strain signal during the test at D1 and D2 road. The data is recorded for 60 second with sampling frequency 500Hz gives 30000 discrete data[9]. This sampling frequency is chosen to be at 500Hz as the value was enough to record the data because of the strain gauge sensitivity in recording the changes of the strain response that exit during the test.

Figure 1shows the sequence method for data collection toward car suspension system. The strain gauge is attached at high stress region of the suspension and lower arm. The high stress region can be determined using finite element analysis (FEA). Strain gauge then connected to the data acquisition system for the purposed of strain signal data collection. The strain data is collected in the time history domain. The strain signal data after that is analysed using global statistical analysis in order to find the statistical parameter i.e kurtosis, root mean square (r.m.s)and fatigue damage that use for data clustering.

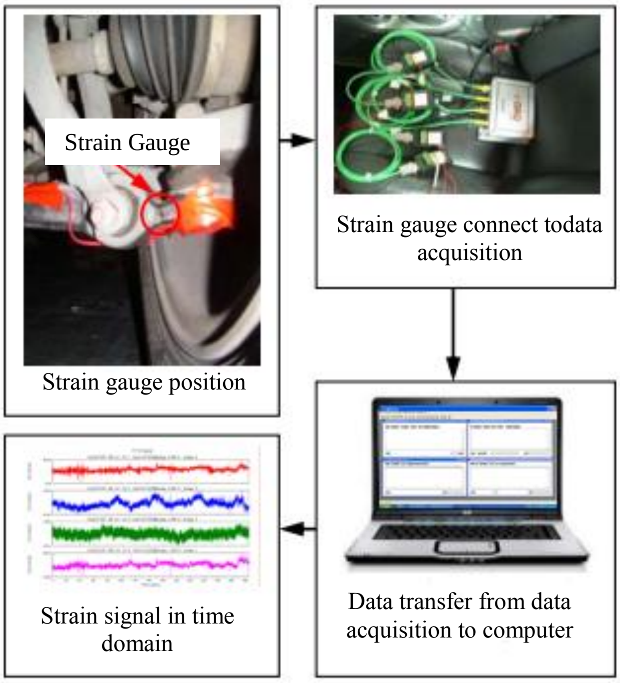


Figure 1. Sequence method for data collection toward car suspension system

Data segmentation and clustering is done by using Matlabsoftware. Data segmentation one of the processes divides data into certain section or segment. Segmentation method can be grouped into three categories such as sliding windows, top-down and bottom-up [14]. Segmentation using bottom-up method is the best approach segmentation for time history signal. Clustering data is analysed using K-mean approach as a clustering algorithm. Cluster with certain centroid is formed with the different colour for every object. The clustering is analysed based on the changing of fatigue damage. The best cluster is determined based on function objective value through Euclidean distance calculation. Euclidean distance gave the distance value for every cluster to centroid. The classification later is generated from the clustering data. The fatigue failure indexes are generated based on classification result.

4. Results and Discussion

Figure 2 (a) shows the data distribution with two group point for D1 road. The positions of the centroid are very close between each other and concentratewith the increasing of kurtosis. It show that the K-means clustering susceptible in initial centroid because the intake of random centroid point. Roughly, it can been seen more than two groups could be formed from the distribution of the data because there is a point that has far distances away from it centroid

The formation of the clusters can clearly see for three point centroid as shows in Figure 2 (b). There was a little difference in the distribution according to the groups of the fatigue damage in red and yellow. The four centroid point of data distribution as shows in Figure 2 (c) more organized than Figure 2 (d). Each of the centroid has a data distribution with uniform distances. The five centroid point of data distribution shows a little visible difference compared to the four centroid as the data seemingly too close between each group with low fatigue damage. Number of different group used to see the density distribution of the data in group. The data with the lower number of centroid have a high density distribution and the objective function also high.

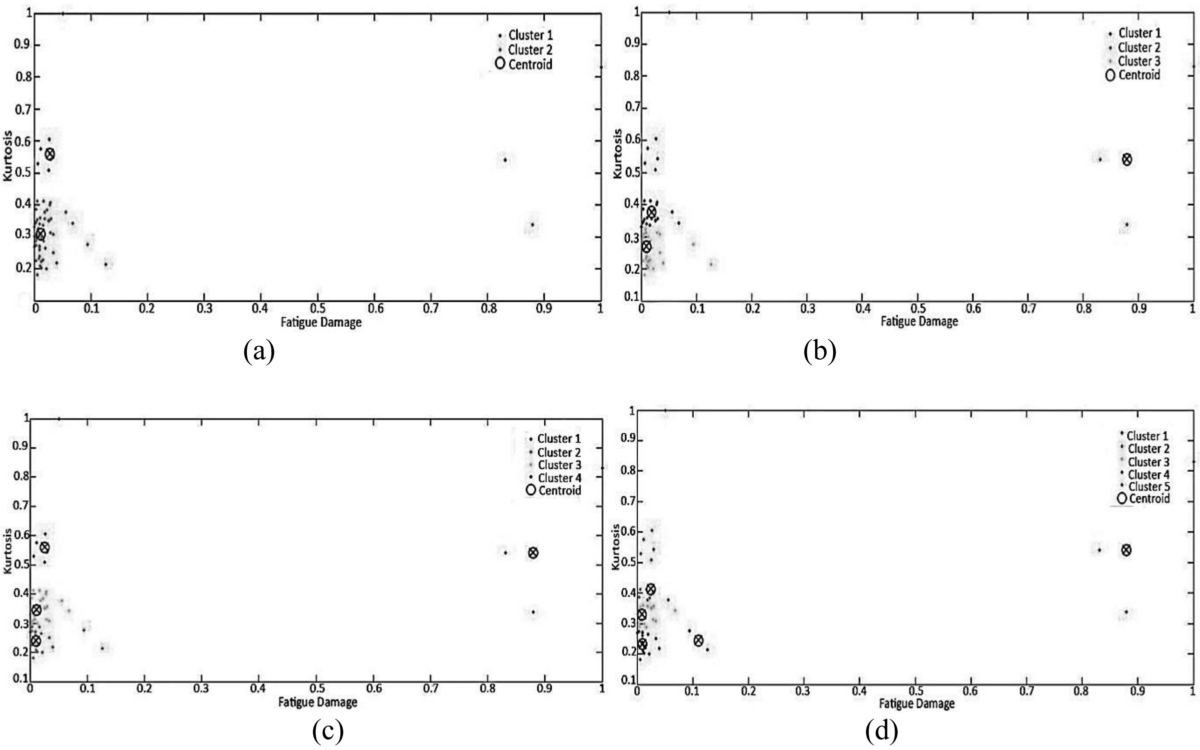
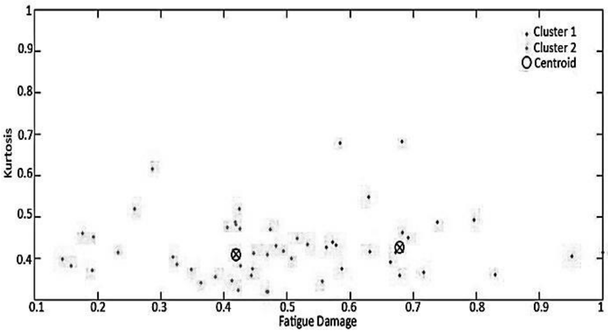


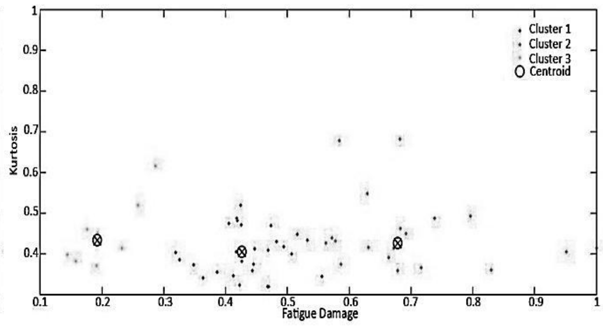
Figure 2. Kurtosis distributions of fatigue damage for D1 road condition; (a) two centroid, (b) three centroid, (c) four centroid, (d) five centroid

The clustering for the D2 road is start from two groups to five groups. Figure 3 (a) shows that the two centroid points is far between each other. However, both of the groups have a clear boundary between each other. This situation happen because of the data distribution is not too dense with each other because the surface D2 road always changing. Figure 3 (b) shows the group's formation seen more clearly for three points centroid. For the each group, there are a clear separation of data distribution and all of the data linear with the fatigue damage value. It also has a high kurtosis value with the ratio 1:1 after normalised in the data distribution and it namely outlier data

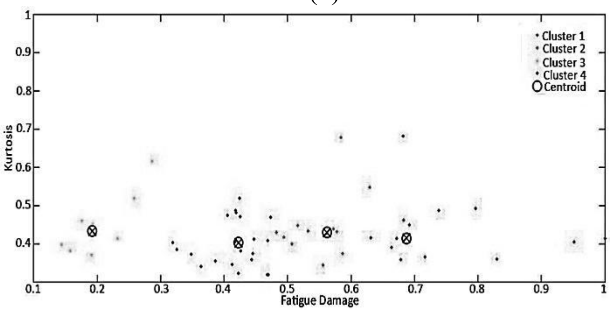
The distribution data for four and five centroid point as shown in figure 3 (c) and (d) is more organised compare to the two and three cnteroid point. Each of the centroid have a distribution data with a uniform distances. The boundary of the groups is clearly showed and the kurtosis ratio 1:1 still retained as a outlier data. However, the data with five groups is chosen as a best group because it have a small objective function.



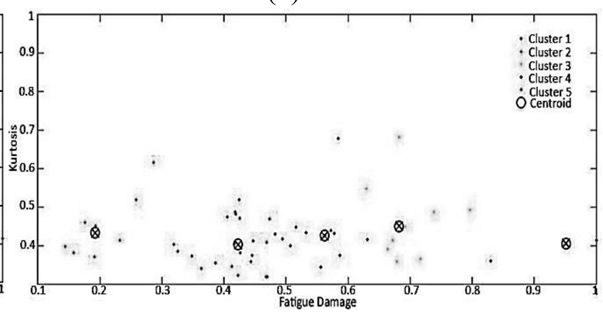
(a)



(b)



(c)



(d)

Figure 3. Kurtosis distributions of fatigue damage for D2 road condition; (a) two centroid, (b) three centroid, (c) four centroid, (d) five centroid

The maximum objective function value for D1 road occurs at data with two centroid as shown in Table 1. The objective function value is decrease with the increasing of centroid number but only for the data with five centroid has less objective function value than the data with four centroid. It show that the data with five group not necessary to generated because it only scattered the data getting away from the centroid. Therefore, the four clusters is enough for the D1 road. For the D2 road, the maximum objective function value also occurs at data with two centroid as shown in Table 1. The objective function value for five clusters is the best for D2 road. The increase to more than five centroid may cause the objective function value decreasing, but the clusters distribution after that not have many difference with the five clusters data.

Table 1. Average objective function value for D1 and D2 road.

	Objective Function value for D1 Road	Objective Function value for D2 Road
C=2	6.89	8.48
C=3	4.96	7.91
C=4	4.23	6.52
C=5	4.33	5.84

Fatigue failure index is generated using the data with four centroid as shown in Figure 4 for D1 road. The index from 1 to 4 is given for each group. The indexes that have a high value represent the group data that have high fatigue damage. Thus, it give a high probability of fatigue failure occurs. Index 1 to 3 is classification as an index with low fatigue damage. Index 4 is classified as index with high fatigue damage.

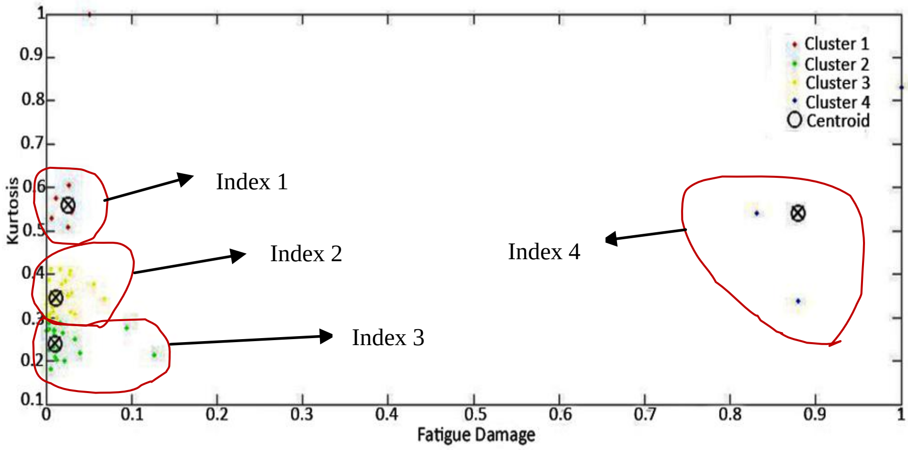


Figure 4. Selection of group to generate the index for D1 road

For D2 road, the fatigue failure index is generated using the data with five centroid as shown in Figure 5. Indexes from 1 to 5 are given for each group and have a high value represent the group data that have high fatigue damage. Group 1 with a green distribution point has a low fatigue damage value and it namely as an index 1. The second group with purple distribution data is namely as index 2. The third group with blue distribution data is namely as index 3. The second last group with the yellow distribution data start to show the high fatigue damage and it namely as index 4. For the last group with red distribution data is namely as index 5 shows the high fatigue damage. The outlier data is recognised for the data distribution at kurtosis ratio 0.7 and above. Although, the outlier data is not include in the group that given an index, this data cannot remove because it give the high amplitude for overall data.

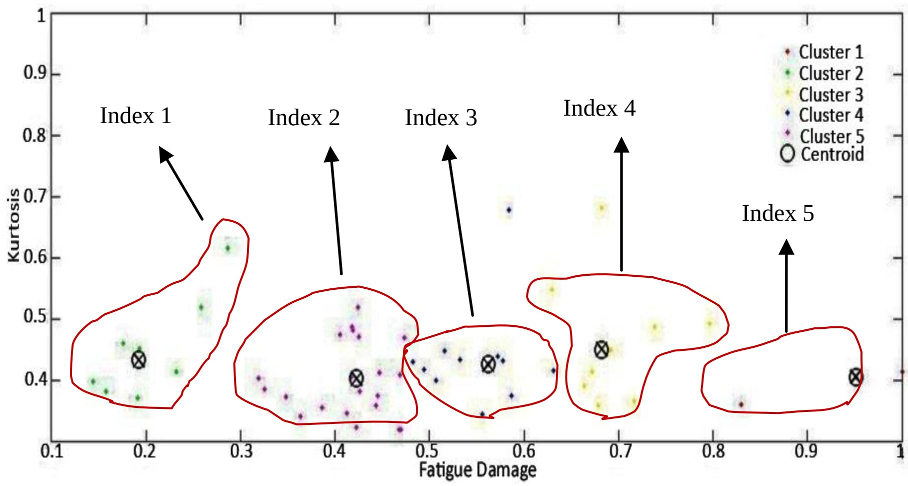


Figure 5. Selection of group to generate the index for D2 road

5. Conclusion

The fatigue failure index that generated using K-means clustering method is a new approach in fatigue study. Based on clustering process the data is cluster according to the required parameter and the best group can be determined. The objective function is calculated and the lower value of objective function show the K-means clustering is well organised data. From the results, it show that index 1 represent the low fatigue damage that occurs for the component. The increasing of the index number show more dominant of the effect of fatigue damage for the component. Index 4 for D1 road and index 5 for D2 road represent the high fatigue damage for the component. The high index number represent the high fatigue damage occurs for the system.

Acknowledgments
Authors would like to express their gratitude to Universiti Kebangsaan Malaysia (Research funding: UKM-KK-FRGS0188-2010 & UKM-OUP-NBT-28-135/2011) for supporting this research

6. References

[1] Stephens R I, Fatemi A, Stephens R R& Henry O F 2001 Metal Fatigue in Engineering. *John Wiley & Son, Inc: New York.*
[2] Fourlaris G, Ellwood R and Jones T B 2007 The reliability of the test results from simple test samples predicting the fatigue performance of automotive component *J. Material & Design***28(4)** 1198-1210
[3] Dong C L, Chang S H 2009 CAE (computer aided engineering) driven durability model verification for the automotive structure development*J. Finite Element in Analysis Design***45(5)** 324-332.
[4] Mao W, Rychlik I &Storhaug G 2008 Safety Index of Fatigue Failure for Ship Structure Details Department of Mathematical Science Chalmers University of Technology, University of Gothenburg.
[5] Bengston A And Rychlik I 2009 Uncertainty in fatigue life prediction of structures subject to Gaussian loads *J. Probabilistic Engineering Mechanics***24** 224-235
[6] Romaniv O N and Tkach A N 1987 A structural analysis of the kinetic fatigue failure curve of constructional steels *J. Materials Science***23(5)** 441-453
[7] Nuawi M Z, Abdullah S, Abdullah S, Arifin A and Haris S M 2009Matlab: a comprehensive reference for engineers, *Malaysia: McGraw-Hill Sdn. Bhd*
[8] Stephens R I, Dindinger P M and Gunger J E 1997 Fatigue damage editing for accelerated durability testing using strain range and SWT parameter criteria. *Int. J. of fatigue***19(8)** 599-606
[9] Oh C S 2001 Application of wavelet transform in fatigue history editing *Int J. of fatigue***23(3)** 241-250
[10] Abdullah S 2005 Wavelet Bump Extraction (WBE) for editing variable amplitude fatigue loadings *Ph.D, Thesis the University of Sheffield*
[11] Jeffrey D S 2007 Joint segmentation of multivariate astronomical time series: Bayesian sampling with a hierarchical model. *IEEE Trans on Signal Processing***55** 414-420
[12] Lopatka M, Laplanche C, Adam O, Motsch J F and Zarzycki J 2005 Non-stationary time-series segmentation based on the Schur prediction error analysis *IEEESP 13th Workshop on Statistical Signal Processing 2005*
[13] Lemire D 2007 A better alternative to piecewise linear time series segmentation Proc. of SIAM Data Mining 2007
[14] Keogh E, Selina C, David H and Pazzani M 2001 Segmenting time series: insight and recommendations *J. of IEEE Trans on Fuzzy Systems*
[15] Nopiah Z M, Abdullah S and Khairir M I 2008 Segmentation and scattering of fatigue time series data by kurtosis and root mean square *Proc. Int. Conf. on Signal Processing System* 64-68
[16] Alvarez F M, Troncoso A, Riquelme J C and Riquelme J M 2007 Partitioning-clustering technique applied to the electricity price time series *Proc. IDEAL'07 8th Int. Conf. on Intelligent Data Eng. And Automated Learning*

.