2005 American Control Conference
June 8-10, 2005. Portland, OR, USA

Online Detection of Fatigue Failure via Symbolic Time Series Analysis

Shalabh Gupta szg107@psu.edu
Asok Ray axr2@psu.edu
Eric Keller eek105@psu.edu
The Pennsylvania State University
University Park, PA 16802

***Abstract*— This paper examines the efficacy of symbolic time series analysis for online detection of fatigue failure in mechanical structures. The detection algorithm is formulated on the principles of *Symbolic Dynamics* and *Automata Theory*. The performance of this method is evaluated based on the information extracted from available sensor data for early detection of small anomalies in the observed data sequence. This concept is experimentally validated on a fatigue damage test apparatus. The time series data, generated from ultrasonic sensor and optical microscope, have been used for detection of small fatigue crack growth in ductile alloy 7075-T6 aluminium specimens.**

I. INTRODUCTION

Engineering applications require real time health monitoring for maintenance of quality performance, repair of damaged parts and and reliable operation. Material irregularities, variances in the usage patterns, overloads, sudden jerks and variations in temperature, humidity and other atmospheric conditions severely affect the service life of mechanical systems causing unexpected failures and undesirable events. Moreover the stochastic nature of fatigue phenomenon makes real time health monitoring an essential task [1] [2]. Online estimate of health fulfills the following objectives:

- Development of an insight into anomaly progression patterns at various time epochs;
- Reduction of the risk of sudden catastrophic failures and prevention of undesirable events by tracking system state in real time;
- Prevention of excessive retardation of machinery due to conservative design and overly high safety factors;
- Formation of an information basis for life extension and damage mitigation.

Generally, purely model-based approach for anomaly detection is infeasible because it is very difficult to achieve the requisite accuracy in modelling of complex dynamical systems over the desired range of operation. Consequently, the analysis of time series data, collected from available sensors, is essential for anomaly detection [3] [4]. Since fatigue crack growth in metallic alloys is a complex mechanism and has a random nature, analysis of time series data obtained in real time for early detection of small cracks and determination of system health by appropriate signal processing techniques is a challenging task.

This work has been supported in part by Army Research Office (ARO) under Grant No. DAAD19-01-1-0646.

Symbolic time series analysis requires conversion of continuous data to discrete symbols [5] [6] [7]. The symbol sequence is treated as a transform of the original time series data of measurements such that no significant information is lost [4]. Then, tools of Computational Mechanics are used to find patterns in the symbolic sequences through construction of a finite-state machine [8] [9]. This method of anomaly detection is built upon a fixed-structure, fixed-order Markov chain, and is called the *D-Markov machine* [6].

II. ANOMALY DETECTION PROBLEM FORMULATION

Anomaly detection in dynamical systems is formulated as a two-time-scale problem, in which it is assumed that the process has stationary dynamics at the *fast time scale* and any observable non-stationary behavior is associated with changes occurring on the *slow time scale*. From this perspective, the problem of anomaly detection is categorized into two subsets, 1) Forward problem and 2) Inverse problem.

A. Forward Problem

The primary objective of the forward problem is to identify the patterns followed by the system as anomaly develops slowly. Solution of the forward problem requires the following steps:

- Generation of time series data for self excitation as well as under various external stimuli
- Partitioning of the phase space (or wavelet space) for generation of symbolic sequences (on the fast time scale) at different epochs of the slow time scale
- Finite state machine construction from the symbol sequences and computation of the respective state probabilities
- Computation of anomaly measures as the norms of the differences of the respective state probability vectors from that under the nominal condition.

B. Inverse Problem

The inverse problem concentrates on inferring the anomalies based on the observed time series data and system response. This problem might be ill-posed or in other words there might be no unique solution. In this paper attempt is made to approach the forward problem.

This paper focuses on application and experimental validation of the anomaly detection method, where the source of possible anomalies is fatigue crack damage in mechanical structures. The specific objective is early detection of small cracks and fatigue damage occurring in ductile alloys before the onset of large crack propagation phase.

III. Description of the Experimental Apparatus

A. Test Bed

Fig.1 shows the test setup of the fatigue testing machine with all the required instrumentation [10]. The system is loaded by a hydraulic cylinder which is driven by hydraulic pressure controlled by electro hydraulic servo valve. The

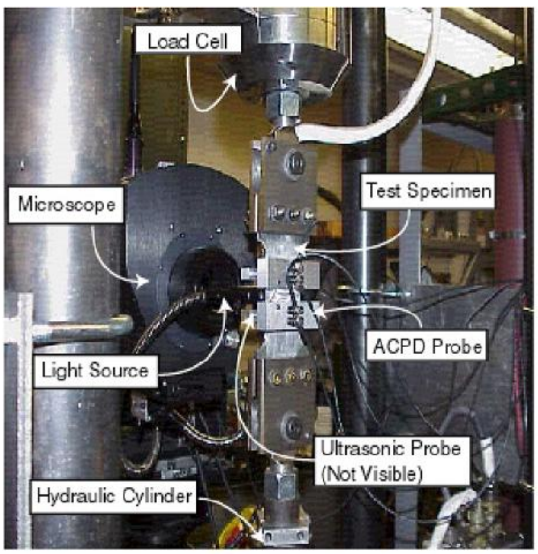

Fig. 1. *Fatigue Testing Machine Setup*

machine is equipped with a 10000-lbf load cell for performing desired load testing. The output of the load cell is fed to the analogue to digital (A/D) converter in the computer and directs the controller to generate a control signal, which governs the hydraulic servo valve for load control. The position of the cylinder is detected by a Linear Variable Differential Transducer(LVDT). The output of load cell and LVDT are sent to the main computer and are used for load and position limit control.

B. Specimen

Fig.2 shows the typical 7075-T6 aluminum specimen used for testing in the fatigue damage sensing machine. The central notch is made to increase the stress concentration factor in the middle and it guarantees crack propagation at notch ends. This specimen undergoes sinusoidal loading under tension-tension mode with a frequency of 12.5 Hz. Since inclusions and flaws are randomly distributed across the material small cracks form at these defects and propagate and join at the machined surface of the notch.

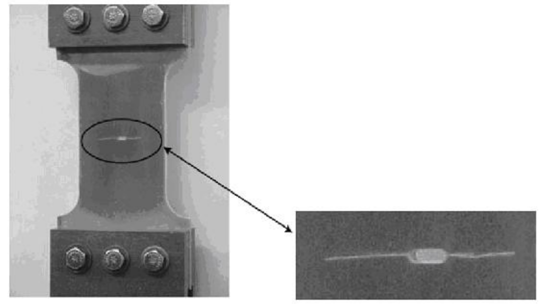

Fig. 2. *Cracked Specimen with a Center Notch*

C. Visual Data by Microscope

Two types of damage sensors are primarily used for data collection and fault detection. The first one is QUESTAR QM100 Step Zoom Long Distance travelling optical microscope which enables a direct measurement of the visible portion of the crack. Fig.3 shows a snapshot of the images taken by the microscope with a cracked specimen. The growth of crack is traced constantly by the images taken from the microscope at regular intervals. The microscope shifts from left to right side of central notch and vice versa after every 200 cycles to capture crack formation on either side of central notch. In order to take pictures the machine slows down to less than 5 Hz to get better images.

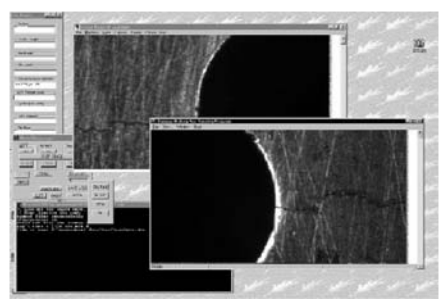

Fig. 3. *Microscope Images for Detection of Fatigue Cracks*

D. Ultrasonics Flaw Detection Technique

The second sensor used is the ultrasonics flaw detector which functions by emitting high frequency acoustic pulses which travel through the specimen and return back through the receiver transducers. A fault in the specimen changes the signature of the signal and thus it can capture some of the minute details and small changes during the crack initiation phase which are difficult to be detected by the microscope. Therefore advantages of using ultrasonic transducers are the ease of installation and detection of early anomalies before the onset of crack propagation phase. The ultrasonic transducers are placed on two sides of the specimen both above and below the central notch so as to send the signal through the region of crack propagation and receive it on the other side.

IV. Symbolic Dynamics and D-Markov Machine Construction

A data sequence (e.g., time series data) can be converted to a symbol sequence by partitioning the phase space Ω (over which the data evolves) into finitely many discrete blocks as shown in Fig.4. Let $\Phi = \{\Phi_1, \Phi_2, ...\Phi_n\}$ be a partitioning of phase space Ω, where Φ is a set of blocks covering the entire phase space such that it is exhaustive and mutually exclusive, i.e,

$$\bigcup_{j=1}^{n} \Phi_j = \Omega \tag{1}$$

and

$$\Phi_j \bigcap \Phi_k = \phi; \forall j \neq k \tag{2}$$

Each block $\Phi_j \epsilon \Phi$ is labelled as the symbol $\sigma_j \epsilon Z$, where the symbol set Z is called the *alphabet set* consisting of n different symbols. As the system evolves in time it travels through various blocks in its phase space and the corresponding symbol $\sigma_j \epsilon Z$ is assigned to it, thus converting a data sequence to a symbol sequence $\sigma_0 \sigma_1 \ldots \sigma_k \ldots$.

A. Wavelet Space Partitioning

Symbolic Sequence as a representation of the dynamics of the system can be obtained by some alternative methods because phase space partitioning might be a difficult task in case of high dimensions. The approach used in this paper is called the wavelet space partitioning method where time series data is first converted to the wavelet transform data [11] and thereafter the wavelet coefficients versus scale at time shift t_k are stacked after the ones at time shift t_{k+1} to obtain the so-called *scale series data*. Then, the range of wavelet *scale series* space is partitioned into equal horizontal slabs where each slab is assigned a particular symbol and thereby converting wavelet transform data to a symbolic sequence. The advantage of using wavelets over Fourier series is highlighted in the fact that it can scale up slight deformations capturing small behavioral changes.

B. Determining Machine States

A finite state machine is constructed from the symbol sequence, where the states of the machine are defined corresponding to a given *alphabet* set of size $\hat{A}$ and window length D where $\sigma_j \epsilon Z$ and $\hat{A}$ is the size of symbol set Z [6]. Alphabet size $\hat{A}$ determines the total number of partitions while window length D determines the fixed length of words forming the states of the machine. The states are chosen as words of length D from the symbol sequence, thereby making the total number of states to be equal to the total permutations of the alphabet symbols within word of length D i.e. $\hat{A}^D$. The number of partitions $\hat{A}$ and window length D determine the accuracy of small fault detection and their choice depends on specific experiments, noise level and available computation power. High *alphabet* set of size $\hat{A}$ can capture noise in signals while low $\hat{A}$ can miss minute details. Similarly, high D recognizes correlation patterns but would lead to larger number of states requiring more computation power. In this paper, a combination of $\hat{A}$=22 and D=2 was able to capture early anomalies. Once the parameters $\hat{A}$ and D are determined they are fixed for all time epochs. The states of machine are defined as $q_j \epsilon Q$ where $Q = \{q_1, q_2....q_N\}$ is the set of all states. Fig.4 shows the partitioning of phase space to obtain symbol sequence and the construction of finite state machine from the symbol sequence.

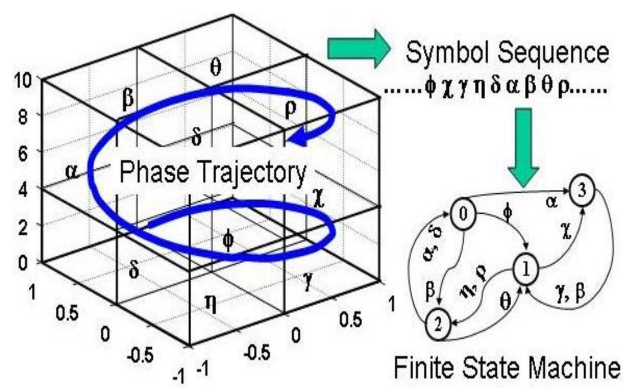

Fig. 4. *Continuous to Symbolic Dynamics and Finite State Machine*

For machine construction, the window of length D is shifted to the right by one symbol on the symbol sequence $\sigma_0 \ldots \sigma_i \ldots$, such that it retains the last (D-1) symbols of the previous state and appends it with the new symbol σ_j in the end. The symbolic permutation in the current window gives rise to a new state.

C. Example

As an example, let us choose D=2 and Z={0,1}, i.e., $\hat{A} = 2$. Consequently, the number of states are: $\hat{A}^D = 4$ and the states are Q={00, 01, 10, 11}. Fig.5 shows the construction of the finite state machine for the above example.

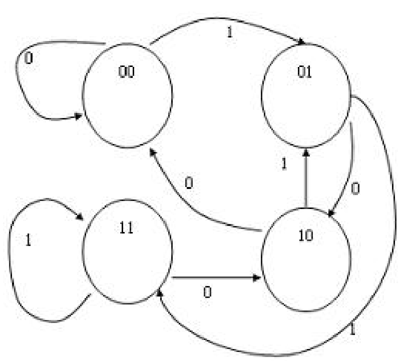

Fig. 5. *Finite State Machine with D=2 and Z={0,1}, i.e.,* $\hat{A} = 2$

D. D-Markov Property

The machine constructed in this fashion is called **D-Markov machine** because of its Markov properties. A symbolic stationary process is called D-Markov if the probability of the next symbol depends only on the previous D symbols i.e.

$$P(\sigma_i/\sigma_{i-1}\sigma_{i-2}\ldots\sigma_{i-D}\ldots) = P(\sigma_i/\sigma_{i-1}\sigma_{i-2}\ldots\sigma_{i-D}) \tag{3}$$

The finite state machine constructed above has D-Markov properties because the probability of occurrence of a symbol σ_i on a particular state depends only on the configuration of that state i.e. previous D symbols.

E. Calculation of Anomaly Measure

The probability of transitions from state q_j to state q_k is defined as

$$\pi_{jk} = P\left(\sigma \in Z \mid q_j \in Q, (\sigma, q_j) \to q_k\right); \sum_k \pi_{jk} = 1; \tag{4}$$

Thus, for a D-Markov machine, the stochastic matrix $\Pi \equiv [\pi_{ij}]$ describes all transition probabilities between states such that it has at most $\hat{A}^{D+1}$ nonzero entries. The left eigenvector $\mathbf{p}$ corresponding to the unit eigenvalue of Π is the state probability vector under the (fast time scale) stationary condition of the dynamical system [6].

On a given symbol sequence $\sigma_{i_1}\sigma_{i_2}...\sigma_{i_k}....$, a window of length (D) is slided by keeping a count of occurrences of word sequences $\sigma_{i_1}\cdots\sigma_{i_D}\sigma_{i_{D+1}}$ and $\sigma_{i_1}\cdots\sigma_{i_D}$ which are respectively denoted by $N(\sigma_{i_1}\cdots\sigma_{i_D}\sigma_{i_{D+1}})$ and $N(\sigma_{i_1}\cdots\sigma_{i_D})$. Note that if $N(\sigma_{i_1}\cdots\sigma_{i_D}) = 0$, then the state $q \equiv \sigma_{i_1}\cdots\sigma_{i_D} \in \mathbf{Q}$ has zero probability of occurrence. For $N(\sigma_{i_1}\cdots\sigma_{i_D}) \neq 0$, the transitions probabilities are then obtained by these frequency counts as follows:

$$\begin{aligned}\pi_{jk} \equiv P[q_k|q_j] = \frac{P[q_k,\ q_j]}{P[q_j]} &= \frac{P(\sigma_{i_1}\cdots\sigma_{i_D}\sigma)}{P(\sigma_{i_1}\cdots\sigma_{i_D})} \\ \Rightarrow \pi_{jk} &\approx \frac{N(\sigma_{i_1}\cdots\sigma_{i_D}\sigma)}{N(\sigma_{i_1}\cdots\sigma_{i_D})}\end{aligned} \tag{5}$$

where the corresponding states are denoted by $q_j \equiv \sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_D}$ and $q_k \equiv \sigma_{i_2}\cdots\sigma_{i_D}\sigma$.

The time series data under the nominal condition(set as a benchmark) generates the *State Transition Matrix* Π_{nom} that, in turn, is used to obtain the *State Probability Vector* $\mathbf{p}_{nom}$ whose elements are the visiting probabilities of all states, where $\mathbf{p}_{nom}$ is the left eigenvector of Π_{nom} corresponding to the (unique) unit eigenvalue. Subsequently, State Probability Vectors $\{\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_k \ldots\}$ are obtained at slow-time epochs $\{t_1, t_2, \ldots t_k \ldots\}$ based on the respective time series data. The length of data at each slow time epoch t_k must be sufficiently large so as to saturate the *State Probability Vector*. Machine structure and partitioning should be the same at all slow time epochs.

The *Anomaly Measures* at slow-time epochs $\{t_1, t_2, \ldots t_k \ldots\}$ are obtained as:

$$\mathcal{M}_k \equiv d\left(\mathbf{p}_k, \mathbf{p}_{nom}\right)$$

where the $d(\bullet, \bullet)$ is an appropriately defined distance function. In our analysis, the distance function d is chosen as the 2-norm such that the anomaly measure is given by:

$$\mathcal{M}_k \equiv \|\mathbf{p}_k - \mathbf{p}_{nom}\|_2 \tag{6}$$

The following steps summarize the procedure of anomaly detection in complex systems:

- Time Series Data collection and de-noising by appropriate de-noising tools
- Stacking of the wavelet transform coefficients of the time series data (obtained with suitable choice of mother wavelet) for various scales at different time shifts to obtain the so called *scale series data*.
- Appropriate Partitioning of the *scale series data* into $\hat{A}$ slabs to obtain the symbol sequence.
- Determination of the D-Markov Machine states from chosen alphabet size $\hat{A}$ and window length D based on desired sensitivity to detect changes and available computation power.
- Calculating Anomaly measures at different slow time epochs $\{t_1, t_2, \ldots t_k \ldots\}$ by the above mentioned procedure.

V. Online Experimental Validation

The tools of symbolic dynamics can be used for detection and characterization of small faults in mechanical structures. This section makes an assessment of the D-Markov machine method for early detection of fatigue damage, using the collected ultrasonic sensor data, as described in the experimental apparatus, Section III. The fatigue test that is considered here was performed at 12.5 Hz frequency under three different loading conditions: a) Low Cycle fatigue loading, b) High Cycle fatigue loading and c) Variable amplitude loading. There is no sharp demarcation between low cycle and high cycle fatigue. However it is considered that in low cycle fatigue the load is high enough and elasto-plastic behavior is observed in the ductile alloy 7075-T6. In case of high cycle fatigue it is mostly elastic behavior because of low load. Variable Amplitude loading was performed by block loading ranging between high cycle and low cycle loading conditions. For Low Cycle and High Cycle loading conditions, sinusoidal load with constant amplitude was applied. Variable amplitude loading consisted of a sequence of periodically repeating blocks of the same duration. Amplitude was constant inside each block and was varied from high load to low load over periodically repeating blocks.

For anomaly detection methodology the alphabet size was chosen as $\hat{A} = 22$ and window length was chosen as $D = 2$ while mother wavelet chosen was Haar Wavelet for all three cases. Increasing the value of D and $\hat{A}$ further up did not improve the results. The ultrasonics device was triggered at 5MHz during every load cycle at the peak of sinusoidal load to obtain 100 data points in each cycle. After every 1500 cycles the data chunk of N=10000 points(collected over 100 previous cycles) was chosen to generate the anomaly measure in real time at that slow time epoch with reference to the nominal condition. The nominal condition was chosen as a benchmark when the system obtained steady state conditions and ideal healthy behavior is assumed at this point. The anomaly measure at nominal condition was chosen to be zero. It is assumed that during fast time scale of 100 cycles (over which data is analyzed at each slow time epoch) system obeys stationarity conditions and no major changes occur in the dynamic behavior. The normalized results for anomaly measure and energy are presented here where anomaly measure was normalized by dividing by the maximum value of anomaly measure and energy was normalized by the nominal value of energy. It is to be emphasized that anomaly measure is a relative measure from the nominal condition of that particular data set and should not be confused with the actual damage. Any value of anomaly measure greater than zero just indicates some deviation from the nominal condition and it signifies that some small faults have occurred inside the specimen. However, inferring anomalies and setting a threshold on anomaly measure to take the control step for life extension is an inverse problem and it is considered for future work.

VI. Results and Discussion

This section presents the results obtained by analyzing the ultrasonics data with the D-Markov anomaly detection methodology. Fig. 6, 7 and 8 correspond to three loading conditions: 1) Low Cycle, 2) High Cycle and 3) Variable Amplitude Loading respectively and Fig. 9, 10 and 11 show the corresponding energy plots of the data where energy is defined as: $E = \sum_{j=1}^{N} s(j)^2$; where s(j) is the amplitude of data points and N=10000 is the length of data chunk chosen at one slow time epoch for calculation of anomaly measure and energy. The vertical lines in Fig. 6, 7 and 8 correspond to the point of crack detection by microscope, however, it was noticed that the specimen underwent small changes even before the onset of crack propagation phase. The energy plots Fig. 9, 10 and 11 show similar trends, however energy is just a scalar measure and cannot capture small distortions in signal behavior. Anomaly Measure based on symbolic dynamics is based on statistical analysis of data sets and is capable of detecting small changes in the signal profile. Three different cases are presented here: 1) *Low Cycle fatigue*(Fig. 6 and 9)-Microscope detected crack at $5.5 * 10^4$ cycles but under low cycle fatigue the system undergoes a lot of elasto-plastic deformations causing a large number of dislocations and growth of small cracks at an early stage. Therefore, energy showed a significant decrease while anomaly measure further magnified these changes and clearly indicated anomalies at an early stage; 2) *High Cycle fatigue*(Fig. 7 and 10)- Microscope detected crack at $2.23 * 10^5$ cycles. Anomaly measure detected small changes in the beginning and a rapid progress was observed at the beginning of crack propagation phase at around $2.25 * 10^5$ cycles. Energy did not show significant changes in the beginning but detected the bifurcation point at which rapid crack propagation started. 3)*Variable Amplitude Loading*(Fig. 8 and 11)- Microscope detected crack at $1.12 * 10^5$ cycles but anomaly measure showed significant changes starting from nominal to $1.12 * 10^5$ cycles; after that rapid growth started because of crack propagation. Energy could not efficiently capture small faults but detected the crack propagation phase.

VII. Conclusions and Future Work

A. Conclusions

This paper presents the application of a novel concept of anomaly detection in complex dynamical systems. This anomaly detection algorithm relies on symbolic time series analysis of process variable(s) and is built upon the principles of *Symbolic Dynamics and Automata Theory*. The efficacy of this approach was verified on the ultrasonic sensor data generated from fatigue damage test apparatus for early detection of fatigue failure in ductile alloy 7075-T6 aluminum specimens.

B. Future Work

- Development of Damage mitigation controller on the basis of inferred anomalies from sensor data by time series analysis in real time.
- Development of suitable partitioning techniques for maximum information extraction from sensor data.

References

[1] K. Sobczyk and B.F. Spencer, *Random Fatigue: Data to Theory*, Academic Press, Boston, MA, 1992.
[2] M. Klesnil and P. Lukas, *Fatigue of Metallic Materials*, Material Science Monographs 71, Elsevier, 1992.
[3] H.D.I. Abarbanel, *The Analysis of Observed Chaotic Data*, Springer-Verlag, New York; 1996.
[4] C.S. Daw and C.E.A. Finney and E.R. Tracy, A review of symbolic analysis of experimental data, *Review of Scientific Instruments*, vol. 74,2,2003, pp 915-930.
[5] R. Badii and A. Politi, *Complexity hierarchical structures and scaling in physics*, Cambridge University Press, United Kingdom; 1997.
[6] A. Ray, Symbolic Dynamic Analysis of Complex Systems for Anomaly Detection, *Signal Processing*, vol 84, 7, 2004, pp 1115-1130.
[7] D. Lind and M. Marcus, *A Introduction to Symbolic Dynamics and Coding*, Cambridge University Press, United Kingdom, 1995.
[8] J.P. Crutchfield and K. Young, Inferring Statistical Complexity, *Physical Review Letters*, vol 63, 1989, pp 105-108.
[9] J.P. Crutchfield, The Calculi of Emergence: Computation dynamics and induction, *Physica*, vol D, 75, 1994, PP 11-54.
[10] E.E. Keller, *Real time sensing of fatigue crack damage for information-based decision and control*, PhD. Thesis, Department of Mechanical Engineering, Pennsylvania State University, State College, PA, 2001.
[11] Martin, Vetterli and K. Jelena, *Wavelets and Subband Coding*, Prentice Hall PTR, 1995.

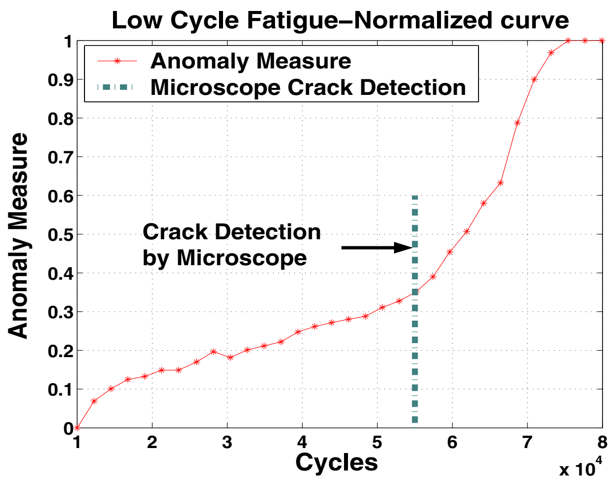

Fig. 6. *Anomaly measure vs cycles in case of LowCycle Loading. The vertical line shows first crack detection by microscope at* $5.5 * 10^4$ *cycles.*

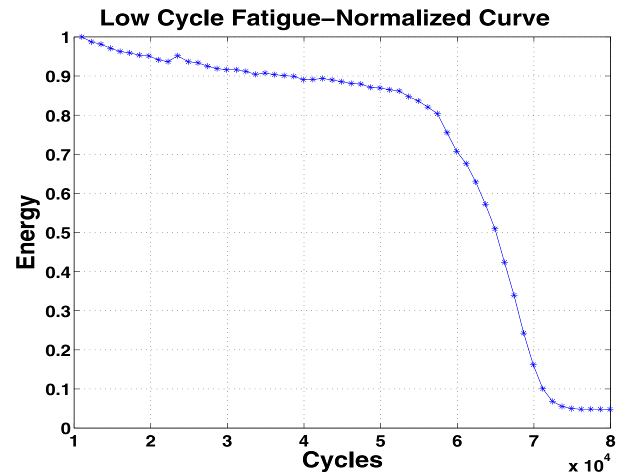

Fig. 9. *Energy of the ultrasonics data vs cycles for LowCycle Loading.*

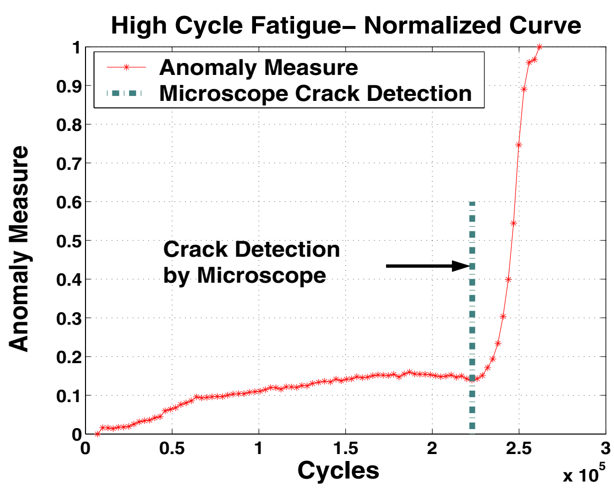

Fig. 7. *Anomaly measure vs cycles in case of HighCycle Loading. The vertical line shows first crack detection by microscope at* $2.23 * 10^5$ *cycles.*

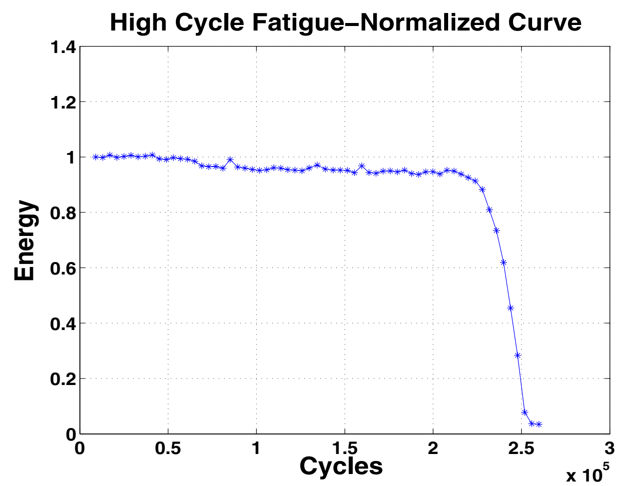

Fig. 10. *Energy of the ultrasonics data vs cycles for HighCycle Loading.*

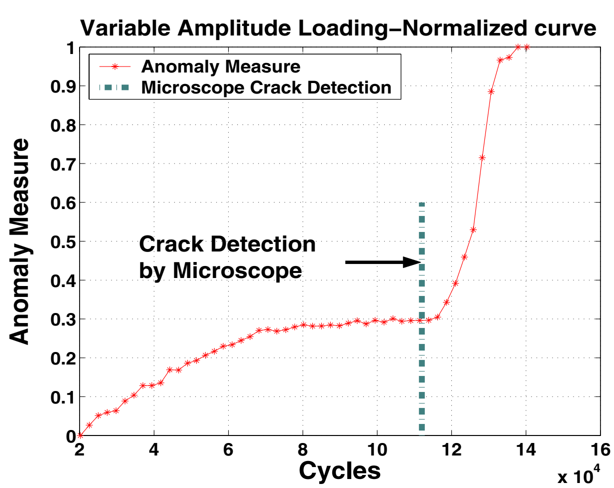

Fig. 8. *Anomaly measure vs cycles in case of Variable Amplitude Loading. The vertical line shows first crack detection by microscope at* $1.12 * 10^5$ *cycles.*

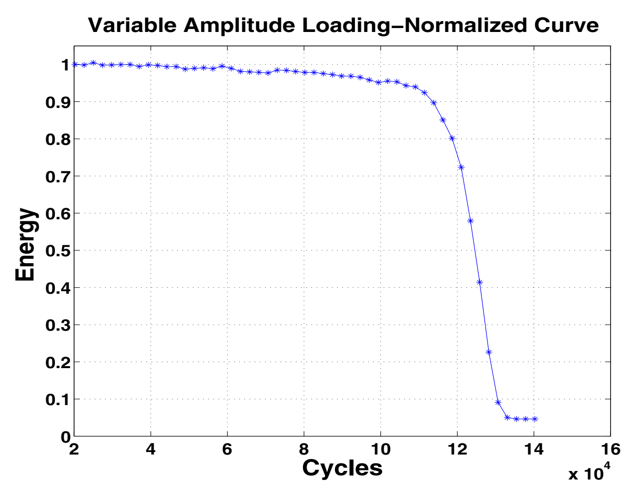

Fig. 11. *Energy of the ultrasonics data vs cycles for Variable Amplitude Loading.*