Real-time fatigue life estimation in mechanical structures*

Shalabh Gupta and Asok Ray

Mechanical Engineering Department, The Pennsylvania State University, University Park, PA 16802, USA

E-mail: szg107@psu.edu and axr2@psu.edu

Abstract
The paper addresses the issue of online diagnosis and prognosis of emerging faults in human-engineered complex systems. Specifically, the paper reports a dynamic data-driven analytical tool for early detection of incipient faults and real-time estimation of remaining useful fatigue life in polycrystalline alloys. The algorithms for fatigue life estimation rely on time series data analysis of ultrasonic signals and are built upon the principles of *symbolic dynamics*, *information theory* and *statistical pattern recognition*. The proposed method is experimentally validated by using 7075-T6 aluminium alloy specimens on a special-purpose fatigue test apparatus that is equipped with ultrasonic flaw detectors and an optical travelling microscope. The real-time information, derived by the proposed method, is useful for mitigation of widespread fatigue damage and is potentially applicable to life extending and resilient control of mechanical structures.

Keywords: symbolic time series analysis, anomaly detection, fatigue damage, ultrasonic sensing

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Recent developments and maturity of scientific theories have demonstrated that information-based detection, diagnosis and prognosis of failures is essential for online estimation of structural damage and quantification of structural integrity for maintenance of safe and reliable operation of human-engineered complex systems. This paper addresses the issues of early diagnosis and prognosis for advanced warning of incipient and gradually emerging faults in complex dynamical systems. Fatigue damage is one of the most commonly encountered sources of structural degradation and could potentially cause catastrophic failures in mechanical systems [1].

    The current state-of-the-art in fatigue life estimation is largely based on model-based analysis; however, no existing fatigue damage model, solely based on the fundamental principles of physics, can adequately capture the dynamical behaviour of fatigue damage at the grain level [2]. In general, model-based approaches are critically dependent on initial defects in the material microstructure, which may randomly form crack nucleation sites and are difficult to model [1, 3]. Small deviations in initial conditions and critical parameters may produce large bifurcations in the expected dynamical behaviour of fatigue damage [4]. In addition, fluctuations in usage patterns (e.g., random overloads) and environmental conditions (e.g., temperature and humidity) may adversely affect the service life of mechanical systems. From this perspective, a data-driven statistical method is presented for online estimation of remaining useful fatigue life in mechanical structures as augmentation of the authors' earlier work [5] that addressed early detection of incipient fatigue damage in polycrystalline alloys.

    The random distribution of micro-structural flaws in identically manufactured components can produce a wide uncertainty in the crack initiation phase [1]. For example, inclusions, casting defects and machining marks possibly originated during fabrication may cause stress augmentation at certain locations. These unavoidable surface and sub-surface

defects constitute integral parts of the material microstructure of the operating machinery. As such, evolution of fatigue damage is described as a stochastic phenomenon [1, 3] and a stochastic measure of fatigue crack growth has also been proposed in the literature [6, 7]. This stochastic nature of fatigue damage emphasizes the need for online updating of information using sensing devices that are sensitive to small microstructural changes and are capable of issuing early warnings during fatigue damage evolution [8]. As such, the above discussions evince that time series analysis of sensor data is essential for real-time monitoring of fatigue in mechanical structures [9].

Impedance of the ultrasonic signals has been shown to be very sensitive to small microstructural changes occurring during the early stages of fatigue damage [10–13]. As such, ultrasonic sensing has been adopted to monitor the evolution of fatigue damage in ductile aluminium alloy 7075-T6 test specimens. The paper presents the application of the symbolic time series analysis (STSA) method [14, 15] to the ultrasonic signals for real-time detection of fatigue damage and estimation of the remaining useful fatigue life. STSA for anomaly detection is an information-theoretic pattern identification tool that is built upon a fixed-structure, fixed-order Markov chain [15]. Recent literature [16, 17] has reported experimental validation of STSA-based pattern identification by comparison with other existing techniques such as principal component analysis (PCA) and artificial neural networks (ANN); STSA has been shown to yield superior performance in terms of early detection of anomalies, robustness to noise, and real-time execution in different applications such as electronic circuits, mechanical vibration systems and fatigue damage in polycrystalline alloys.

This paper addresses the issues of fatigue damage monitoring and remaining life estimation as the following two related problems:

(1) The forward (analysis) problem of anomaly detection and assimilation of the respective statistical information.
(2) The inverse (synthesis) problem of anomaly identification (e.g., determination of ranges of anomalous parameters).

As a partial solution to the forward problem, recent literature has reported successful application of the STSA method [15] to early detection of fatigue damage [5, 17]. The procedure is experimentally validated by using 7075-T6 aluminium alloy specimens on a special-purpose fatigue test apparatus that is equipped with ultrasonic flaw detectors and an optical travelling microscope [18]. This paper addresses both the forward and inverse problems as augmentation of the reported work on part of the forward problem (i.e., early detection of fatigue damage [5]). The underlying statistical concepts and the theory of fatigue life estimation are experimentally validated.

The paper is organized in six sections, including the present section. Section 2 reviews the underlying concepts and essential features of symbolic time series analysis for anomaly detection [15, 19]. Section 3 formulates the anomaly detection problem. Section 4 describes the experimental apparatus on which the anomaly detection method is validated in real time. Section 5 presents the pertinent results of both the forward problem and the inverse problem. The paper is concluded in section 6 along with recommendations for future research.

2. Review of STSA-based anomaly detection

This section presents the underlying concepts and essential features of STSA [14] for anomaly detection in complex dynamical systems [15]. While the details are reported in previous publications [15, 19], the key features of STSA are briefly summarized here for clarity and completeness of this paper.

In the STSA procedure, a data sequence is converted to a symbol sequence by partitioning a compact region of the phase space of the dynamical system, over which the trajectory evolves, into finitely many discrete blocks. Each block is labelled as a symbol, where the symbol set Σ is called the *alphabet* that consists of $|\Sigma|$ different symbols. (Note: $|\Sigma| \geqslant 2$.) As the system evolves in time, it travels through or touches various blocks in its phase space and the corresponding symbol $\sigma \in \Sigma$ is assigned to it, thus converting the data sequence into a symbol sequence.

2.1. The two-time scale modelling

The sampling frequency and data acquisition process is required to be faster than the time period of damage evolution. Therefore, fatigue damage monitoring is formulated as a two-time-scale problem: the *fast time scale* is related to the response time of machinery operation. Over the span of a given time series data sequence, the structural dynamic behaviour of the system is assumed to remain invariant, i.e., the process has stationary dynamics at the fast time scale. In other words, the variations in the internal dynamics of the system are assumed to be negligible on the fast time scale. The *slow time scale* is related to the time span over which the process may exhibit non-stationary dynamics. An observable non-stationary behaviour can be associated with anomalies evolving at the slow time scale. In general, a long time span in the fast time scale is a tiny (i.e., several orders of magnitude smaller) interval in the slow time scale. For example, evolution of fatigue damage in structural materials (causing a detectable change in the dynamics of the system) occurs on the slow time scale (possibly in the order of months or years); the fatigue damage behaviour is essentially invariant on the fast time scale (approximately in the order of seconds or minutes). Nevertheless, the notion of fast and slow time scales is dependent on the specific application, loading conditions and operating environment. As such, with the perspective of fatigue monitoring, sensor data acquisition is done on the fast time scale at different slow time epochs separated by regular intervals. Thereby, anomaly detection is done over the slow time scale from the information generated by the STSA of the data collected at different slow time epochs. A pictorial view of the two time scales is presented in figure 1.

The key concepts of STSA-based anomaly detection procedure are summarized below.

(1) *Time series data acquisition from appropriate sensor(s)* at time epoch t_0, i.e., the nominal condition, at which the system is assumed to be in the healthy state.
(2) *Generation of wavelet transform coefficients, obtained with an appropriate choice of the wavelet basis* [20]. A crucial step in symbolic time series analysis is partitioning of the phase space for symbol sequence generation [14].
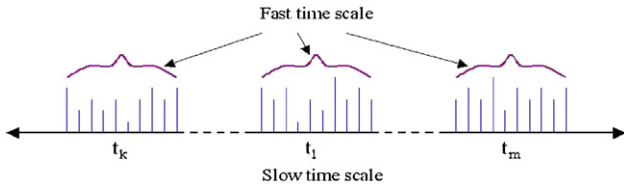
Figure 1. Pictorial view of the two time scales: (i) *slow time scale* of anomaly evolution and (ii) *fast time scale* for data acquisition and signal conditioning.

Figure 2. Measurement of the physical process by a set of observable parameters $\mathbf{o} = \{o^1, o^2, \ldots, o^r\}$.

Since the structure of phase space can become very complex for high dimensional systems or may even be unknown for un-modelled systems, an alternative way is to extract information from the time series data of appropriate sensors. Once the time series data set is collected, it is transformed into the wavelet domain by appropriate choice of scale(s) and basis [19]. The paper has adopted a wavelet-based partitioning approach [15] for the construction of symbol sequences from the time series data. The wavelet transform [20] largely alleviates the difficulties of phase-space partitioning in the case of high dimensions and is particularly effective with noisy data from high-dimensional dynamical systems [19].

(3) *Maximum entropy partitioning of the wavelet space at the nominal condition* [19]. The partitioning is done such that the regions with more information are partitioned finer and those with sparse information are partitioned coarser. This is achieved by maximizing the Shannon entropy, which is defined as:

$$S = -\sum_{i=1}^{|\Sigma|} p_i \log(p_i)$$

where p_i is the probability of the ith segment of the partition and summation is taken over all segments. The partitioning is fixed for subsequent time epochs. Each segment of the partitioning is assigned a particular symbol and a symbol sequence is generated.

(4) *Calculation of the probability vector* $\mathbf{p}^0$ at time epoch t_0 whose elements represent the visiting probabilities of the segments of the partitioning. As a consequence of maximum entropy, $\mathbf{p}^0$ has uniform distribution (i.e., all elements of $\mathbf{p}^0$ are equal).

(5) *Time series data acquisition at subsequent slow time epochs,* $t_1, t_2, \ldots, t_k, \ldots$, and their conversion to the wavelet domain to generate respective symbolic sequences based on the partitioning at time epoch t_0.

(6) *Generation of the probability vectors* $\mathbf{p}^1, \mathbf{p}^2, \ldots, \mathbf{p}^k, \ldots$ at slow time epochs, $t_1, t_2, \ldots, t_k, \ldots$ from the respective symbolic sequences

(7) Computation of scalar anomaly measures $\psi^1, \psi^2, \ldots, \psi^k, \ldots$ at time epochs, $t_1, t_2, \ldots, t_k, \ldots$ based on evolution of these probability vectors and by defining an appropriate distance function with respect to the nominal condition (see [5]).

3. Anomaly detection problem formulation

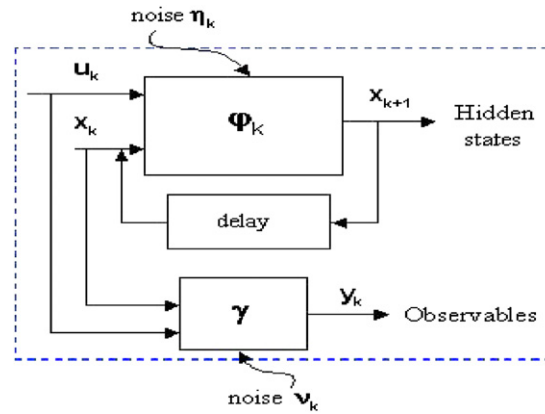Having discussed the tools and techniques of STSA, this section outlines the formulation of the anomaly detection problem (including fault diagnosis and prognosis). In diverse applications of science and engineering, high dimensionality of the phase space, uncertain chaotic behaviour [4], nonlinear spatial-temporal stochastic processes and random noisy excitations often restrict the applications of the fundamental laws of physics to accurately determine a dynamical model for such complex emergent systems because of the infeasibility in achieving requisite modelling accuracy and precision [21]. As such, the problem is simplified using observation-based estimation of the underlying mathematical structure of the system and its relevant parameters. As an illustration, a map of the dynamical process (in discrete time) is described as:

$$\mathbf{x}_{k+1} = \varphi_k(\mathbf{x}_k, \mathbf{u}_k, \eta_k) \tag{1}$$

$$\mathbf{y}_k = \gamma(\mathbf{x}_k, \mathbf{u}_k) + \upsilon_k \tag{2}$$

where k is the discrete time index; φ describes the time evolution of the state trajectory; γ represents the measurement model; $\mathbf{x}$ is the state vector in the phase space; $\mathbf{u}$ is the input excitation vector; $\mathbf{y}$ is the measurement vector; η is the (possibly non-additive) process noise and υ is the measurement noise. Evolution of the dynamical process generates time series data of system outputs $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_k, \ldots$ starting from an initial point $\mathbf{x}_0$. Since $\mathbf{x}$ is usually hidden and φ is generally unknown especially for anomalous systems, the problem needs to be investigated by alternative means of extraction of relevant information from the time series data set $\mathbb{Y} = \{\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_k, \ldots\}$ of selected observable outputs (e.g., sensor data), as shown in figure 2.

As stated earlier in section 1, the issues of fatigue damage monitoring and remaining life estimation are addressed as two problems: (1) the *forward (or analysis) problem* and (2) the *inverse (or synthesis) problem. The forward problem* consists of prediction of outcomes, given *a priori* knowledge of the underlying model parameters. In the absence of an existing model this problem requires generation of behavioural patterns of the system evolution through an off-line analysis of an ensemble of the observed time series data. On the other hand, *the inverse problem* consists of estimation of critical parameters characterizing the system under investigation using the actual observations. Inverse problems arise in different engineering disciplines such as geophysics, structural health monitoring, weather forecasting and astronomy. Inverse

problems often become ill-posed and challenging due to the following reasons: (a) high dimensionality of the parameter space under investigation and (b) in the absence of a unique solution where change in multiple parameters can lead to the same observations.

In the presence of sources of uncertainties, any parameter inference strategy requires estimation of parameter values and also the associated confidence intervals, or the error bounds, to the estimated values. As such, inverse problems are usually solved using the Bayesian methods that allow observation-based inference of parameters and provide a probabilistic description of the uncertainty of inferred quantities. A good discussion of inverse problems is presented by Tarantola [22].

In the context of fatigue damage, the tasks and solution steps of these two problems as followed in this paper are discussed below.

(1) *Forward problem:* the primary objective of the forward problem is identification of the changes in behavioural patterns of the system dynamics due to evolving fatigue damage on the slow time scale. Specifically, the forward problem aims at detecting the deviations in the statistical patterns in the time series data, generated at different time epochs in the slow time scale, from the nominal behaviour pattern. The solution procedure of the forward problem requires the following steps:

(F1) Collection of time series data sets (at fast time scale) from the available sensor(s) at different slow time epochs;

(F2) Analysis of these data sets using the STSA method as discussed in section 2 to generate pattern vectors defined by the probability distributions at the corresponding slow time epochs. The profile of anomaly measure is then obtained from the evolution of this pattern vector from the nominal healthy condition [5];

(F3) Generation of a family of such profiles from multiple experiments performed under identical conditions to construct a statistical pattern of anomaly growth. Such a family represents the uncertainty in the evolution of anomalies in mechanical systems due to its stochastic nature. This step is required in systems where there is a source of parametric or non-parametric uncertainty. In the case of fatigue damage, the uncertainty arises from the random distribution of microstructural flaws in the body of the component leading to stochastic behaviour [1].

(2) *Inverse problem:* the objective of the inverse problem is to infer the anomalies and to provide estimates of the remaining useful fatigue life from the observed time series data and system response in real time. The decisions are based on the information derived in the forward problem. Specifically, in the context of fatigue damage, identical structures operated under identical loading and environmental conditions show different trends in the evolution of fatigue due to surface and sub-surface material uncertainties. Therefore, as a precursor to the solution of the inverse problem, generation of an ensemble of data sets is required during the forward problem for multiple fatigue tests conducted under identical operating conditions. Damage estimates can be obtained at any particular instant in a real-time experiment with certain confidence intervals using the information derived from the ensemble of data sets of fatigue damage evolution generated in the forward problem [15]. Furthermore, a control policy can be designed based on these estimates for life extension and prevention of catastrophic failures [23]. The solution procedure of the inverse problem requires the following steps:

(I1) Collection of time series data sets (in the fast time scale) from the available sensor(s) at different slow time epochs up to the current time epoch in a real-time experiment as in step F1 of the forward problem;

(I2) Analysis of these data sets using the STSA method as discussed in section 2 to generate pattern vectors defined by probability distributions at the corresponding slow time epochs. The value of anomaly measure at the current time epoch is then calculated from the evolution of this pattern vector from the nominal healthy condition [5]. The procedure is similar to the step F2 of the forward problem. As such, the information available at any particular instant in a real-time experiment is the value of the anomaly measure calculated at that particular instant;

(I3) Detection, identification and estimation of an anomaly (if any) based on the computed anomaly measure and the statistical information derived in step F3 of the forward problem. More details of this step are discussed in section 5.

A schematic of the entire framework for the anomaly detection problem in mechanical systems is shown in figure 3. As shown in figure 3, the forward problem section involves the generation of ultrasonic data sets from fatigue experiments which are analysed using the STSA method to produce a profile of the anomaly measure (see section 2) that represents the evolution of fatigue damage. Following the same procedure, several experiments are conducted under identical conditions to generate a family of anomaly measure profiles. Such a family represents the stochastic behaviour of the fatigue damage evolution on a slow time scale (see section 5.1).

This family of anomaly measure profiles is analysed in the inverse problem section to generate the requisite statistical information (see section 5.2). The information available in real time is the value of the anomaly measure obtained from the analysis of ultrasonic data at any particular time epoch. This information is entered in the inverse problem section that provides the estimates of the *expended life fraction*. The estimates can only be obtained within certain bounds at a particular confidence level. The online statistical information on the damage status is significant because it can facilitate early scheduling for the maintenance or repair of critical components or to prepare an advance itinerary of the damaged parts. The information can also be used to design control policies for damage mitigation and life extension. The entire architecture has been built within the MATLAB and C++ environment. The proposed architecture in figure 3 is also applicable to other engineering fields for anomaly detection.
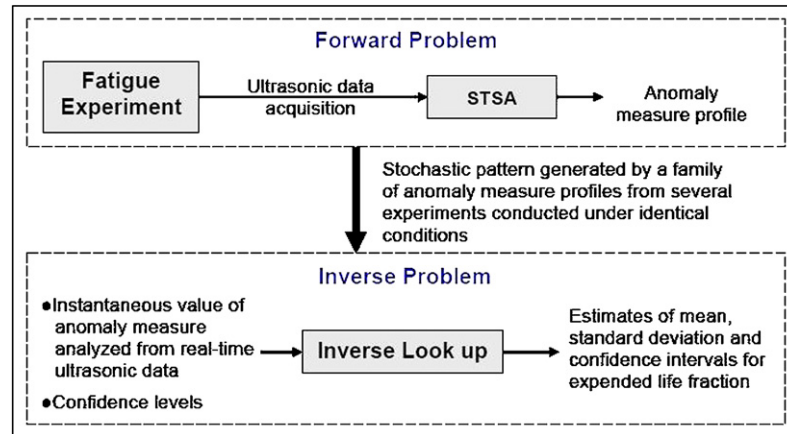
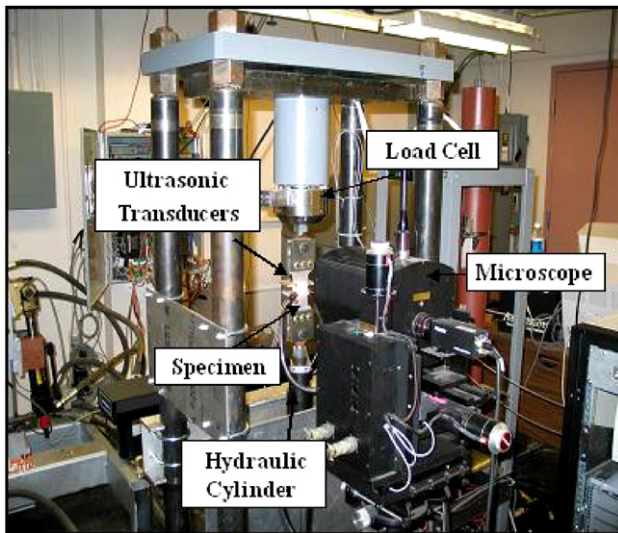**Figure 3.** Framework of STSA based anomaly detection in mechanical systems.

Figure 4. Computer-instrumented apparatus for fatigue testing.

Figure 5. Cracked specimen with a side notch.

Figure 6. Schematic of ultrasonic sensors on a test specimen.

4. Description of the experimental apparatus

The experimental apparatus, shown in figure 4, is a special-purpose uniaxial fatigue testing machine, which is operated under load control or strain control at speeds up to 12.5 Hz; a detailed description of the apparatus and its design specifications are reported in [18]. The test specimens are subjected to tensile–tensile cyclic loading by a hydraulic cylinder under the regulation of computer-controlled electro-hydraulic servo-valves. The feedback signals that are generated from the load cell and the extensometer are processed by signal conditioners that include standard amplifiers and signal processing units. The controller governs the hydraulic servo-valve for operation under specified load and position limits. The damage estimation and life prediction subsystem consists of data analysis algorithms and the associated computer hardware.

Figure 5 shows a single-edge notched specimen of 7075-T6 aluminium alloy used for testing in the fatigue damage test apparatus. The spec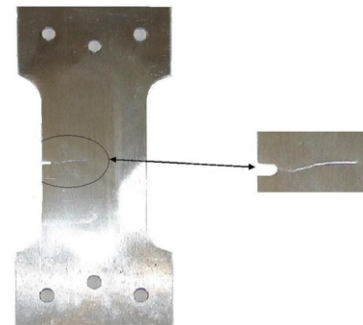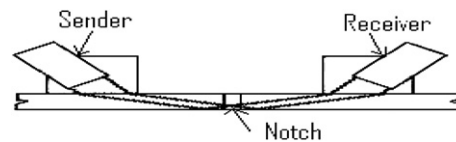imens are 3 mm thick and 50 mm wide with a slot on one side of 1.58 mm diameter and 4.57 mm length. The notch is made to increase the stress concentration factor that ensures crack initiation and propagation at the notch end. The travelling optical microscope, shown as a part of the test apparatus in figure 4, provides direct measurements of the visible portion of a crack. The test specimens have been subjected to sinusoidal loading under tension–tension mode (i.e., with a constant positive offset) at a frequency of 12.5 Hz. The dc offset is provided in the load cycling to ensure that the specimen is always under tension.

A piezoelectric transducer is used to inject ultrasonic waves in the specimen and a receiver transducer is placed on the other side of the notch to measure the transmitted signal, as seen in figure 6. The ultrasonic signals produced are 10 MHz sinusoidal waves and they are triggered during a very short portion at the peak of every load cycle. Since material characteristics (e.g., voids, dislocations and short cracks) influence ultrasonic impedance, a small fault in the specimen is likely to change the signature of the signal at

the receiver end. Therefore, the signal can be analysed using STSA to capture minute changes during the early stages of fatigue damage [17]. A significant amount of damage (e.g., dislocations, short cracks and microstructural defects) occurs before the crack appears on the surface of the specimen when it is observed by the microscope [24]. The resulting damage can cause detectable attenuation and/or distortion of the ultrasonic waves [10, 17]. The crack propagation stage starts when this microstructural damage eventually develops into a single large crack. Subsequently, the crack growth rate increases rapidly and when the crack becomes sufficiently large, complete attenuation of the ultrasonic signal occurs at the receiver end. An elaborate description of the properties of ultrasonic waves in solid media is provided by Rose [25].

5. Experimental procedure, application of STSA and results and discussion

This section provides a detailed description of experimental procedure, application of the STSA method to the time series data of ultrasonic signals and the results and discussion for both the forward and the inverse problems. The solution procedures for these problems are provided below.

5.1. Solution procedure for the forward problem and results

This section presents a detailed description of the solution procedure for the forward problem. As discussed earlier, the primary objective of the forward problem is to identify the behavioural pattern of damage evolution in a complex dynamical system involving the uncertainties (if any) which can be both parametric or non-parametric in nature.

5.1.1. Sources of uncertainties.
In the case of fatigue damage, as considered in this paper, the sources of uncertainties are assumed to include:

(a) material inhomogeneities such as voids or inclusions,
(b) surface defects including finishing marks that usually develop from the machining process, nonuniform polishing and other deformities,
(c) sub-surface defects originating due to random distribution of microstructural flaws like dislocations and grain boundaries,
(d) variations in the critical dimensions of the components resulting from the non-zero tolerances of the cutting tools used in the fabrication process,
(e) small fluctuations in the environmental conditions such as humidity and temperature,
(f) small fluctuations in the operating conditions due to noisy environment and finite precision of the mechanical system.

In the presence of the above uncertainties, a complete solution of the anomaly detection problem cannot be obtained in the deterministic setting because the profile of anomaly progression would not be identical for similarly manufactured components. In that case, the problem can be represented in the stochastic setting, where a family of anomaly progression profiles is generated from multiple experiments conducted under identical conditions [26]. As such, the requirement

of the forward problem is to generate a pattern that consists of a family of anomaly progression profiles. Each member of this family represents the anomaly measure profile of a particular sample. This profile is generated from a fatigue test that is conducted to observe the entire service life of the specimen from the starting healthy condition to the eventual failure.

5.1.2. Experimentation and data analysis procedure using STSA.
The fatigue tests have been conducted on 7075-T6 aluminium specimens at 12.5 Hz frequency. The specimens are subjected to a sinusoidal load cycling where the maximum and minimum loads are 89.3 MPa and 4.85 MPa at the nominal condition. Ultrasonic waves with a frequency of 10 MHz are triggered at the peak of each sinusoidal load cycle where the stress is maximum and the crack is open causing maximum attenuation of the ultrasonic waves. Since the ultrasonic frequency is much higher than the load cycling frequency, data collection is performed for a very short interval in the time scale of load cycling. The slow time epochs have been chosen to be 1000 load cycles (i.e., ~80 s) apart. At the onset of each slow time epoch, the ultrasonic data points are collected on the fast time scale of 50 cycles (i.e., ~4 s), which produced a string of $N = 15\,000$ data points. It is assumed that during the fast time scale of 50 cycles, the system remains in a stationary condition and no major changes occur in the fatigue damage behaviour. These sets of time series data points collected at different slow time epochs are analysed using the STSA method, as described in section 2, to calculate the anomaly measures at those slow time epochs.

The nominal condition at the slow time epoch t_0 is chosen to be ~0.5 kilocycles to ensure that the electro-hydraulic system of the test apparatus had come to a steady state and it is assumed that no significant damage occurred till that point. This nominal condition is chosen as a benchmark where the anomaly measure is chosen to be zero. The anomalies at subsequent slow time epochs, $t_1, t_2, \ldots, t_k, \ldots$, are then calculated using STSA to yield a profile of anomaly measure representing the progression of fatigue damage on the slow time scale. It is emphasized that the anomaly measure is relative to the nominal condition which is fixed in advance and should not be confused with the actual damage at an absolute level. The data collection is stopped at a time epoch t_f considered as the final failure point where the ultrasonic energy is attenuated to 2% of the nominal condition. The energy of the signal is defined as:

$$E = \sum_{i=1}^{N} |s(i)|^2|$$

where $|s(i)|$ is the magnitude of the ith data point of the ultrasonic signal. Once the failure point is reached the specimen is already under crack propagation stage and a sufficiently large crack has developed such that it is no longer useful and is considered broken. Following the above procedure, a family of profiles is generated for multiple experiments conducted under identical experimental conditions.

For the STSA procedure, the alphabet size for partitioning has been chosen to be $|\Sigma| = 8$ while the wavelet basis is chosen to be 'gaus2' [27]. Increasing the alphabet size $|\Sigma|$ does not improve the results and creates a large number of states
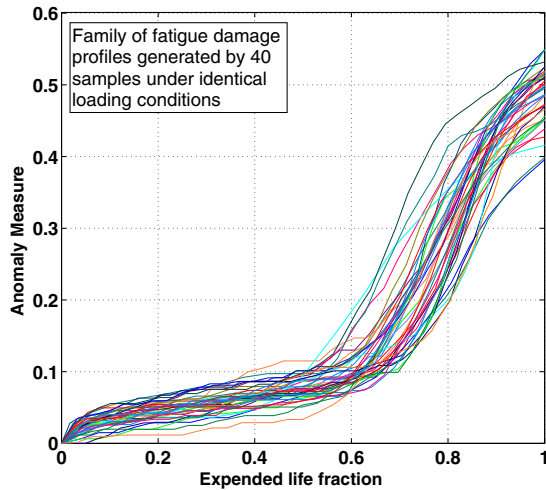
Figure 7. Statistical behaviour of fatigue damage represented by a family of anomaly measure profiles generated by 40 identical experiments.

many of them having very small or zero probabilities. The algorithm is readily implemented in real time. This algorithm enables detection of crack initiation with only eight states and is computationally very fast in the sense that the code execution time is several orders of magnitude smaller than the interval between two adjacent slow time epochs. The wavelet basis, 'gaus2', provides better results than the wavelet bases of the Daubechies family [20] because the 'gaus2' wavelet base closely matches the shape of the ultrasonic signals [19].

5.1.3. Generation of statistical patterns. Similar to the procedure described above, ultrasonic time series data are generated under both nominal and anomalous conditions at different slow time epochs for multiple experiments conducted on identically manufactured specimens under identical experimental conditions. STSA-based analysis of the data from each of these experiments produces a profile of anomaly measure, thereby generating an ensemble of anomaly measure profiles for multiple experiments. This family of profiles represents a stochastic pattern of the progression of fatigue damage under identical experimental conditions. To this effect, $\ell = 40$ experiments have been conducted and the profiles of anomaly measures are shown in figure 7. The family of the anomaly measure profiles of these experiments is plotted versus a normalized variable, *expended life fraction*, $\tau_e = \left(\frac{t-t_0}{t_f-t_0}\right)$, where t is the actual number of cycles, t_0 is the nominal condition chosen to be ~ 0.5 kilocycles for each experiment and t_f is the final time of failure for each experiment as described in the previous section. (Note: the *expended life fraction* τ_e is normalized between 0 and 1.) It is also noted that an infinite number of experiments are required to generate the true stochastic pattern of fatigue damage. This paper reports the results based on $\ell = 40$ experiments for limitations in a laboratory environment. It is assumed that the statistics generated by these finite numbers of experiments closely represent the behaviour of the true population consisting of infinite experiments.

For each individual experiment, the state probability vector $\mathbf{p}^0$ is generated at the nominal condition t_0 by

partitioning the wavelet domain using the maximum entropy principle [5]. As a consequence, $\mathbf{p}^0$ has uniform distribution, i.e. each element has equal probability. This fact demonstrates the fact that there is maximum uncertainty about the origin of the fatigue damage or the available information on potential damage is minimal at the nominal condition. In contrast, for the completely broken stage of the specimen, the entire probability distribution is concentrated on only one element of the state probability vector, i.e. delta distribution, which indicates a very large attenuation of the ultrasonic signal [17]. This phenomenon of the sample being completely broken signifies certainty of information and hence zero entropy. Therefore, as the fatigue crack damage evolves, the uniform distribution (i.e., maximum entropy) under nominal condition degenerates towards the delta distribution (i.e., zero entropy) for the broken specimen. Since, $\mathbf{p}^0$ has uniform distribution for all experiments, the statistical property is identical for all experiments at the nominal condition. The variations in the evolution of fatigue damage, as seen in figure 7, are caused due to various uncertainties listed earlier.

As seen in figure 7, each profile has a smaller slope of the anomaly measure during the initial period of fatigue damage, i.e., the crack initiation region. Anomaly measure gradually increases during this period where small microstructural damage occurs in the specimen. During the end stage of this period small micro cracks eventually develop into a single large crack leading to a transition from the crack initiation stage to the crack propagation stage (approx. from $\tau_e = 0.5$ to $\tau_e = 0.7$). This phenomenon is observed by a sharp change in the slope of the anomaly measure profile of each sample. Once the crack propagation stage starts the fatigue damage occurs rapidly eventually leading to the final failure. The two regions of crack initiation and crack propagation can be clearly identified from figure 7. The family of anomaly measure profiles, generated in the forward problem, is used in the analysis of the inverse problem for estimating the *expended life fraction* (τ_e) at any particular time epoch as presented in the next section.

5.2. Solution procedure of the inverse problem and results

The objective of the inverse problem is identification of anomalies and estimation of the fault parameters based on the family of curves generated in the forward problem. It is essential to detect the evolving fatigue damage and to estimate the remaining useful life during the operating period of the mechanical system, so that appropriate remedial action(s) can be taken before the onset of widespread fatigue propagation leading to complete failure. Therefore, estimation of fatigue damage is crucial for scheduled maintenance. The steps required for the solution procedure of the inverse problem are listed in section 3 and are further elaborated in this section.

5.2.1. Generation of the pattern matrix. In an online experiment, time series data sets (at fast time scale) of the ultrasonic sensors are generated at different slow time epochs up till the current time epoch. These data sets are analysed using the STSA method as discussed in section 2 to generate the probability distributions at the corresponding slow time epochs. The value of the anomaly measure at the current time
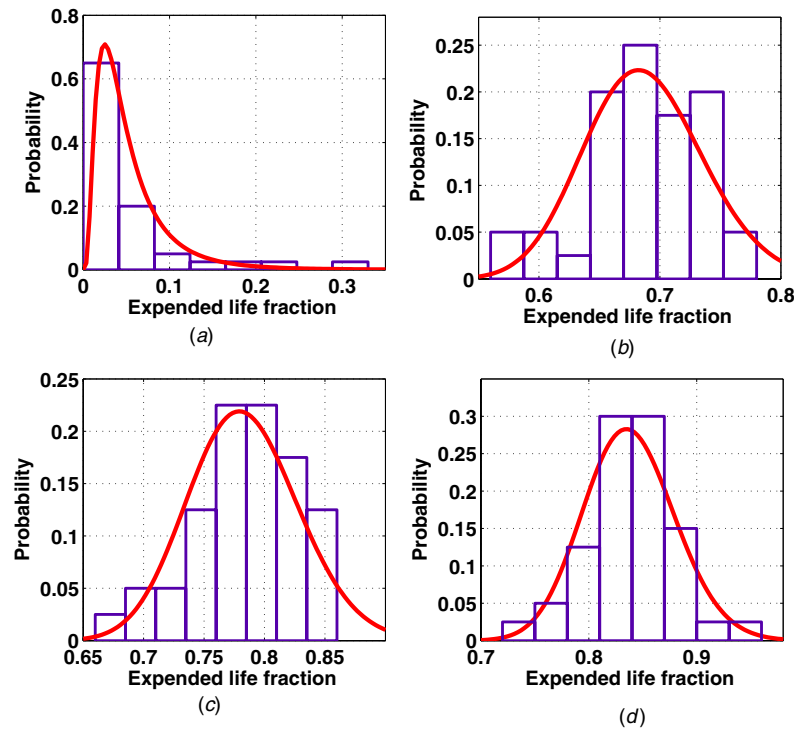
Figure 8. Plots of actual probability distribution of expended life fraction and corresponding *log-normal* fit at four different values of anomaly measure: (*a*) $\psi = 0.025$, (*b*) $\psi = 0.160$, (*c*) $\psi = 0.280$, (*d*) $\psi = 0.350$.

epoch is then calculated from the evolution of this probability vector from the nominal healthy condition. As such, the information available at any particular instant in a real-time experiment is the value of the anomaly measure calculated at that particular instant. Based on this derived value of the anomaly measure the exact determination of the *expended life fraction* (τ_e) is not possible due to the variations observed in the statistical family as seen in figure 7. Therefore, due to uncertainty in determining its exact value at a particular value of anomaly measure, τ_e can be treated as a random variable [24].

The range of anomaly measure (i.e. the ordinate in figure 7) is partitioned into $h = 100$ uniformly spaced levels. A pattern matrix $\mathcal{T}$ of dimension $\ell \times h$ is then derived from the anomaly measure profiles shown in figure 7. The elements of $\mathcal{T}$ are derived such that each column of $\mathcal{T}$ corresponds to the values of τ_e measured for ℓ samples at the corresponding anomaly measure. As such, the elements of each column of $\mathcal{T}$ describe a distribution of the random variable τ_e. That is, for each value of the anomaly measure, there is a distribution associated with τ_e over a certain range.

5.2.2. Estimation of the expended life fraction with confidence intervals. In order to estimate the value of τ_e by statistical means, a two-parameter *lognormal* distribution [6, 28, 29] is hypothesized for each column of $\mathcal{T}$. The lognormal probability density function of the random variable τ_e is defined as:

$$p_{\tau_e}(x) = \frac{1}{\sqrt{2\pi}\sigma x}\exp\left(\frac{-(\ln(x) - \mu)^2}{2\sigma^2}\right)\mathcal{U}(x) \qquad (3)$$

where $\mathcal{U}(\bullet)$ is the standard Heaviside unit step function; and μ and σ are respectively the mean and standard deviation of the

Gaussian distributed random variable $\ln(\tau_e)$. The lognormal distribution is obtained for each column of $\mathcal{T}$ over the mean and the variance of τ_e. The goodness of fit is examined by both χ^2 and Kolmogorov–Smirnov tests [30]. The number of bins was taken to be $r = 8$ for the data set of each column of $\mathcal{T}$. With $f = r - 2 - 1 = 5$ degrees of freedom, the χ^2-test shows that, for each of the h data sets, the hypothesis of the two-parameter lognormal distribution passed the 20% significance level [30] which suffices for the conventional standard of 5% significance level. Also, for each of the h data sets, the hypothesis passed the 20% significance level of the Kolmogorov–Smirnov test which again suffices for the conventional standard of 5% significance level. A good discussion of these statistical tests is provided in [30].

Figure 8 shows the histograms of τ_e generated at four different anomaly measure values and the associated lognormal distribution. Figure 8 shows that the lognormal distribution is a good approximation of the histograms of τ_e. This is because the lognormal distribution easily models the failure rate shapes. The lognormal distribution for failure analysis is commonly used in the literature [29]. The rationale for the choice of the lognormal distribution is presented below.

- The fact that the lognormal distribution is one directional on the positive axis is consistent with the fatigue life of a structure;
- The probability of failure is in general high for a certain range of fatigue cycles and gradually decreases thereafter. This behaviour is easily modelled by the lognormal distribution, i.e., the failure rate shapes and the pdf of lognormal distribution are physically meaningful;
- Since the variable $\log(\tau_e)$ is Gaussian, many standard statistical tools are available for further analysis. This

Table 1. Statistical information on expended life fraction (τ_e).

Anomaly measure (ψ)	Actual expended life fraction (τ_e) for three new test samples			Information generated from the family of forward data sets							
				Statistical parameter estimates					Confidence interval bounds		
	No. 1	No. 2	No. 3	μ	σ	μ_{τ_e}	σ_{τ_e}	$\widehat{\tau}_e$	95%	85%	75%
0.0127	0.0315	0.0109	0.0296	−3.7305	0.6732	0.0301	0.0228	0.0152	0.0065, 0.0927	0.0093, 0.0652	0.0113, 0.0535
0.0545	0.3687	0.1839	0.2623	−1.3815	0.4630	0.2796	0.1367	0.2027	0.1028, 0.6232	0.1305, 0.4903	0.1490, 0.4289
0.0724	0.4392	0.4908	0.4151	−0.8192	0.2664	0.4567	0.1238	0.4106	0.2621, 0.7430	0.3009, 0.6470	0.3250, 0.5989
0.0962	0.6447	0.5809	0.6263	−0.5467	0.1286	0.5837	0.0754	0.5693	0.4501, 0.7449	0.4812, 0.6967	0.4994, 0.6712
0.1256	0.6800	0.6689	0.6758	−0.4287	0.0789	0.6534	0.0517	0.6473	0.5580, 0.7604	0.5814, 0.7298	0.5949, 0.7133
0.1654	0.7143	0.7008	0.7305	−0.3596	0.0703	0.6997	0.0493	0.6945	0.6081, 0.8011	0.6307, 0.7723	0.6437, 0.7567
0.2068	0.7440	0.7288	0.7688	−0.3090	0.0651	0.7357	0.0479	0.7311	0.6463, 0.8340	0.6685, 0.8063	0.6813, 0.7912
0.2415	0.7633	0.7515	0.7885	−0.2740	0.0613	0.7618	0.0467	0.7575	0.6743, 0.8574	0.6961, 0.8305	0.7086, 0.8159
0.2738	0.7797	0.7712	0.8038	−0.2482	0.0586	0.7816	0.0458	0.7776	0.6956, 0.8752	0.7171, 0.8489	0.7294, 0.8346
0.3025	0.7952	0.7876	0.8215	−0.2231	0.0548	0.8012	0.0440	0.7976	0.7185, 0.8908	0.7393, 0.8657	0.7511, 0.8521
0.3465	0.8181	0.8149	0.8518	−0.1850	0.0510	0.8322	0.0425	0.8290	0.7521, 0.9185	0.7723, 0.8944	0.7838, 0.8813
0.3875	0.8472	0.8397	0.8818	−0.1456	0.0515	0.8657	0.0446	0.8622	0.7816, 0.9563	0.8028, 0.9311	0.8148, 0.9173
0.4258	0.8811	0.8704	0.9133	−0.0988	0.0548	0.9073	0.0497	0.9032	0.8137, 1.0000	0.8372, 0.9803	0.8506, 0.9648
0.4521	0.9145	0.8932	0.9374	−0.0660	0.0500	0.9373	0.0469	0.9338	0.8487, 1.0000	0.8711, 1.0000	0.8838, 0.9915
0.4702	0.9450	0.9202	0.9624	−0.0434	0.0419	0.9584	0.0402	0.9559	0.8821, 1.0000	0.9015, 1.0000	0.9125, 1.0000

μ = mean of $\ln \tau_e$, σ^2 = variance of $\ln \tau_e$, $\mu_{\tau_e} = \exp(\mu + \sigma^2/2)$; $\sigma_{\tau_e}^2 = (\exp(\sigma^2) - 1)\mu_{\tau_e}^2$; and maximum likelihood estimate $\widehat{\tau}_e = \exp(\mu - \sigma^2)$.
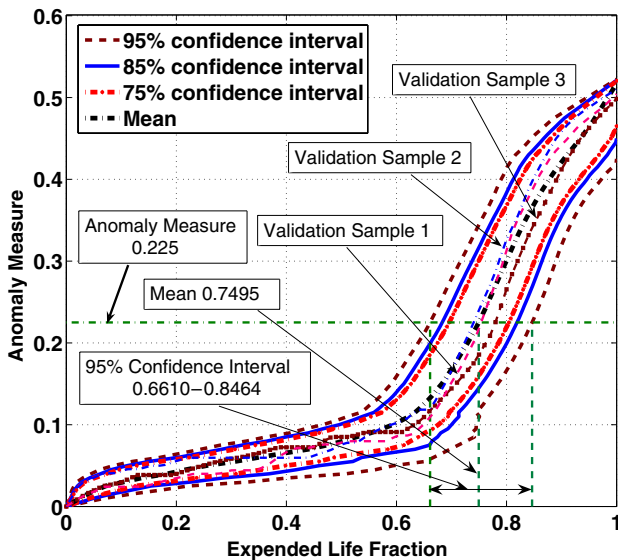

Figure 9. Plots of confidence interval bounds are shown at three different confidence levels of 95%, 85% and 75%. Profiles of anomaly measure are also shown for three new validation test specimens.

makes the lognormal distribution a very natural choice for failure analysis.

Once the lognormal distributions are obtained, the confidence interval bounds at different confidence levels can be computed from the properties of the distribution using elementary statistics [30, 31]. The confidence level signifies the probability that the estimated parameter will lie within the corresponding confidence interval. As an example, for a confidence level of 95%, the probability that the actual parameter will lie between the specified confidence intervals is 95%. Figure 9 provides the plots of confidence interval bounds at three different confidence levels of 95%, 85% and 75%. As an illustration in figure 9, the confidence interval bounds at 95% confidence level are shown for an arbitrary value of anomaly measure equal to 0.225 (lower bound =

0.6610 and upper bound = 0.8464). The estimate $\widehat{\tau}_e$ of the expended life fraction τ_e can be obtained at the point of highest probability, i.e. the mean of the distribution. The other useful parameter is the *remaining life fraction* whose estimate $\widehat{\tau}_r$ is obtained at any instant as: $\widehat{\tau}_r = 1 - \widehat{\tau}_e$. The information on the remaining life estimate in a real-time experiment is useful for development of life extending control and resilient control strategies for prevention of widespread structural damage and catastrophic failures.

5.2.3. Experimental validation and discussion. The proposed methodology is validated by fatigue experiments on three new test specimens. The profiles of anomaly measure are computed using the STSA method for these three test specimens. Figure 9 also shows the profiles of the three test samples along with the plots of confidence interval bounds derived from the statistical ensemble. The estimates of the mean $\widehat{\tau}_e$ of the expended life fraction with the standard deviation $\widehat{\sigma}$ are obtained at (arbitrary) different values of the anomaly measure using the procedure described in the previous section. The results are interpolated for values of the anomaly measure that lie in between the two columns of the pattern matrix $\mathcal{T}$. Confidence interval bounds are obtained at three different confidence levels of 95%, 85% and 75%. Table 1 provides the actual and the estimated values of the expended life fraction for the three test specimens along with the various confidence bounds for arbitrarily chosen values of anomaly measures.

As seen in figure 7, there are two phases of crack growth which can be identified from the slope of the anomaly measure profiles. The two phases can be broadly classified as the crack initiation phase and the crack propagation phase. At lower values of anomaly measure the profiles of all specimens indicate a relatively smaller slope than for higher values of anomaly measure. The change in the rate of fatigue damage progression as observed in figure 7 occurs at the transition from the crack initiation phase to the crack propagation phase (approx from $\tau_e = 0.5$ to $\tau_e = 0.7$). Figure 9 shows that the uncertainty in the fatigue damage is higher in the

crack initiation phase as indicated by the width of confidence intervals for any particular value of the anomaly measure. Subsequently, upon onset of the crack propagation phase, the confidence intervals are significantly more tight than those in the crack initiation phase. This observation is explained by the fact that the uncertainty in the crack initiation phase depends on the random distribution of flaws in the specimen [26]. Furthermore, the uncertainty in the estimation of expended life is amplified in the crack initiation region because of the fact that the anomaly measure profiles have a smaller slope. During this crack initiation phase, small cracks originate from the microstructural damage (e.g. dislocations, voids and inclusions) at multiple sites in the entire body of the material structure causing a high uncertainty in fatigue damage behaviour. These multiple small cracks eventually develop into a single large crack leading to the onset of the crack propagation phase. Therefore, the uncertainty in the crack initiation phase relates to the inhomogeneity in the material and non-uniform distribution of the initial conditions in the specimen causing stress augmentation at certain locations which directly affects the formation of small cracks. The small slope of the profiles of fatigue damage growth further enhances the uncertainty in fatigue life estimation. Once a single large crack develops, i.e., the crack propagation stage starts, then the dynamics of subsequent damage are relatively simple due to high stress concentration at the crack tip.

Table 1 compiles the pertinent statistical information on expended life fraction. Column 1 lists randomly selected values of anomaly measure ψ as a representation of the information generated from observed time series data. The results are interpolated for values of the anomaly measure that lie in between the two columns of the pattern matrix $\mathcal{T}$. Columns 2–4 show the observed values of expended life fraction τ_e for three additional test specimens that are not included in the original ensemble of $\ell = 40$ specimens. Columns 5–12 list the statistical information, extracted from the ensemble of original $\ell = 40$ specimens, and the details are described below.

- Two parameters μ and σ of the lognormal distribution (see equation (3)) of the expended life fraction τ_e, which are respectively the mean and standard deviation of $\ln(\tau_e)$. Correspondingly, the mean and variance of τ_e are respectively obtained as:

  $$\mu_{\tau_e} = \exp(\mu + \sigma^2/2) \quad \text{and} \quad \sigma_{\tau_e}^2 = (\exp(\sigma^2) - 1)\mu_{\tau_e}^2.$$

- Maximum likelihood estimate $\widehat{\tau}_e = \exp(\mu - \sigma^2)$ of the expended life fraction τ_e, i.e., where the probability density function p_{τ_e} attains the maximum value.

- Confidence interval bounds for the expended life fraction τ_e, at three different confidence levels of 95%, 85% and 75%.

As seen in table 1, the confidence intervals are relatively wider during the crack initiation region, i.e. for low anomaly measure values. This is due to the high uncertainty associated with the crack initiation phase. After the onset of crack propagation phase when a single large crack starts progressing, the uncertainty reduces and the confidence bounds come closer to the mean (see, for example, the rows of table 1

for anomaly measure values $\gtrsim 0.12$). As such, the information that is derived during the crack initiation phase can act as an early warning of the onset of widespread fatigue in the crack propagation phase. The information from table 1 and figure 9 (including the estimate of τ_e and different confidence intervals) can be utilized for real-time monitoring of the fatigue damage and to obtain advance warning of complete failure. Furthermore, the information can be used in future research for the development of probabilistic robust control strategies for damage mitigation and prevention of catastrophic failures. This work is still under active investigation and will be reported in forthcoming publications.

6. Summary, conclusions and future work

This paper presents a dynamic data-driven statistical method for detection of fatigue damage at an early stage and online estimation of the remaining useful fatigue life in polycrystalline alloy structures. The proposed method is built upon the principles of *symbolic dynamics*, *information theory* and *statistical pattern recognition*. Specifically, ultrasonic signals are analysed for detection of microstructural changes in 7075-T6 alloy specimens on a fatigue damage test apparatus.

This paper has addressed damage monitoring and remaining useful life estimation for a single source of faults. In human-engineered complex systems, there exist multiple correlated sources of anomalies that may evolve to catastrophic failures. The proposed tool is potentially applicable to dynamical systems with multiple sources of correlated faults, where it would be necessary to fuse the multivariate statistical information into a univariate distribution of the states of a finite-state automaton model [15]. Nevertheless, further theoretical and experimental research is necessary before its application to such problems that are prevalent in industry.

The reported work is a step towards building a reliable instrumentation system for early detection of fatigue damage and real-time estimation of remaining useful fatigue life in mechanical structures of human-engineered systems. For example, the information on current health status and remaining useful life could be used to update the decision and control laws online to enhance plant performance or avert forthcoming failures. As such future research areas include life extending control and resilient control in complex engineering systems. While there are many research issues that need to be addressed, the following potential future tasks are envisioned:

- Solution of the inverse problem under more complicated scenarios for multiple parameter estimations and under variable amplitude and random loading conditions.
- Development of a probabilistic robust control strategy based on the statistical information derived from the forward problem for real-time life extension, damage mitigation and prevention of catastrophic failures [23].

References

[1] Sobczyk K and Spencer B F 1992 *Random Fatigue: Data to Theory* (New York: Academic)
[2] Pathria R K 1996 *Statistical Mechanics* (Amsterdam: Elsevier)

[3] Ozekici S 1996 *Reliability and Maintenance of Complex Systems (NATO Advanced Science Institutes (ASI) Series F Computer and Systems Sciences* vol 154*)* (Dordrecht: Kluwer)

[4] Ott E 1993 *Chaos in Dynamical Systems* (Cambridge: Cambridge University Press)

[5] Gupta S, Ray A and Keller E 2006 Symbolic time series analysis of ultrasonic signals for fatigue damage monitoring in polycrystalline alloys *Meas. Sci. Technol.* **17** 1963–73

[6] Ray A 2004 Stochastic measure of fatigue crack damage for health monitoring of ductile alloy structures *Struct. Health Monit.* **3** 245–63

[7] Scafetta N, Ray A and West B 2006 Correlation regimes in fluctuations of fatigue crack growth *Physica* A **59** 1–23

[8] Keller E and Ray A 2003 Real time health monitoring of mechanical structures *Struct. Health Monit.* **2** 191–203

[9] Abarbanel H 1996 *The Analysis of Observed Chaotic Data* (New York: Springer)

[10] Rokhlin S and Kim J-Y 2003 *In situ* ultrasonic monitoring of surface fatigue crack initiation and growth from surface cavity *Int. J. Fatigue* **25** 41–9

[11] Kenderian S, Berndt T, Green R and Djordjevic B 2003 Ultrasonic monitoring of dislocations during fatigue of pearlitic rail steel *Mater. Sci. Eng.* **348** 90–9

[12] Anson L, Chivers R and Puttick K 1995 On the feasibility of detecting pre-cracking fatigue damage in metal matrix composites by ultrasonic techniques *Compos. Sci. Technol.* **55** 63–73

[13] Vanlanduit S, Guillaume P and Linden G 2003 Online monitoring of fatigue cracks using ultrasonic surface waves *NDT&E Int.* **36** 601–7

[14] Daw C, Finney C and Tracy E 2003 A review of symbolic analysis of experimental data *Rev. Sci. Instrum.* **74** 915–30

[15] Ray A 2004 Symbolic dynamic analysis of complex systems for anomaly detection *Signal Process.* **84** 1115–30

[16] Chin S, Ray A and Rajagopalan V 2005 Symbolic time series analysis for anomaly detection: a comparative evaluation *Signal Process.* **85** 1859–68

[17] Gupta S, Ray A and Keller E 2007 Symbolic time series analysis of ultrasonic data for early detection of fatigue damage *Mech. Syst. Signal Process.* **21** 866–84

[18] Keller E E 2001 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

[19] Rajagopalan V and Ray A 2006 Symbolic time series analysis via wavelet-based partitioning *Signal Process.* **86** 3309–20

[20] Mallat S 1998 *A Wavelet Tour of Signal Processing 2/e* (New York: Academic)

[21] Badii R and Politi A 1997 *Complexity Hierarchical Structures and Scaling in Physics* (Cambridge: Cambridge University Press)

[22] Tarantola A 2005 *Inverse Problem Theory* (Philadelphia, PA: Society for Industrial and Applied Mathematics)

[23] Khatkhate A, Gupta S, Ray A and Keller E 2006 Life extending control of mechanical structures using symbolic time series analysis *American Control Conf. (Minnesota, USA)*

[24] Suresh S 1998 *Fatigue of Materials* (Cambridge: Cambridge University Press)

[25] Rose J 2004 *Ultrasonic Waves in Solid Media* (Cambridge: Cambridge University Press)

[26] Klesnil M and Lukas P 1991 *Fatigue of Metallic Materials (Material Science Monographs* vol 71*)* (Amsterdam: Elsevier)

[27] Mathworks Inc. 2006 *MATLAB Wavelet Toolbox*

[28] Bogdanoff J and Kozin F 1985 *Probabilistic Models of Cumulative Damage* (New York: Wiley)

[29] http://www.itl.nist.gov/div898/handbook/eda/section3/eda3669.htm, *NIST/SEMATECH e-Handbook of Statistical Methods*

[30] Brunk H 1995 *An Introduction to Mathematical Statistics* 3rd edn (Lexington, MA: Xerox)

[31] Snedecor G W and Cochran W G 1989 *Statistical Methods* 8th edn (Ames, IA: Iowa State University Press)