A MODEL RESIDUAL BASED SEQUENTIAL PROBABILITY RATIO TEST FRAMEWORK FOR STRUCTURAL DAMAGE DIAGNOSIS

Fotis Kopsaftopoulos, Kyriakos Vamvoudakis-Stefanou, and Spilios Fassois

ABSTRACT

A model residual based Sequential Probability Ratio Test (SPRT) framework for vibration based structural damage diagnosis is introduced. This employs the residual sequences obtained from a single statistical time series model of the healthy structure, and its performance is pre-determined via the use of the Operating Characteristic and Average Sample Number functions in combination with baseline experiments. The approach postulated in this framework is shown to achieve early damage detection and identification (classification) via its application to damage diagnosis on a GARTEUR scale aircraft skeleton structure. Comparisons with a non-parametric Power Spectral Density based method are also presented.

Keywords: Damage Detection, Damage Identification, Structural Health Monitoring, Time Series Methods, Sequential Methods

1. INTRODUCTION

Statistical time series methods form an important and rapidly evolving class, within the broader vibration based family of Structural Health Monitoring (SHM) methods [1–4]. Their main elements are: (i) random excitation and/or vibration response signals (time series), (ii) statistical model building, and (iii) statistical decision making for inferring the health state of a structure. They offer a number of potential advantages, including no requirement for physics based or finite element models as they are data based (inverse type) methods, no requirement for complete modal models, effective treatment of uncertainties, and statistical decision making with specified performance characteristics [1, 2].

The vast majority of statistical time series SHM methods is based on Fixed Sample Size (FSS) hypothesis testing procedures used in the statistical decision making. On the other hand, sequential methods have the feature that the number of observations required is not determined in advance, but depends, at each stage, on the results of the observations previously made. Thus, the number of observations required by the test is not predetermined, but a random variable. A merit of the sequential approach, as applied to testing statistical hypotheses, is that test procedures can be constructed which require, on average, a substantially smaller number of observations than equally reliable test procedures based on a predetermined (fixed)
number of observations [5]. Moreover, a potential advantage of a damage diagnosis method based on
sequential procedures is its straightforward extension for online implementation, which is of high interest
within the SHM context. In this context, preliminary – with respect to the use of a Sequential Probability
Ratio Test (SPRT) scheme – studies include [6, 7], where the binary form of the SPRT based on AR-ARX
model residuals has been applied for damage detection in a laboratory three-story building model and an
eight-degree-of-freedom mass-spring system, respectively.

The goal of the present study is the introduction and experimental assessment of a model residual based
SPRT framework for SHM capable of achieving early and robust damage detection and identification
under uncertainties. This framework is based on the statistically optimal SPRT (both its binary and mul-
tihypothesis versions [8, 9]), while taking advantage – for the first time in the context of vibration based
SHM – of its properties and capabilities. The basis of the proposed framework consists of the residual
sequences obtained through a single stochastic time series model of the healthy structural dynamics. Its
effectiveness is validated and experimentally assessed via its application to a GARTEUR aircraft scale
skeleton structure and damages that correspond to loosening of various bolts that connect its structural ele-
ments. This structure has been employed by the authors and their collaborators in various SHM studies
using either similar loosened-bolt damage types [10] or other damage types [11, 12].

The results presented herein are for three vibration response measurement positions, with a single mea-
surement used at a time. The random force excitation is provided via an electromechanical shaker, while
the vibration responses are measured via lightweight accelerometers. The main features and operational
characteristics of the framework are discussed along with practical issues. Comparisons with a non-
parametric Power Spectral Density (PSD) based method [1, 4] are also made.

2. THE MODEL RESIDUAL BASED SPRT FRAMEWORK

The postulated framework consists of two phases: (a) An initial baseline phase, which includes the
modeling of the healthy structure, and (b) the inspection phase, which is performed during the structure’s
service cycle or continuously (online), and includes the functions of damage detection and identification.

2.1. Baseline phase

Data records from the healthy structure are employed for the identification of an appropriate parametric
time series model. Specifically a scalar (univariate) model is needed in case of a single vibration response
measurement location, or a vector (multivariate) model (or suboptimally an array of scalar models) is
needed in case that more vibration response measurement locations are to be simultaneously used. In the
linear response-only case, AutoRegressive (AR) or AutoRegressive Moving Average (ARMA) models
may be employed [13]. In the present study a single measurement location, and a corresponding scalar
AutoRegressive (AR) model, is used at a time.

2.2. Inspection Phase

Damage detection and identification are based on the binary and multihypothesis versions of the SPRT,
respectively [8, 9], which are used in order to detect a change in the standard deviation \( \sigma \) of the model
residual sequence obtained by driving the current (unknown) response signals through a single baseline
healthy model. The SPRT allows for the specification of two values \( \sigma_0 \) and \( \sigma_1 \) for the standard deviation,
so that the structure is determined in healthy state iff \( \sigma \leq \sigma_0 \), and in damaged state iff \( \sigma \geq \sigma_1 \). The
zone between \( \sigma_0 \) and \( \sigma_1 \) constitutes an uncertainty zone, thus for \( \sigma \) lying in it the decision is postponed
and data collection continues. The values of \( \sigma_0 \) and \( \sigma_1 \) are user defined and express the increase of the
standard deviation ratio \( q = \sigma_1 / \sigma_0 \) for which the structure is considered to be damaged. For example, a
ratio of \( q = 1.1 \) means that the structure is considered damaged whenever there is an increase of 10% in
the standard deviation \( \sigma \) of the current residual sequence compared to a threshold value \( \sigma_0 \).

Damage detection is based on the binary hypothesis testing problem implemented via the SPRT of
strength \( (\alpha, \beta) \), with \( \alpha, \beta \) designating the type I (false alarm) and II (missed damage) error probabil-
ities, respectively:

\[
H_0 : \quad \sigma \leq \sigma_0 \quad \text{(null hypothesis – healthy structure)}
\]

\[
H_1 : \quad \sigma \geq \sigma_1 \quad \text{(alternative hypothesis – damaged structure)}
\]  

(1)
with \(\sigma\) designating the standard deviation of a scalar model residual signal \(e[t]\) obtained by driving the current response signal through the healthy structural model, and \(\sigma_o, \sigma_1\) user defined values. Under the null hypothesis of a healthy structure the residuals \(e[t]\) are iid zero mean Gaussian with variance \(\sigma^2\), hence \(e[t] \sim \mathcal{N}(0, \sigma^2)\).

The basis of the SPRT is the logarithm of the likelihood ratio function which is computed at data sample \(t\) (presently coinciding with discrete time) as follows:

\[
\Lambda[t] = \log \frac{f(e[1], \ldots, e[t]|H_1)}{f(e[1], \ldots, e[t]|H_0)} = \sum_{i=1}^{t} \log \frac{f(e[t]|H_1)}{f(e[t]|H_0)} = t \cdot \log \frac{\sigma_o}{\sigma_1} + \frac{\sigma_o^2 - \sigma_1^2}{2\sigma_o^2 \sigma_1^2} \sum_{i=1}^{t} e^2[t], \; t = 1, 2, \ldots
\]

with \(\Lambda[t]\) designating the decision parameter of the method and \(f(e[t]|H_i)\) the probability density function (normal distribution) of the residual sequence under hypothesis \(H_i\) \((i = 0, 1)\).

Decision making is then based on the test (of strength \((\alpha, \beta)\)):

\[
\begin{align*}
\Lambda[t] &\leq \log B \quad \text{accept } H_o \quad \text{(healthy structure)} \\
\Lambda[t] &\geq \log A \quad \text{accept } H_1 \quad \text{(damaged structure)} \\
\log B < \Lambda[t] &\leq \log A \quad \text{no decision is made} \quad (\text{continue the test})
\end{align*}
\]

with \(A = (1 - \beta)/\alpha\) and \(B = \beta/(1 - \alpha)\). Following a decision at a stopping sample (time) \(\hat{T}\), it is possible to continue the test by resetting \(\Lambda[\hat{T}+1]\) to zero and continuing by collecting additional residual samples.

For any value of the residual standard deviation \(\sigma\), the Operating Characteristic (OC) function of the SPRT denotes the probability that the test will terminate with the acceptance of the null hypothesis \(H_o\) [8]. Similarly, the Average Sample Number (ASN) function represents the average number of inspection samples required by the SPRT to reach a decision [8]. The ASN is an approximation of the expected value \(E_o[T]\) of the number of residual samples required by a sampling plan of strength \((\alpha, \beta)\) and standard deviations \(\sigma_o, \sigma_1\) in order to reach a terminal decision.

Damage identification is based on the multihypothesis sequential test, which is based on the Armitage test [8, 9]. Then, considering \(k\) hypotheses \((k\) potential damage states), the multihypothesis test to be implemented may be expressed as follows:

\[
\begin{align*}
H_A & \colon \sigma = \sigma_A \quad \text{Hypothesis A – damage is of type A} \\
H_B & \colon \sigma = \sigma_B \quad \text{Hypothesis B – damage is of type B} \\
\vdots & \vdots
\end{align*}
\]

The standard deviation values \(\sigma_A, \sigma_B, \ldots\) are user defined and may be determined based on available baseline data obtained from the structure under damage types \(A, B, \ldots\), respectively. A typical selection of \(\sigma_A, \sigma_B, \ldots\) could be as the mean values of the residual standard deviations estimated from the available baseline data records under the corresponding damage structural states. By denoting the log likelihood under hypothesis \(H_i\) \((H_i \text{ is true, } i = A, B, \ldots)\) as \(L_i\) there are \(\frac{1}{2}k(k - 1)\) log likelihood ratios for the various pairs of hypotheses, with each one expressed in terms of \(k - 1\) independent likelihood ratios [9, 14]:

\[
\Lambda_{ij}[t] = \frac{L_j[t]}{L_i[t]} = t \cdot \log \frac{\sigma_j}{\sigma_i} + \frac{\sigma_j^2 - \sigma_i^2}{2\sigma_j^2 \sigma_i^2} \sum_{i=1}^{t} e^2[t], \; i, j = A, B, \ldots \; \text{and} \; i \neq j.
\]

Then, the multihypothesis test termination is defined by the pair \((T, \delta)\), with \(T\) indicating the stopping time and \(\delta\) the final decision [14, pp. 237–238]:

\[
\hat{T} = \min \inf \left\{ \frac{t}{\Lambda_{ij}[t]} \geq \log A_{ij} \quad \forall i \neq j, \; i < j, \; t = 1, 2, \ldots \right\}, \quad \hat{\delta} = \arg \min_{j=1,\ldots,k} T.
\]

Let \(a_{ij}\) the probability of accepting \(H_i\) when in fact \(H_j\) is true (error probabilities), that is \(a_{ij} = P(\delta = H_j|H_i), \; i \neq j\), and let \(a_{ii}\) the probability of accepting \(H_i\) when in fact \(H_i\) is true (correct decision probabilities), that is \(a_{ii} = P(\delta = H_i|H_i)\). The error probabilities \(a_{ij}\) may be controlled via suitable selection of the \(A_{ij}\)’s [8, 9].
3. THE STRUCTURE AND THE EXPERIMENTAL SET-UP

The scale aircraft skeleton laboratory structure was designed by ONERA in conjunction with the GARETEUR SM–AG19 Group and manufactured at the University of Patras (Figure 1). It represents a typical aircraft skeleton design and consists of six solid beams with rectangular cross sections representing the fuselage ($1500 \times 150 \times 50$ mm), the wing ($2000 \times 100 \times 10$ mm), the horizontal ($300 \times 100 \times 10$ mm) and vertical stabilizers ($400 \times 100 \times 10$ mm), and the right and left wing-tips ($400 \times 100 \times 10$ mm). All parts are constructed from standard aluminium (total mass 50 kg).

The structure is suspended through a set of bungee cords and hooks from a long rigid beam sustained by two heavy-type stands. The suspension is designed in a way as to exhibit a pendulum rigid body mode below the frequency range of interest, as the boundary conditions are free-free. The excitation is broadband random stationary Gaussian applied vertically at the right wing-tip (Point X, Figure 1) through an electromechanical shaker (MB Dynamics Modal 50A, max load 225 N). The vertical acceleration responses at Points Y1–Y3 (Figure 1) are measured via lightweight accelerometers (PCB 352A10). The acceleration signals are driven through a conditioning charge amplifier (PCB 481A02) into the data acquisition system based on SigLab 20–42 modules. Five damage types (designated as A, B, . . . , E) are presently considered (Table 1), each one corresponding to the complete loosening of one or more bolts at different joints of the structure.

1600 experiments per structural state are undertaken, 100 of which from each structural state are employed in the baseline phase – the rest are used in the inspection phase (assessment); see Table 1. In each experiment vibration measurements are collected at Points Y1, Y2, Y3 (Figure 1). Further experimental details are provided in Table 1 – worth noting is the very low/limited bandwidth used. The sample mean is subtracted from each signal and scaling by the signal’s sample standard deviation is implemented.

| Table 1 The considered damage types, number of experiments, and vibration signal details. |
|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|
| Structural State | Description | Total Number of Experiments |
| Healthy | — | 1600 (100 baseline) |
| Damage type A | loosening of bolts A1, A4, Z1, and Z2 | 1600 (100 baseline) |
| Damage type B | loosening of bolts D2 and D3 | 1600 (100 baseline) |
| Damage type C | loosening of bolts D1, D2, and D3 | 1600 (100 baseline) |
| Damage type D | loosening of bolt K1 | 1600 (100 baseline) |
| Damage type E | loosening of bolt K1 and K2 | 1600 (100 baseline) |

Sampling frequency: $f_s = 512$ Hz, Signal bandwidth: $[0.5 \mathrm{--} 200]$ Hz

Signal length $N$ in samples (s): Non-parametric analysis: $N = 46 080$ (90 s)

Parametric analysis: $N = 1000$ (1.95 s)
Figure 2 Healthy structure: (a) Operating Characteristic (OC) and (b) Average Sample Number (ASN) functions for various residual standard deviation ratios \( q = \sigma_1/\sigma_o \) and constant strength \((\alpha, \beta) = (0.05, 0.01)\). The vertical colored dashed lines designate the \( \sigma_1 \) values for the corresponding ratios \( q \). The dashed vertical cyan lines represent the residual standard deviation values for each of the 100 baseline healthy data sets.

4. DAMAGE DETECTION AND IDENTIFICATION RESULTS

4.1. Baseline phase: structural identification under the healthy structural state

Parametric AR identification of the structural dynamics is based on \( N = 10 \, 000 \approx 19.5 \, s \) sample-long single response signals (MATLAB function \texttt{arx.m}). The modeling strategy consists of the successive fitting of AR(\( n \)) models (with \( n \) designating the AR order) until a suitable model is selected. Model parameter estimation is based on a Least Squares (LS) estimator [13, p. 206] and model order selection on the BIC and RSS/SSS (Residual Sum of Squares / Signal Sum of Squares) criteria and frequency stabilization diagrams [13]. The selected AR model characteristics are summarized in Table 2 (SPP refers to signal samples per estimated parameter). Note that the identification procedure generally leads to different AR models (including different model orders) for each vibration measurement position.

4.2. Inspection Phase

Prior to implementing the SPRT, an appropriate sampling plan should be selected. The selection of the sampling plan involves the determination of the following three aspects: (i) the nominal residual standard deviation \( \sigma_o \) under which the structure is considered to be in its healthy state, (ii) the standard deviation ratio \( q = \sigma_1/\sigma_o \) which constitutes the standard deviation increase under which the structure is determined to be damaged, and (iii) the SPRT strength \((\alpha, \beta)\).

The determination of the residual standard deviation \( \sigma_o \) under which the structure is considered healthy is based on the available 100 baseline data records obtained from the healthy structure (Table 1). The value \( \sigma_o \) is chosen in order for the probability of \( \sigma \leq \sigma_o \) to be equal to 95\% \((P(\sigma \leq \sigma_o) = 0.95)\). The determination of the residual standard deviation ratio \( q \) may be based on the OC and ASN functions of the SPRT [8] for various \( q \) ratios, along with the use of the baseline data records. Figures 2a and 2b present, for vibration response Y1, the OC and ASN functions, respectively, for various candidate ratios \( q \) and constant SPRT strength \((\alpha, \beta) = (0.05, 0.01)\). In both figures, the \( \sigma_o \) value is shown as gray vertical dashed line, while the \( \sigma_1 \) values corresponding to the considered \( q = \sigma_1/\sigma_o \) ratios are shown in colored vertical dashed lines. Along with the OC and ASN function curves, the standard deviation values obtained from the 100 baseline residual sequences are depicted in vertical cyan dashed lines.

<table>
<thead>
<tr>
<th>Response</th>
<th>Selected Model</th>
<th>No of estimated parameters</th>
<th>SPP</th>
<th>BIC</th>
<th>RSS/SSS (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y1</td>
<td>AR(124)</td>
<td>124 parameters</td>
<td>80.6</td>
<td>-2.82</td>
<td>5.26</td>
</tr>
<tr>
<td>Y2</td>
<td>AR(142)</td>
<td>142 parameters</td>
<td>70.4</td>
<td>-5.82</td>
<td>0.25</td>
</tr>
<tr>
<td>Y3</td>
<td>AR(111)</td>
<td>111 parameters</td>
<td>90.1</td>
<td>-5.78</td>
<td>0.27</td>
</tr>
</tbody>
</table>
Figure 3 Indicative damage detection results (response Y1) at the $(\alpha, \beta) = (0.05, 0.01)$ risk levels ($q = \sigma_1/\sigma_o = 1.08$). The actual structural state is shown above each plot.

In Figure 2a the intersections of the dashed vertical lines, belonging to the residual standard deviation values, with the OC function curves for the various $q$ ratios correspond to the probabilities of acceptance of the null hypothesis $H_0$ (healthy structure) for each ratio, while in Figure 2b correspond to the expected number of residual samples required to reach a decision. The OC function (Figure 2a) is considered more favorable the higher the value of $L(\sigma)$ for $\sigma$ consistent with $H_0$ and the lower the value of $L(\sigma)$ for $\sigma$ not consistent with $H_0$. Thus, by plotting the OC and ASN functions, not only one may have an indication of the probability of acceptance for various residual standard deviations $\sigma$, but one may also obtain an approximation of the number of residual samples that are required for reaching a terminal decision.

**Damage detection results.** Indicative damage detection results for Point Y1 are presented in Figure 3. A damage is detected when the test statistic (vertical axis) exceeds the upper critical point (dashed horizontal lines), while the structure is determined as being in its healthy state when the test statistic exceeds the lower critical point. After a critical point is exceeded a decision is made, while the test statistic is reset to zero and the test continues. Hence, during testing multiple decisions may be made. Evidently, correct detection is obtained in each test case, as the test statistic is shown to exceed multiple times (multiple correct decisions) the lower critical point in the healthy case, while it also exceeds multiple times the upper critical point (multiple correct damage detections) in the damage test cases.

The summarized damage detection results are presented in Table 3. The correct detections and false alarm percentages are based on 1600 inspection experiments. The false alarm rates are extremely low,

<p>| Table 3 Damage detection summary results for the three vibration responses (Y1, Y2, and Y3). |
|-----------------------------------------------|-----------------|----------|----------|----------|----------|----------|----------|----------|</p>
<table>
<thead>
<tr>
<th><strong>Response</strong></th>
<th><strong>Correct detections (%)</strong></th>
<th><strong>False alarms (%)</strong></th>
<th><strong>Missed damage (%)</strong></th>
<th><strong>Damage A</strong></th>
<th><strong>Damage B</strong></th>
<th><strong>Damage C</strong></th>
<th><strong>Damage D</strong></th>
<th><strong>Damage E</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Y1</td>
<td>98.7</td>
<td>0.06</td>
<td>0</td>
<td>8</td>
<td>0.12</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Y2</td>
<td>98.8</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Y3</td>
<td>97.2</td>
<td>0.18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Test strength $(\alpha, \beta) = (0.05, 0.01)$; Residual standard deviation ratio $q = \sigma_1/\sigma_o = 1.08$.
Correct detection and false alarm rates based on 1600 healthy experiments (signal length $N = 1$ 000 samples).
Missed damage rates based on 1600 damage experiments (signal length $N = 1$ 000 samples).
Note that the correct detection and false alarm rates do not sum to 100% as in certain cases no decision is made.


as well as the mean missed damage rates which are zero, except for damage type B which exhibits an increased number of missed damage rate when the response Y2 is used.

**Damage identification (classification) results.** Indicative damage identification results based on the multihypothesis SPRT at the $\alpha_{ij} = 0.001$ error probabilities level are presented in Figure 4 for vibration response Y1. As it may be observed the method is capable of accurately identifying the actual damage type. Summary damage identification results are presented in Table 4. As it may be observed the obtained damage classification results are very accurate for damage types A, B, D and E for at least two vibration measurement positions, as the rates of correct classification are very high. Nevertheless, difficulties are faced in conjunction with damage type C when either the response point Y1 or Y3 is used.

Finally, Table 5 presents damage detection and identification summary results for the non-parametric PSD based method using 46 080 samples (90 s). It is evident (Tables 3 and 4) that the sequential method is able to improve the early damage detection and identification capability by using at maximum 1 000 samples (1.95 s).

5. **CONCLUDING REMARKS**

A model residual based SPRT framework for structural damage diagnosis was introduced. In this framework damage detection and identification were effectively tackled, achieving high performance with practically zero false alarms and missed damage rates. An optimal sampling plan was determined *a priori* via the use of the Operating Characteristic (OC) and Average Sample Number (ASN) functions, selected type I (false alarm) and II (missed damage) error probabilities, and available baseline data records of the structure under various potential states. *Early* (needing at maximum 1.95 s) and robust damage detection were achieved, and “local” and “remote” damage with respect to the sensor position was detected. The multihypothesis test based damage identification procedure faced some difficulties in classifying one damage type, an issue that may be tackled via the baseline modeling of the specific damage type followed by sequential binary hypothesis testing.

<table>
<thead>
<tr>
<th>Table 4 Damage identification summary results for the three vibration responses (Y1, Y2, and Y3).</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Actual</strong></td>
</tr>
<tr>
<td>damage</td>
</tr>
<tr>
<td>Type A</td>
</tr>
<tr>
<td>Type B</td>
</tr>
<tr>
<td>Type C</td>
</tr>
<tr>
<td>Type D</td>
</tr>
<tr>
<td>Type E</td>
</tr>
</tbody>
</table>

Acceptance rates for each one of Y1/Y2/Y3 based on 1600 experiments (signal length $N = 1\,000$ samples).
Table 5 PSD method – damage detection and identification summary results (responses Y1, Y2 and Y3).

<table>
<thead>
<tr>
<th></th>
<th>Damage Detection ($\alpha = 10^{-5}$)</th>
<th>Damage Identification ($\alpha = 10^{-5}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>False alarms (%)</td>
<td>damage A</td>
<td>damage B</td>
</tr>
<tr>
<td>0/0/0</td>
<td>0/0/0</td>
<td>47.5/0/0</td>
</tr>
<tr>
<td>Damage classification success rate (%)</td>
<td>damage A</td>
<td>damage B</td>
</tr>
<tr>
<td>100/87.5/95</td>
<td>97.5/100/97.5</td>
<td>100/75/47.5</td>
</tr>
</tbody>
</table>

False alarm, missed damage and damage misclassification rates for each measurement point (Y1/Y2/Y3) based on 40 experiments (signal length $N = 46,080$ samples).

REFERENCES


