M.S. THESIS

Damage Detection of Structures by Autoregressive Model and Threshold Value Determination

BY

Muhammad Tahir Bashir

August 2010

DEPARTMENT OF CIVIL AND ENVIRONMENTAL ENGINEERING
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

공학석사학위논문

자기회귀 모형을 이용한 구조물 손상 탐지 기법의 임계값 판별

Damage Detection of Structures by Autoregressive Model
and Threshold Value Determination

2010 년 8 월

서울대학교 대학원

건설환경 공학부

Muhammad Tahir Bashir

Damage Detection of Structures by Autoregressive Model and Threshold Value Detrmination

Advisor:
Professor Hae-Sung Lee

Thesis Dissertation by:

Muhammad Tahir Bashir

August 2010

College of Civil and Environmental Engineering

Seoul National University

Submitted to Seoul National University in fulfillment of the requirements for the degree of Master of Science

Approved by:

CHAIRMAN:	**Professor Hyun-Moo Koh**
VICE CHAIRMAN:	**Professor Hae-Sung Lee**
MEMBER:	**Professor Ho-Kyung Kim**

Abstract

Muhammad Tahir Bashir

Department of Civil and Environmental Engineering

Seoul National University

A structure can observe severe loading conditions during its life span which can cause damage to the structure. It has become very important to study the behavior of structures under normal and abnormal conditions to monitor the health of the structure. There is a lot of research going on the damage identification, localization and assessment. Autoregressive model is used to detect the damage in the structure. Covariance between autoregressive coefficients and residual of measured and calculated acceleration is proposed as a damage feature.

Non Modal based Scheme is used for structural health monitoring and measured acceleration data from the sensors is utilized here. Initially data is in time domain which is then converted to frequency domain by applying the transfer function. Non Causal filter is designed to filter the lower frequencies which are resulted by perturbations, operational and over loading conditions of traffic.

As we have the measured signals from the structure, which contains some noise and perturbation resulted by environmental effects, so to remove those perturbations time Windowing Technique is employed. Because those environmental changes occur in long duration of time, while the time window size is so small that, it remains constant in each time window.

An algorithm has been developed to find the damage feature in real time. Two span truss bridge is analyzed for damage detection; a damage scenario is created by reducing the area of the truss members. Loading of the truck, car and bus is applied normally and then increasing the loading to study the damage. Algorithm captured both the damage location and timing.

Extreme value distribution is utilized for an accurate selection of outliers because extreme value distribution is well established for tail distribution. Threshold value based on optimal sample size and significance level is determined. Some mistakes in previous study to determine the threshold value are also highlighted in this paper.

A simplified approach is adopted to get the optimal sample size against each significance level. Observation and monitoring time is also studied to see the effects on the threshold value. Limited available data cannot predict the actual situation of the bridges, which are designed for hundred years. So at least one year observation period should be taken to predict the behavior of bridge which is designed for a century.

Keywords: Damage Detection, Time window, Autoregressive Model, Damage Feature, Non Causal Filter, Threshold value, Extreme value distribution.

Student Number: 2008-23538

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Problem Statement

A structure can observe a severe loading condition in its life span which can cause damage to the structure. It has become important to study the behavior of structures under normal and abnormal conditions to monitor the health of the structure. Significant work has been done in the area of detecting damage in the structure and its health monitoring by using the changes in the dynamic response of structure. Because natural frequency and mode shapes of structure are dependent on the mass and stiffness distributions, many subsequent changes in them should theoretically, be reflected in frequency and mode shape of structure.

Structural damage may be defined as any deviation of geometric or material property defining a structure that may result in unwanted response of the structure. A solution to this problem is important for at least two reasons, firstly damage localization and severity estimation are first two steps in the broader category of damage detection, secondly, a timely damage assessment could produce desirable consequences such as saving of lives, reduction of human suffering, protection of property, increased reliability, increased productivity of operation and reduction in maintenance cost.

Despite these research efforts, however, many problems related to vibration based damage detection remain unresolved today. Outstanding need remain to locate and estimate the severity of damage in the following cases

a) In structures with only few available modes
b) In structures with many members
c) In structures for which baseline modal responses are not available and

d) In environment of uncertainty associated with modeling, measurement and processing errors.

Many designed structures don’t consider the unexpected severe loading and structural defects in material which can result the damage of the structure during its use. Some times improper intention to local damages or improper maintenance can also cause failure of the structure. So it is important to constantly monitor the structure and pay the attentions to the changes while its service life.

Until now the structural damage was accessed by visual inspection by experts or some non destructive methods were used, but due to emergence and development of new technologies some automated damage assessment techniques and systems are becoming more popular. It also decreases the maintenance cost and especially for the cable supported bridges it is not possible physically to monitor each component of structure. So automated systems are being developed and employed to monitor the health and behavior of the structures.

Basically for damage detection of the structure there are two modeling techniques

(1) Structural Model Based Scheme (SMBS)

(2) Non -Structural Model Based Scheme (NMBS)

In structural model based scheme stiffness, damping and mass information are used and system parameters are estimated by inverse analysis based on the sensitivity method from a mathematical model. Structural model based system is a type of inverse problems, which are usually ill posed problem. An ill-posed problem is characterized by the non-uniqueness, non- continuous and instability solutions. Various regularization techniques have been developed to eliminate this ill-posedness of inverse problem. In spite of this, ill-posedness can be alleviated by regularization techniques successfully, model based system identification schemes are not applied in real situation because of modeling error i.e. Difference between mathematical model and real structure. In regularization technique an additional constraint on the system parameters is imposed which is referred to as regularization function, which is imposed to the original minimization problem defined by the error function.. It is very important to define a proper regularization function that is able to describe characteristics of a SI problem in hand. Recently, a lot of structural health monitoring algorithms with statistical pattern recognition using purely measured signals has been attempted in center of Los Almos National Laboratory in USA. System identification is further divided into time domain and frequency domain. Mostly structural dynamic behavior is measured in the form of acceleration in time domain and then it is transformed to frequency domain. Due to the ease in handling the frequency domain SI algorithms are developed and applied. Actually both approaches have some merits and de merits and limitations.

Non structural model based health monitoring consists of four steps which is Data acquisition from sensors of structure, data transmission, data analysis using measured data for damage detection and decision making, whether the considered structure is sound or not. In spite of rapid progress of sensor and IT technology, rigorous damage detection algorithm did not be proposed. Therefore so many measured signals obtained from various structures cannot be used for structural health monitoring. This study is based on non model based damage detection.

The autoregressive model is utilized to study the non structural model based problem. Autoregressive model was used by Yule for the first time to explain the changes and periods of sunspot from simple trigonometric identity. Autoregressive model can be extremely useful in the representation of certain practically occurring time series. In this model current value of the process is expressed as a finite, linear aggregate of previous values and shock.

Various non structural models based structural health monitoring algorithms using static or dynamic responses have been proposed. But the main problem of structural monitoring is how to handle noises in measured signals, whereas measured signals contain a mix of information related to both the damage in the structure and the perturbations due to the environmental changes. The autoregressive modal with time windowing technique is employed to overcome the perturbations of measured signals. Auto regressive model with self assessment of the general process of structural soundness is shown in fig 1.1. First the sensor is placed in the target structure and data is acquired in time series. Measured data may include displacement, velocity, acceleration and strain etc but here we are only using acceleration data. That acceleration data contains the information about the structural behavior which is analyzed by algorithm developed on autoregressive technique and damage indication are made by setting the level of significance.

As mentioned before that measured acceleration data may contain some white noise which should have to be a parted from the true response of the structure. It is noted that the environmental effects are related to some time lapse, so a time windowing technique is employed. In this technique the auto regressive model is estimated sequentially using measured data within a finite time period. Time window advances forward at each time step to update autoregressive model repeatedly. Perturbation due to environment changes, take places gradually during a long time period and time window size is relatively smaller so environmental changes can be neglected within time window.

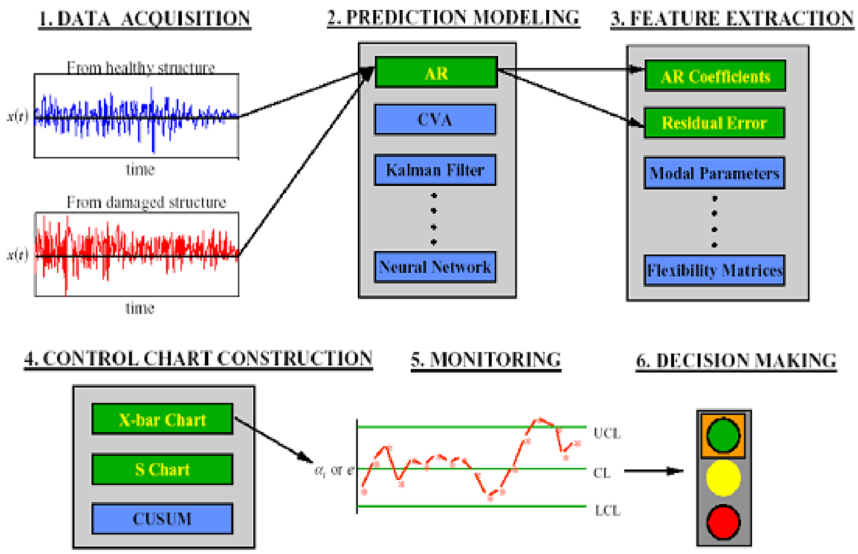


Figure 1.1 Autoregressive Model Based Damage Detection Scheme

Making decision whether the considered structure is sound or not using damage features from each sensor in every time step is also very important. The extreme value distribution is utilized to detect outliers because damage information almost lies in the tail of distribution and extreme value is well established for tail distribution. Threshold value which

1.2 Previous Research

Previous research on the damage detection by autoregressive model was done by kang [PhD thesis Feb 2008] and Park [MS thesis 2010] their approach to the topic is mentioned below

1- Causal filter in time domain

Kang done his study on the damage detection by measured acceleration data in time domain and he utilized the causal filter for damage detection. In his study he utilized the previous value of acceleration data to predict present value He used covariance between the autoregressive coefficients and residual between measured and calculated acceleration as a damage feature. He worked to get some reasonable sample size which can be used in real time monitoring and also give prominent damage feature.

2- Non Causal filter in frequency domain.

Park used the same damage feature i.e. covariance between autoregressive coefficients and residuals but he used the transfer function to convert the kang proposed damage feature in frequency domain. In non causal filter present value of the function is determined by previous and future value of the measured data. He improved the damage feature and damage instant from where one can get the location of the damage. He suggested the sample size of 600 for damage feature.

1.3 Goal and objective

In non structural model based damage detection technique, the biggest problem is to handle the noise in measured data. The causes of noise may be due to temperature or environmental factors, error from measurement of data and sensors functionality problem. Many researchers are finding it difficult to isolate the wrong signals sent by noise and they are providing better techniques to tackle this issue. Although there is a rapid development and improvement of sensor technology, yet the presence of noise in the measured signals cannot be removed perfectly.

Autoregressive model in time domain and frequency domain will be used to study the damage feature and similarly Causal and non casual filter will be studied and comparison of the results will be made to follow the unique technique. Time windowing technique will be used to nullify the effect of noise.

So this study is mainly related to autoregressive model base damage detection and finding the threshold value and try to establish some decision making about the damage by studying the threshold value. There is also some discussion about observation and monitoring time and false alarm which occurs due to non damaged signals.

1.4 Organization of the Thesis

In chapter 1, introduction about the topic and the previous studies is given and the goals and objectives for my studies are set. In chapter 2 time series analysis, autoregressive model, impulse response function (transfer function), covariance as a damage feature is discussed and mathematical modeling is done. In chapter 3 frequency domain autoregressive model is studied. Formulation is developed for design of non causal filter and damage feature is improved in frequency domain.

In chapter 4 extreme value theory and extreme value distribution is studied to get the threshold value and decision making about damage or non damage of the structure. In chapter 5 damage detection algorithm is applied to 2 span bridge and results are obtained, which shows the detection of damage and threshold value for different sensors, set at different nodes at the truss bridge. In chapter 6 conclusion about the study are derived and some discussion is made on further studies.

2 Time Series Analysis

Time series is a sequence of data points, measured at successive times spaced at uniform time intervals. time series analysis comprises methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data, by using time series analysis we can forecast future events based on known past events to predict data points before they are measured. They are basically two methods for time series analysis one is based on time domain and the other is on frequency domain. So we will study both of them to see the effect on the damage feature and other system parameters. Frequency domain includes spectral analysis and wavelet analysis while the time domain analysis include auto-correlation and cross co- relation analysis. The ultimate goal of the time series analysis is to predict the future. As shown in the figure 2.1 based on time series data from the previous pattern we can predict function for future values within the limits. To assess degree of confidence in predictions based on time series statistical distribution can be best estimated.

Estimates are made on the desired confidence level while setting the upper and lower limits on the both sides of the forecasting function. Previous time series data to identify patterns of input data and output data indicating the relationship between advancing transfer function will be used. There are different approaches to model the time series, which depends upon the nature of the problem.

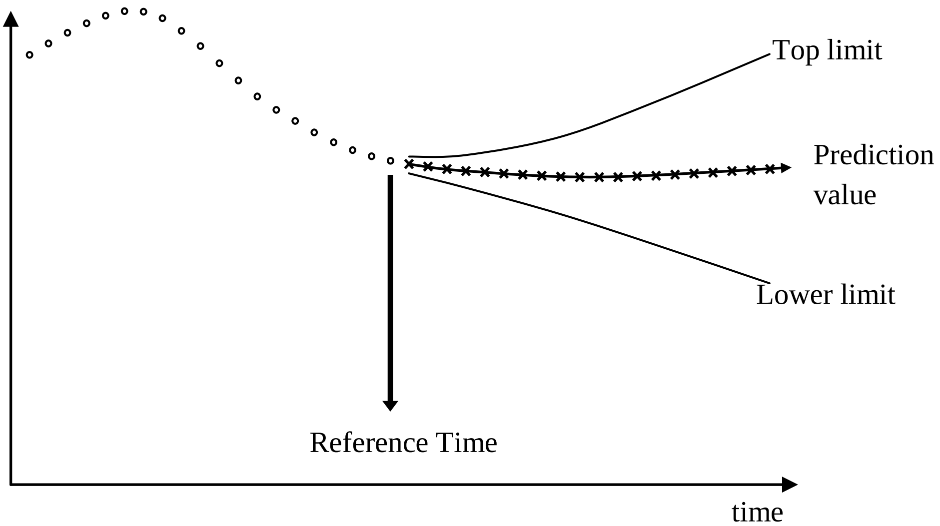


Figure 2.1.1 Prediction Function and Boundaries for Ttime Series data

There are obviously numerous reasons to record and to analyze the data of a time series. Among these is the wish to gain a better understanding of the data generating mechanism, the prediction of future values or the optimal control of a system. The characteristic property of a time series is the fact that the data are not generated independendetly, their dispersion varies in time, they are often governed by a trend and they have cyclic components. Statistical procedures that suppose independent and identically distributed data are, therefore, excluded from the analysis of time series. This requires proper methods that are summarized under time series analysis.

2.1 Autoregressive Model (AR Model)

The major episode in the history of development of time series analysis took place in the time domain, and it began with the two articles of 1927 by Yule and Slutsky. In both articles, we find a rejection of the model with deterministic harmonic components in favor of models more firmly rooted in the notion of the random causes. In a wonderfully fugrative exposition , Yule invited his readers to imagine a pendulum attached to a recording device and left to swing .Then any deviations from perfectly harmonic motion which might be recorded must be the result of errors of observations which could be all but eominated if a long sequence of observations were subjected to a peridogram analysis . Next, Yule enjoined the reader to imagine that the regular swing of pendulum is interrupted by small boys who get into the room and start pelting the pendulum with peas sometimes from one side and sometimes from the other. The motion is now affected not by superposed fluctuations but by true disturbances.

Simple trigonometric functions such as expression 2.1 and the autoregressive model which Yule was proposing takes the form as shown in equation 2.2.

$$\sin(kx) = 2\cos(x)\sin(k-1)x - \sin(k-2)x \tag{2.1}$$

$$y(k) = a\,y(k-1) - y(k-2) + \varepsilon(k) \tag{2.2}$$

Where $\varepsilon(k)$ is a white noise sequence. Now, instead of making the regular periodicity of the pendulum, the white noise has actually become the engine which drives the pendulum by striking it randomly in one direction and another. Auto regressive model for the time series data and statistical analysis is the most popular model, in which the combination of current and past value at a time step is represented as a white noise

2.2 Formula for Autoregressive Model

Regular time interval listed in the time series data for each time $t, t-1, \ldots\ldots t-p$ corresponding to measurement data *y*(*t*), *y*(*t*-1), …, *y*(*t*-*p*). Autoregressive model of order p can be written as

$$y(t) = a_1 y(t-1) + a_2 y(t-2) + \cdots + a_p y(t-p) + e(t) \tag{2.3}$$

Here $a_1, a_2, \cdots, a_p$ are the coefficients of autoregressive model and $e(t)$ is a white noise

$$[y(t) - \mu] = a_1 [y(t-1) - \mu] + a_2 [y(t-2) - \mu] + \cdots + a_p [y(t-p) - \mu] + e(t) \tag{2.4}$$

Backward shift operator

$$q^{-n} y(t) = y(t-n) \tag{2.5}$$

Operational polynomial is given by Equation 2.6

$$A(q) = 1 - a_1 q^{-1} - a_2 q^{-2} - \cdots - a_p q^{-p} \tag{2.6}$$

Operator expression is given by

$$A(q) y(t) = e(t) \tag{2.7}$$

2.2.1 Impulse Response Function and transfer Function

Let $u(t)$ be an input signal and $y(t)$ be the output signal then the impulse response function can be defined as

$$y(t) = \int_{\tau=0}^{\infty} g(\tau) u(t-\tau) d\tau \tag{2.8}$$

$$y(t) = \sum_{k=1}^{\infty} g(k) u(t-k) \quad , \quad t = 0, 1, 2, \ldots \tag{2.9}$$

$$y(t) = \sum_{k=1}^{\infty} g(k)u(t-k) = \sum_{k=1}^{\infty} g(k)\left(q^{-k}u(t)\right)$$

$$= \left[\sum_{k=1}^{\infty} g(k)q^{-k}\right]u(t) = G(q)u(t) \tag{2.10}$$

Where $G(q)$ is called the transfer function $v(t)$

$$v(t) = \sum_{k=0}^{\infty} h(k)e(t-k) \tag{2.11}$$

Where $e(t)$ at time t represents the white noise. Expression 2.11 when applied to the output data, it can be expressed in the form of transfer function as follows

$$v(t) = \sum_{k=0}^{\infty} h(k)q^{-k}e(t) = H(q)e(t) \tag{2.12}$$

Finally the expression of transfer function 2.10, 2.12 can be written as follows

$$y(t) = G(q)u(t) + H(q)e(t) \tag{2.13}$$

2.2.2 Transfer Function Model

More generalized transfer function can be expressed as

$$A(q)y(t) = \frac{B(q)}{F(q)}u(t) + \frac{C(q)}{D(q)}e(t) \tag{2.14}$$

Equation 2.14 is the expression for transfer function model that can be generated by the model number of regular expressions A, B, C, and D depending upon the combination of five transfer functions which can generate 32 branches which are mentioned in table 2.1. The model shown in table 2.1 indicates the weak among the auto regressive process, x is the external input component (exogenous input) and MA indicates moving average process. ARMA is the Autoregressive moving average model it is also called box-jenkins models and are applied to autocorelated time series. Given a time series of data Xt, the ARMA model is a tool for understanding and perhaps, predicting future in this series? The model consists of two parts autoregressive (AR) part and moving average (MA) part and is written as ARMA(pq). Where p is the order of autoregressive part and q is the order of the moving average part.

Table 2.1 commonly used transfer function models

Used polynomials	Name of model structure
A	AR
AB	ARX
ABD	ARMAX
AC	ARMA
ABD	ARARX
ABCD	ARARMAX
B	FIR (Finite Impulse Response)
BF	OE (Output Error)
BFCD	BJ (Box-Jenkins)

2.2.3 Order of Autoregressive Model

To determine the order of the autoregressive model is the top priority issue. If the order for autoregressive model is small then the model and data can not be represented properly. If the order of the autoregressive model is high then it

Cana represents the data adequately but gives numerically unstable results. Therefore optimal order of autoregressive model adequately represents the measured data and model changes.

While selecting the order of the AR model we should find that order which has the minimum residual error between measured and calculated values i.e. the minimum value of residual variance is the best way to determine the order. However variance of the model generally increases the order of the model which cannot give optimal value for the AR order.

To determine the order of AR model the most commonly used method is auto correlation function or partial auto correlation function. These functions especially in time series models are useful for identification and diagnosis. Continuous Random Variables in time series X_{t-2}, X_{t-1}, X

$$\rho_{X_{t-2}X_{t-1}}, \rho_{X_{t-1}X_t}, \rho_{X_{t-2}X_t} \tag{2.15}$$

$$X_{t+h} = \phi_{h1}X_{t+h-1} + \phi_{h2}X_{t+h-2} + \cdots + \phi_{hh}X_t + \varepsilon_{t+h} \tag{2.16}$$

$$\rho_X(k) = \phi_{h1}\rho_X(k-1) + \phi_{h2}\rho_X(k-2) + \cdots + \phi_{hh}\rho_X(k-h) \tag{2.17}$$

$$\varphi_{hh} = \frac{\begin{vmatrix} 1 & \rho_X(1) & \rho_X(2) & \cdots & \rho_X(h-2) & \rho_X(1) \\ \rho_X(1) & 1 & \rho_X(1) & \cdots & \rho_X(h-3) & \rho_X(2) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_X(h-1) & \rho_X(h-2) & \rho_X(h-3) & \cdots & \rho_X(1) & \rho_X(h) \end{vmatrix}}{\begin{vmatrix} 1 & \rho_X(1) & \rho_X(2) & \cdots & \rho_X(h-2) & \rho_X(h-1) \\ \rho_X(1) & 1 & \rho_X(1) & \cdots & \rho_X(h-3) & \rho_X(h-2) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_X(h-1) & \rho_X(h-2) & \rho_X(h-3) & \cdots & \rho_X(1) & 1 \end{vmatrix}} \tag{2.19}$$

To determine the optimal order of autoregressive model many researchers have proposed various criteria like Akaike final prediction error (FPE) (Akaike 1969), Akaike information criteria (AIC) (Akaike ,1974) suggested two ways. AIC is a number of measurement data Rissanen AIAc is the number of measurement data, where N is the infinitely large residual variance may occur because of 0 , and is not a problem to solve a numerical minimum description length (Rissnen,1983) proposed method

$$FPE[p] = \hat{\sigma}_p \left(\frac{N + (p+1)}{N - (p+1)} \right) \tag{2.20}$$

$$AIC[p] = N \ln \hat{\sigma}_p{}^2 + 2p \tag{2.21}$$

$$MDL[p] = N \ln \hat{\sigma}_p{}^2 + p \ln N \tag{2.22}$$

Here $\hat{\sigma}_p$, N, p are the residuals , maximum likelihood estimator of the variance, the number of measurement data and order of autoregressive model respectively Figure 2.2 is the typical example of correlation function, which shows the partial correlation and threshold values for 2 span continuous truss structure under the Kobe ground earthquake data. Partial correlation function which is normally distributed is shown and threshold values are set at 5% significance level.

Fig 2.3 shows the graph between Ardor and RMS of residuals, from here we can see that increasing the order of autoregressive model the RMS values of residual values decreases sharply unto order 2 and becomes stable for higher values so optimal order can be selected , for example 4 is suitable for the above case.

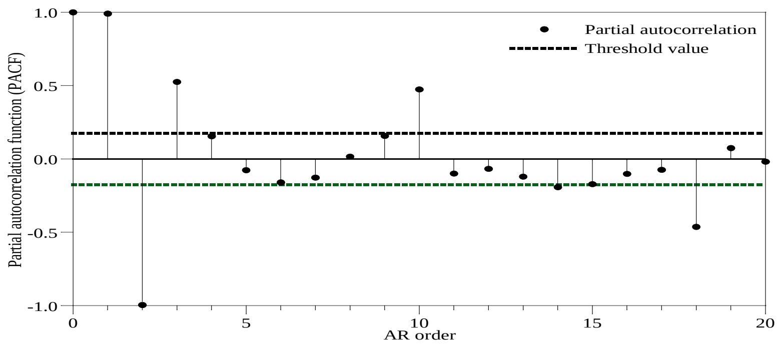


Figure 2.2 Partial Correlation Function

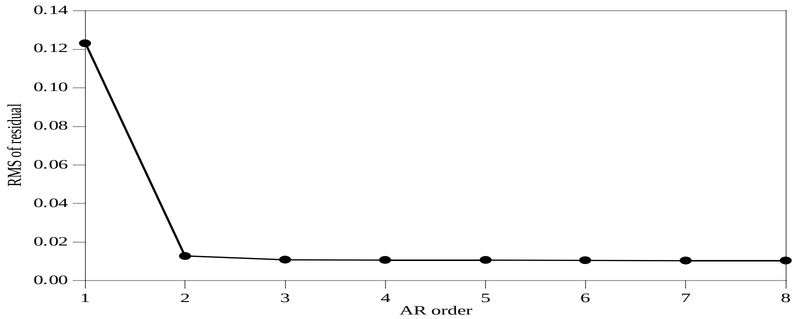


Figure 2.3 Residual Mean Square

2.2.4 Estimation of Autoregressive coefficients

There are several methods to calculate coefficients of autoregressive model which are least square d method, moment method, maximum likelihood, Bayesian theory and so on. Least square method is used here because it is very simple and clear. Lets the values of the process at equally spaced times t, t-1, t-2 .. t-p be y(t), y(t-1), y(t-2).......y(t-p). The prediction value from autoregressive model of order p can be defined as given below

$$\begin{aligned} \hat{y}(t \backslash \boldsymbol{\theta}) &= [1 - A(q)]y(t) \\ &= a_1 y(t-1) + a_2 y(t-2) + \cdots + a_p y(t-p) \\ &= \boldsymbol{\theta}^T \boldsymbol{\phi}(t) = \boldsymbol{\phi}^T(t)\boldsymbol{\theta} \end{aligned} \tag{2.23}$$

Where $\hat{y}$ represents prediction value from autoregressive model, $\boldsymbol{\theta}$ represents system parameter vector and $\boldsymbol{\phi}$ represents regression vector respectively.

$$\begin{aligned} \boldsymbol{\theta} &= [a_1 \quad a_2 \quad \dots \quad a_p]^T \\ \boldsymbol{\phi}(t) &= [y(t-1) \quad y(t-2) \quad \dots \quad y(t-p)]^T \end{aligned} \tag{2.24}$$

The residuals can be defined as the difference between measured signals and prediction values using autoregressive model at each time step

$$e(t) = y(t) - \boldsymbol{\phi}^T(t)\boldsymbol{\theta} \tag{2.25}$$

The linear object function by least squared method can be obtained as shown below

$$V(\boldsymbol{\theta}) = \frac{1}{2}\sum_{k=p+1}^{N}\left[y(k) - \boldsymbol{\phi}^T(k)\boldsymbol{\theta}\right]^2 \tag{2.26}$$

Where N represents the total number of measured signals. The optimal solution of equation 2.26can be obtained by least squared method as shown below

$$\hat{\boldsymbol{\theta}}_N = \left[\sum_{k=p+1}^{N}\boldsymbol{\phi}^T(k)\boldsymbol{\phi}(k)\right]^{-1}\sum_{k=p+1}^{N}\boldsymbol{\phi}^T(k)y(k) \tag{2.27}$$

The above equation can be written as $\mathbf{R}(N)$ which is a PxP matrix and $\mathbf{f}(N)$ which is vector of order p.

$$\mathbf{R}(N) = \sum_{k=p+1}^{N}\boldsymbol{\phi}^T(k)\boldsymbol{\phi}(k) \tag{2.28}$$

$$\mathbf{f}(N) = \sum_{k=p+1}^{N}\boldsymbol{\phi}^T(k)y(k) \tag{2.29}$$

$$\hat{\boldsymbol{\theta}}_N = \mathbf{R}(N)^{-1}\mathbf{f}(N) \tag{2.30}$$

Multiplying equation 2.3 by $y(t\text{-}k)$ on both sides we get

$$\begin{aligned} y(t-k)y(t) = a_1 y(t-k)y(t-1) + a_2 y(t-k)y(t-2) + \cdots \\ + a_p y(t-k)y(t-p) + y(t-k)e(t) \end{aligned} \tag{2.31}$$

$$\begin{aligned} \gamma(k) = a_1\gamma(k-1) + a_2\gamma(k-2 + \cdots) \\ + a_p\gamma(k-p) + E\left[y(t-k)e(t)\right] \qquad k > 0 \end{aligned} \tag{2.32}$$

$$\gamma(k) = E\left[y(t)y(t+k)\right] = \frac{1}{N-k}\sum_{t=1}^{N-k} y(t)y(t+k) \tag{2.33}$$

$$\rho(k) = a_1\rho(k-1) + a_2\rho(k-2) + \cdots + a_p\rho(k-p) \qquad k > 0 \tag{2.34}$$

For k=1,2,3……..,p putting the value in equation 2.35 we get (2.35)

$$
\begin{array}{ccccccccc}
\rho_1 & = & a_1 & + & a_2\rho_1 & + & \cdots & + & a_p\rho_{p-1} \\
\rho_2 & = & a_1\rho_1 & + & a_2 & + & \cdots & + & a_p\rho_{p-2} \\
\vdots & & \vdots & & \vdots & & \ddots & & \vdots \\
\rho_p & = & a_1\rho_{p-1} & + & a_2\rho_{p-2} & + & \cdots & + & a_p
\end{array}
\tag{2.36}
$$

Equation 2.36 is a Yule- walker equation from there we can get the coefficient of autoregressive model

$$\boldsymbol{\theta} = \mathbf{P}^{-1}\boldsymbol{\rho} \tag{2.37}$$

Where $\boldsymbol{\theta}$, $\boldsymbol{\rho}$, $\mathbf{P}$ are given by the following expressions

$$\boldsymbol{\theta} = \begin{bmatrix} a_1 & a_2 & \cdots & a_p \end{bmatrix}^T, \ \boldsymbol{\rho} = \begin{bmatrix} \rho_1 & \rho_2 & \cdots & \rho_p \end{bmatrix}^T \tag{2.38}$$

$$\mathbf{P} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \cdots & 1 \end{bmatrix} \tag{2.39}$$

This method is suitable for the sufficient amount of data and but generally moment method is difficult to use on a small data. In this paper least square method is used to estimate the parameter of autoregressive model.

2.3 Time Window Technique

In non structural model based damage detection scheme the main focus is on the reliability of the data obtained from the measurement equipments such as accelerometer and sensors. The key difficulty in structural health monitoring is perturbation of measured signals by unknown effects such as environmental and instrumental effects. Measurement errors can be reduced by improvement in the sensor technology but perturbation of environment cannot be reduced by this way. The change in measured signals occurs gradually due to environmental effect such as temperature effect. Even if there is no damage in the considered structure, measured signals can be swayed by environment. Every previous method suffered from this difficulty of environmental data in laboratory.

A time window technique is utilized to overcome this problem. Environmental perturbations are commonly changed during relatively long time period. In autoregressive model with time window technique, the autoregressive model is estimated sequentially within a finite time window period. The time window overlaps and advances forward at each time step to update autoregressive model. Time window size is very small as compared to the period of environmental changes so it can be assumed that perturbations of measured signals from environment within the time window cannot be happened.

In autoregressive model with time windowing technique , the autoregressive model will be estimated sequentially not using all of the signals but measured signals within a finite time period which is called time window as shown in figure 2.4.

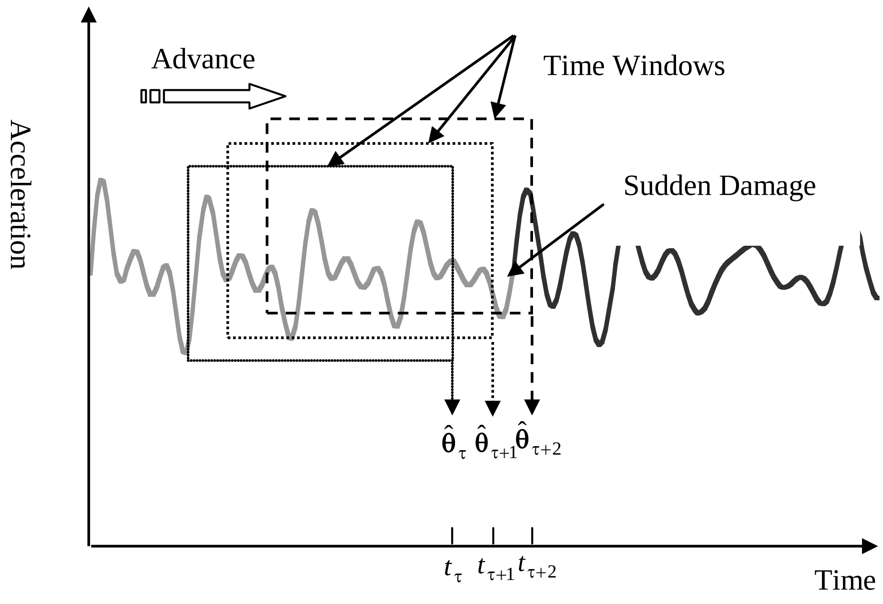


Figure 2.4 Outline of Time Window Technique

2.3.1 Application of Time Window Technique

Time windows as shown in the fig 2.4 are set at each time step and data points are included in each time window. Time window contains the time data at equal time steps. Number of data points in each time window shows the size of time window. It is also very important to select the optimal size of time window, as small time window will take more calculation time, while longer time window cannot remove the perturbation of environmental effects. So time window moves each time step and predicts the value for next step. So at each time step new autoregressive coefficients are predicted as time window moves ahead which are $\cdots, \hat{\boldsymbol{\theta}}_{\tau}\ , \hat{\boldsymbol{\theta}}_{\tau+1}\ , \hat{\boldsymbol{\theta}}_{\tau+2}\ , \cdots$.

From equation 2.39 we have the object function for the autoregressive model i.e.

$$V(\boldsymbol{\theta}) = \frac{1}{2} \sum_{k=(t-nw)+p+1}^{t} \left[y(k) - \boldsymbol{\phi}^T(k)\boldsymbol{\theta} \right]^2 \tag{2.39}$$

Where *nw* denotes the time window in this time windowing technique which is also called causal filter, we predict present value of the autoregressive coefficients from the previous values of time data

$$\hat{\boldsymbol{\theta}}_t = \left[\sum_{k=(t-nw)+p+1}^{t} \boldsymbol{\phi}^T(k)\boldsymbol{\phi}(k) \right]^{-1} \sum_{k=(t-nw)+p+1}^{t} \boldsymbol{\phi}^T(k) y(k) \tag{2.40}$$

So in autoregressive model with time windowing technique the present values of the system is determined when the time window moves ahead, so this procedure is repeated for the whole data. In this study time window of size hundred is selected which has the size of 2.5seconds, as the data used in this study is sampled at the frequency of 400 Hz.

$$e(t+1) = y(t+1) - \hat{y}(t+1 \mid \hat{\boldsymbol{\theta}}_t) \tag{2.41}$$

2.3.2 Time Window size and effect on damage feature

Unstable AR coefficients are obtained by the minimization of LSE, so for removing the instability and noise from the data it's very important to select the optimum size of time window and small amount of data. So therefore for the stability and sensitivity to the changes, appropriate time window size is determined. In the figure 2.5, 2.6, 2.7 damage features for a two span truss structure is shown which is under the Kobe earthquake loading as shown below. We have used different time window size and we can see while comparing the damage feature that, increasing the time window size

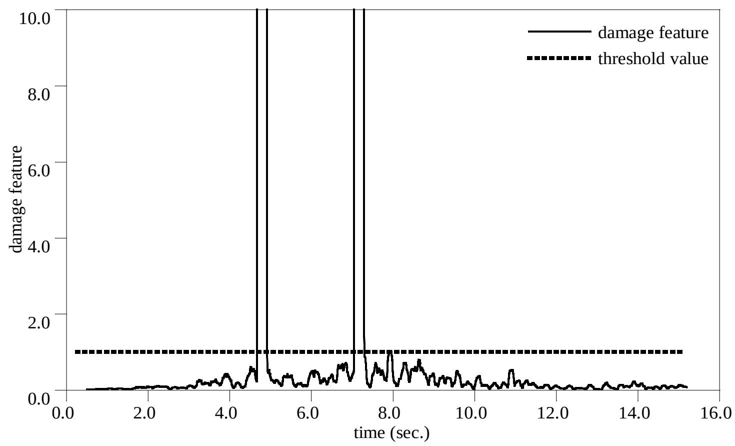


Figure 2.5 Damage Feature for the Truss under Kobe earthquake loading (tw=50)

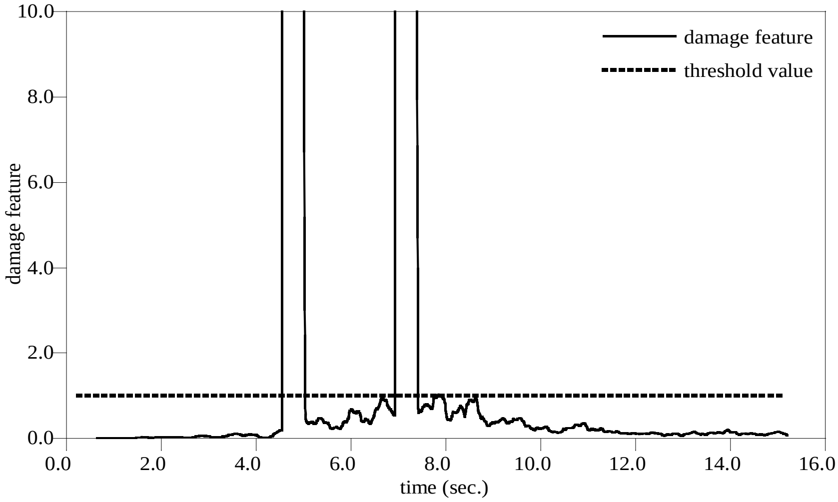


Figure 2.6 Damage Feature for the Truss under Kobe Earthquake loading (tw=100)

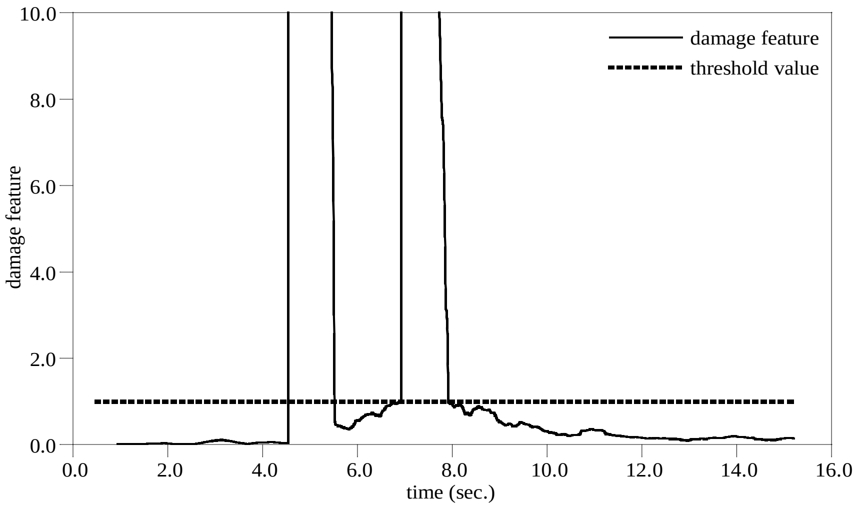


Figure 2.7 Damage feature for the truss under Kobe earthquake loading (For time window size=200)

Gives the distinct damage feature and we can reduce the white noise from the given signals. From this we can see that increasing window size, reduces the sensitivity of damage feature to perturbations and we can separate white noise from damage feature.

2.4 Regularization Technique

The autoregressive coefficients which are estimated by the minimization of least squared errors are extremely unstable. Since the number of measured signals within time window cannot be increased, regularization technique must be adopted to alleviate instability of autoregressive coefficients. The regularized least square estimator is shown in equation 2.42. The regularization function is added to the error function to overcome ill-posedness of inverse problems.

$$\Pi_R = \frac{\beta}{2}\left\|\boldsymbol{\theta}_t - \boldsymbol{\theta}_t^*\right\|_2^2 \tag{2.43}$$

Where $\beta, \boldsymbol{\theta}_t, \boldsymbol{\theta}_t^*$ are regularization factor, autoregressive coefficient and mean value of autoregressive coefficient respectively.

$$V(\boldsymbol{\theta}) = \frac{1}{2}\sum_{k=(t-nw)+p+1}^{t}\left[y(k) - \boldsymbol{\phi}^T(k)\boldsymbol{\theta}\right]^2 + \Pi_R + \frac{\beta}{2}\left\|\boldsymbol{\theta}_t - \boldsymbol{\theta}_t^*\right\|_2^2 \tag{2.43}$$

The regularization factor has critical effect on the stability of the solution. The optimal regularization factor is determined by the geometric mean scheme (gms) as given below

$$\beta = \sqrt{S_{\min} \cdot S_{\max}} \tag{2.44}$$

Where S is a singular value obtained from singular value decomposition of system matrix $S_{\min}, S_{\max}$ is minimum and maximum singular value respectively. It has been proved by Kang that without proper regularization and time window technique we cannot get the distinct damage feature due to noise and instability of autoregressive coefficients. So proper regularization technique is very important to get the clear and distinct damage feature.

2.5 Damage Feature

All the damage detection models and techniques are required to detect the changes in the structural response in case of any abnormal loading or earthquakes. As non structural model based scheme is based on the vibration of the structure. So any change in both frequency and amplitude should be captured by damage detection schemes. There are two possible damage features using autoregressive model. One is residual and the other is autoregressive coefficients. Residual has good characteristic which is very stable and sensitive to both amplitude and frequency change of measured data. But the main draw back of residual is that, we cannot identify the sources of changes.

The autoregressive coefficients are system parameters of autoregressive model. They are sensitive to frequency change and insensitive to amplitude change. The main drawback of autoregressive coefficient is that they suffer ill posedness while minimization of least square error.

The covariance between residuals and autoregressive coefficients is proposed as a new damage feature as shown in equation 2.46 in order to use the information of residuals and autoregressive coefficients instantaneously. The absolute value of residuals and autoregressive coefficients are used because the directional information of damage feature is not necessary

$$\begin{aligned} D_t &= \mathrm{cov}[|e_t|, |\theta_t^1|] = E[|(e_t - \mu_{e_t})||(\theta_t^1 - \mu_{\theta_t^1})|] \\ &= \frac{1}{nw-1} \sum_{k=t-nw+1}^{t} \left[|(e_k - \mu_{e_t})||(\theta_k^1 - \mu_{\theta_t^1})| \right] \qquad t \geq 2nw \end{aligned} \tag{2.45}$$

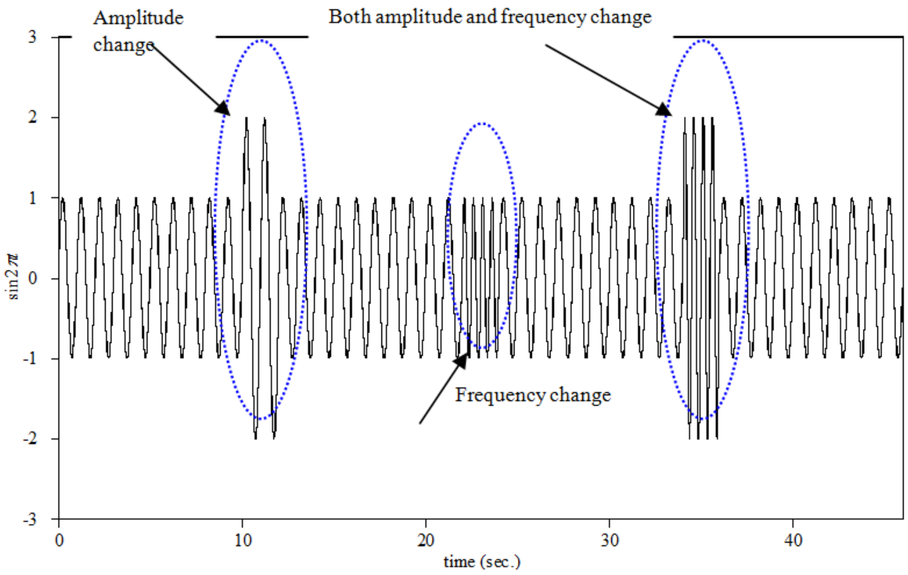


Figure 2.8 Harmonic curve during 46 seconds

In figure 2.17 a harmonic curve is shown it can be seen that there are three locations where changes have occurred. At instant one amplitude have been doubled from 1 to 2, in instant two, frequency have been changed and at third instant both frequency and amplitude have been changed. Now by using the Autoregressive model and damage features such as residual, autoregressive coefficients and covariance, we will try to detect the changes in above mentioned case.

Now as shown in the figure 2.18 that residual is sensitive to both frequency and amplitude changes, which the algorithm catches successfully, and we can see in figure 2.19, that Ist order autoregressive coefficients are sensitive to frequency changes. While covariance shown in figure 2.20 can detect the changes in both frequency and amplitude

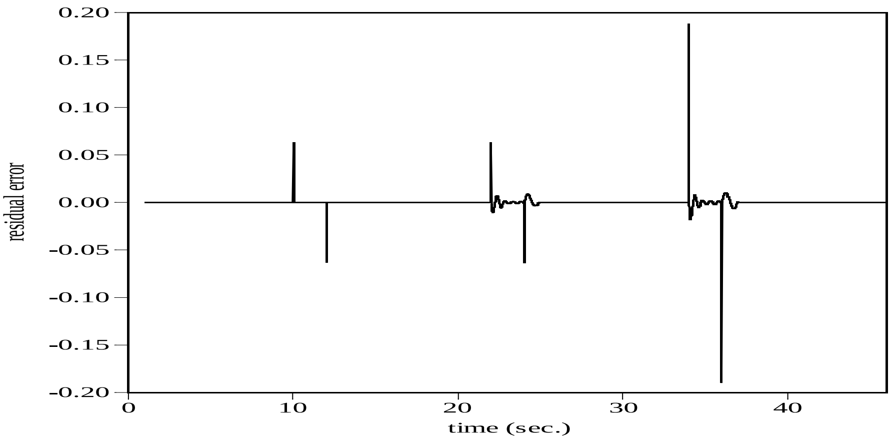


Figure 2.9 Residual as Damage Feature

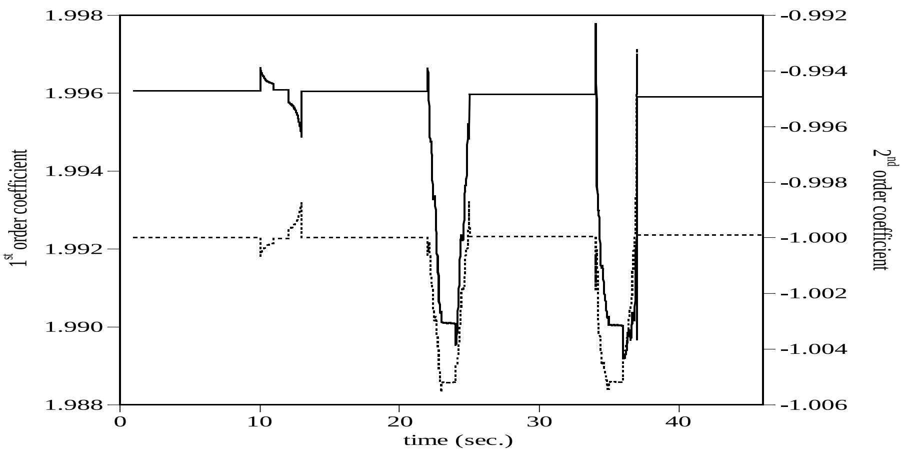


Figure 2.10 AR Coefficients as Damage Feature

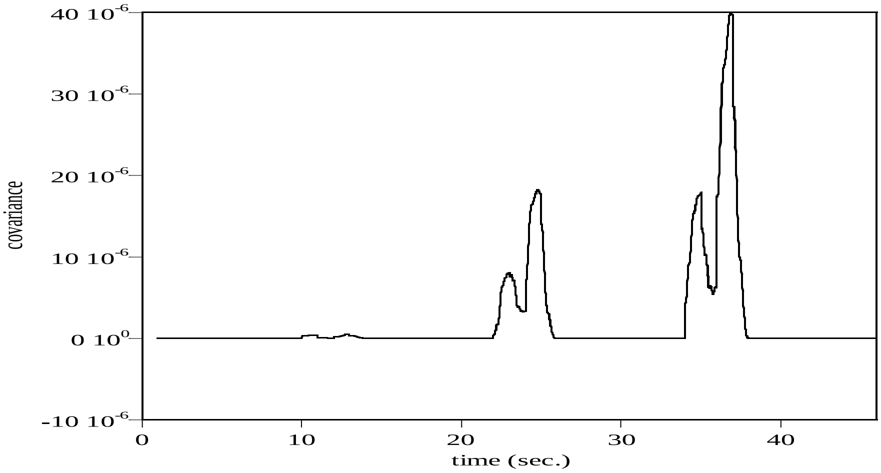


Figure 2.11 Covariance between Residual and AR Coefficients

3 Transfer Function and Filter Design

As discussed in previous chapters that damage detection can be done by using the autoregressive model in both time and frequency domain. Structures at damage instant shows high frequency value, so an algorithm can be developed easily to detect those high frequencies and decision about the damage then can be made by statistical treatment. In the previous studies. In the previous studies the high frequency effects on the damage detection was not explained clearly. So frequency domain analysis is carried out by using autoregressive model to detect the damage.

3.1 Frequency Domain Analysis.

As in previous chapter we studied the autoregressive model in time domain here we will transfer the function from time domain to frequency domain and for this purpose Fourier transforms are used

$$X(f) = \int_{-\infty}^{\infty} e^{-2\pi i f t} x(t)\, dt \tag{3.1}$$

Fourier transform is an operation that transforms one complex-valued function of a real variable into another. The domain of the original function is typically time and is accordingly called the time domain, that of the new function is frequency and so the Fourier transform is often called the frequency domain representation of the original function. So by applying the transfer function we will get the data in frequency domain and then frequency spectrum can be utilized for our damage detection purposes

3.1.1 Transfer function

Transfer function are commonly used in the analysis of single input single out put filters and it is referred to linear time invariant systems(LTI). In its simplest form for continuous time input signal $x(t)$ and out put $y(t)$, the transfer function is the linear mapping of the Laplace transform of the input $X(f)$, to the output $Y(f)$.

$$Y(f) = H(f)X(f) \tag{3.2}$$

Then the output is related tot input by the transfer function $|H(f)|$

$$|H(f)| = \frac{|Y(f)|}{|X(f)|} \tag{3.3}$$

$$\arg[H(f)] = \arg[Y(f)] - \arg[X(f)] \tag{3.4}$$

3.2 Transfer function for Autoregressive model

The autoregressive model for damage detection studied in time domain by Kang only used the past data of time invariant system. So it can be said as a causal filter as it utilized previous values to predict the present value, mathematically it can be written as follows

$$\hat{y}(\Delta t \cdot n / \theta) = a_1 y(\Delta t(n-1)) + a_2 y(\Delta t(n-2)) + \cdots + a_p y(\Delta t(n-p)) \tag{3.5}$$

By taking the Fourier transform of equation 3.5we get the transfer function of casual filter, and frequency $\omega = 2\pi f$

$$F(\hat{y}(\Delta t \cdot n / \theta)) = (a_1 e^{-i\omega\Delta t} + a_2 e^{-2i\omega\Delta t} + \cdots + a_p e^{-pi\omega\Delta t}) F(y(\Delta t \cdot n)) \tag{3.6}$$

$$H(\omega) = a_1 e^{-i\omega\Delta t} + a_2 e^{-2i\omega\Delta t} + \cdots + a_p e^{-pi\omega\Delta t} \tag{3.6b}$$

Equation 3.6b is the autoregressive model equation in frequency domain where $a_1 .\ a_2 \ldots\ldots\ldots a_p$ are the coefficients of autoregressive model. In order to develop causal filter for the autoregressive model acceleration data of 2 span trusses is utilized. So we use 400 Hz frequency and fourth order autoregressive model. Figure 3.1 and 3.2 shows

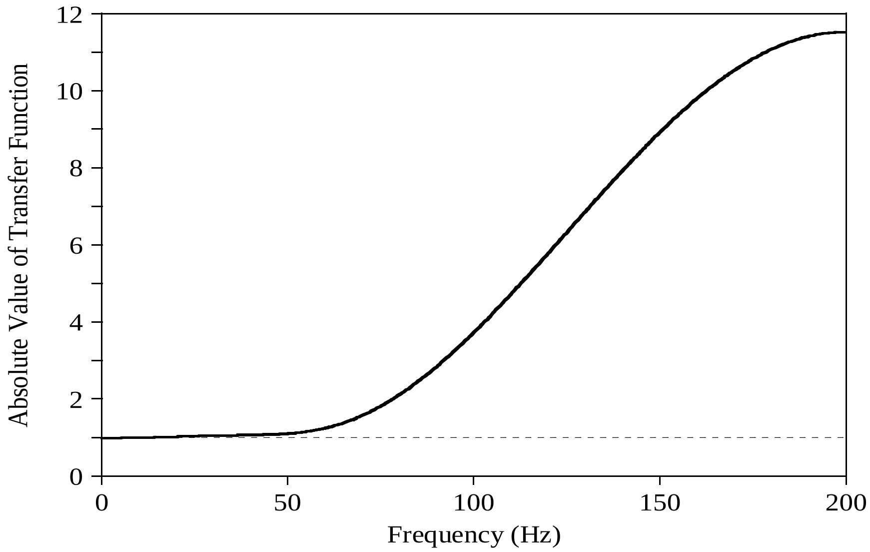


Figure 3.1 Casual filter and its absolute value

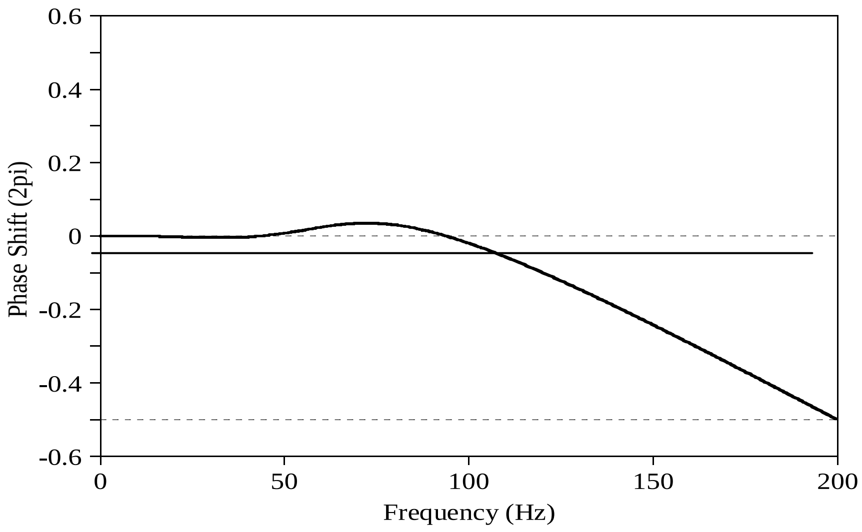


Figure 3.2 Phase Shift for Truss Structure data

In this dissertation non causal filtering technique is utilized to get the damage feature and finally decide for the threshold value for the acceleration data and using frequency domain. In Equation 3.7 a non causal autoregressive model is shown, from the expression we can see that in non causal filter the present value is accessed by the previous and future values of measured acceleration.

$$\begin{aligned}\hat{y}(\Delta t\, n) &= a_{-p} y(\Delta t(n+p)) + \cdots + a_{-1} y(\Delta t(n+1)) \\ &\quad + a_1 y(\Delta t(n-1)) + \cdots + a_p y(\Delta t(n-p))\end{aligned} \tag{3.7}$$

We can see from the equation 3.7 that autoregressive process works for the data from $\Delta t(n-p)$ to $\Delta t(n+p)$,expression for the non causal filter in frequency domain can be given as follows

$$H(\omega) = (a_{-p} e^{piw\Delta t} + \cdots + a_{-1} e^{iw\Delta t} + a_1 e^{-iw\Delta t} + \cdots + a_p e^{-piw\Delta t}) \tag{3.8}$$

Fig 3.3 and 3.4 shows the non causal filter for the figure 3.1 and 3.2; we can compare the changes in both phase shift and transfer function

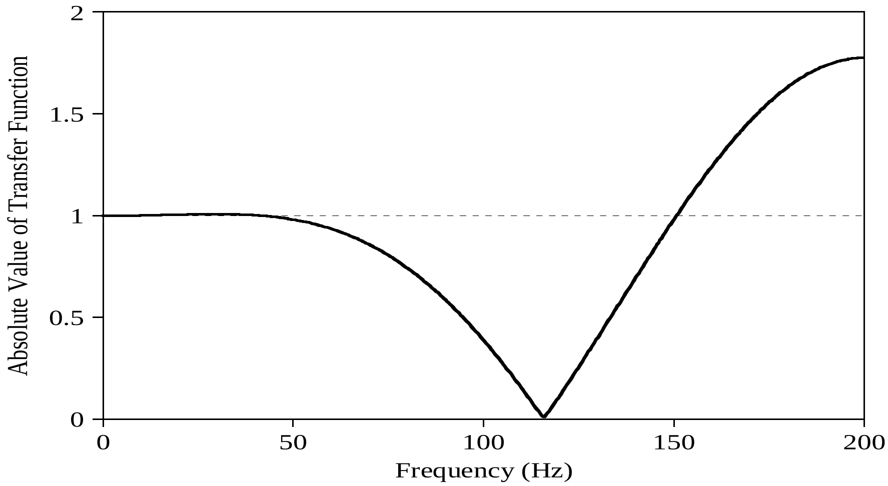


Figure 3.3 Transfer function for truss structure for non causal filter

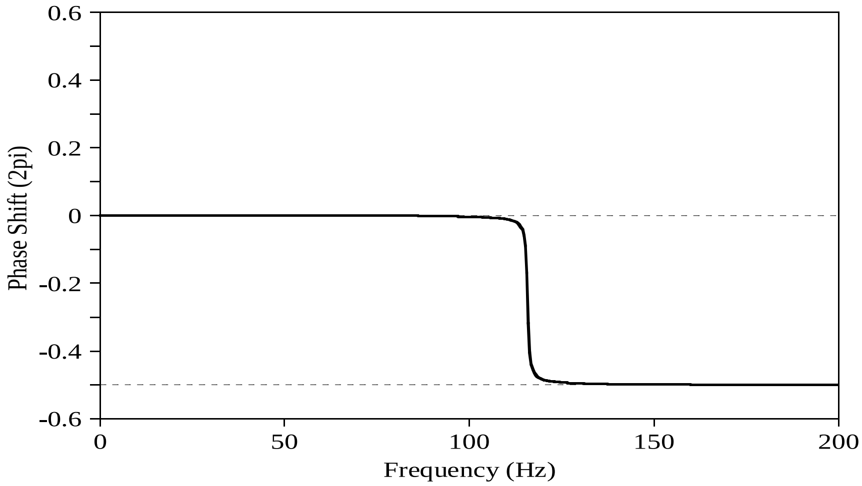


Figure 3.4 Phase shift for acceleration data of truss structure in non casual filter

As we have to remove the noise frequency we use frequency band filter, which is a device that passes frequencies within certain range and rejects frequencies outside that range. An ideal band pass filter would have a completely flat pass band (e.g., without no gain / attenuation thought) and will completely attenuate all frequencies outside the pass band. Additionally the transition out of the pass band would be instantaneous I frequency. Generally, the design of filter seeks to make the roll-off as narrow as possible, thus allowing the filter to perform as close as possible to its intended design. Often this is achieved at the expense of pass band or stop band ripple. The bandwidth of the filter is simply the difference between the upper and lower cut off frequencies. A filter pass band test was conducted by using non casual filter for frequencies 6Hz, 15Hz, 30Hz and for the frequency of 400Hz by using fourth order autoregressive model.

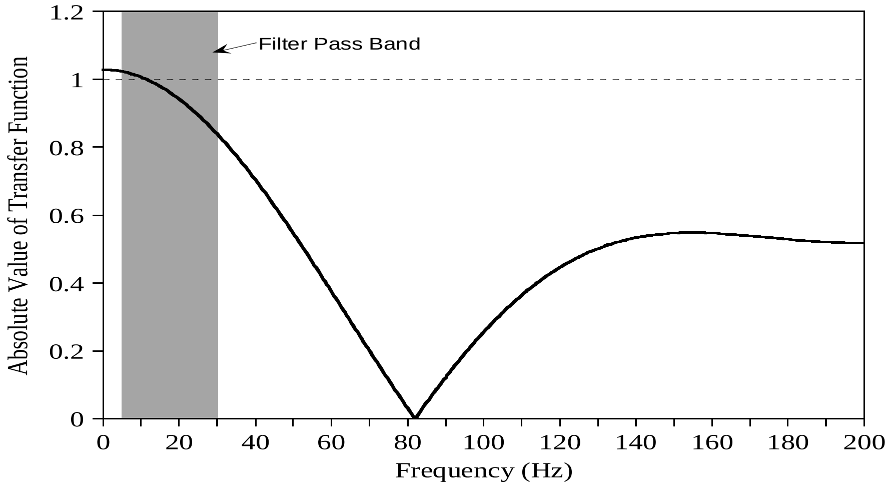


Figure 3.5 Frequencies of 6Hz, 15Hz, and 30Hz applied to the transfer function of autoregressive model

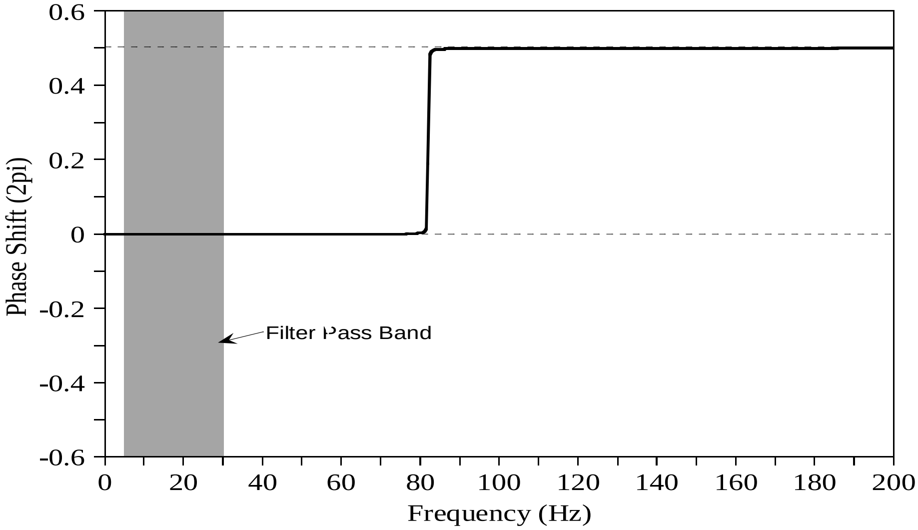


Figure 3.6 Phase shift of AR model which applies 6Hz, 15Hz, and 30Hz frequencies

3.3 Damage feature in Frequency Domain

Transfer function of the autoregressive model can be used for the damage detection of the structures. Using transfer function in frequency domain and its sensitivity to frequency are discussed in detail in the previous section. Transfer function can give us good damage feature as from previous discussion, the location where the phase difference is zero. And transfer function of autoregressive model spreads more widely where the transfer function has the value equal to one. When we get coefficients of AR model we get one transfer function here we use absolute value of transfer function. As residual is sensitive to amplitude change and coefficients are sensitive to frequency change, we apply covariance between them as damage feature. It should be noted that when there is damage in the structure both frequency and amplitude gets changed. In fig 3.11, it is shown that when the vehicle enters the structure there is abrupt change in the frequency of vibration of the bridge which can be seen in the frequency spectrum. So we have to make distinction between the frequency increase due to vehicle and damage frequency of the structure. From the frequency spectrum it can be seen that the structural frequency after entering of the vehicle ranges from 5HZ ~ 15 Hz and including other environmental effects frequency can be in the range of 20Hz ~60Hz

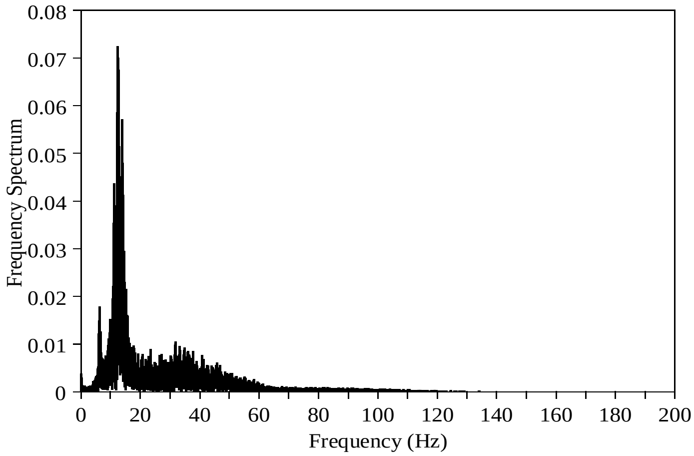


Figure 3.7 Normal Frequency Spectrum for 2span Truss Bridge

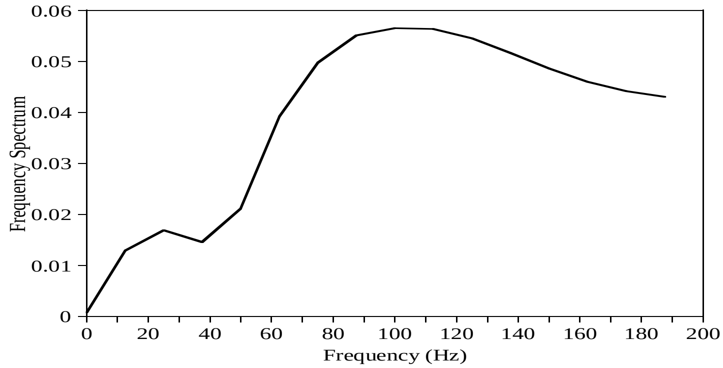


Figure 3.8 Damage Frequency of the Truss Bridge

In fig 3.8 for the truss bridge we can see that frequency for the damaged reaches to 100 HZ.

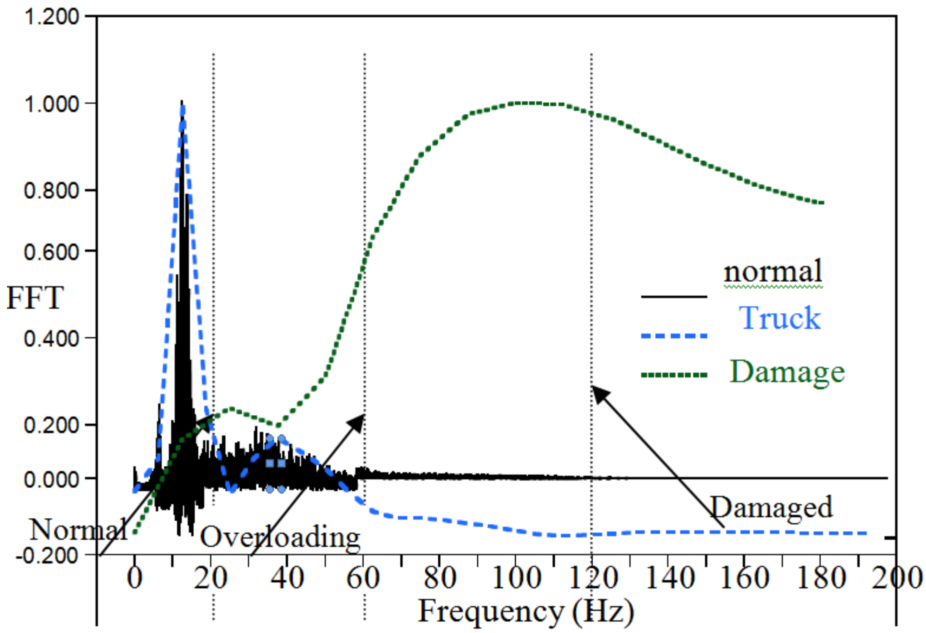


Figure 3.9 Frequency spectrum for normal, overloading and damaged conditions

As shown in the fig 3.9 that signal contain mix type of frequencies, for different loading conditions of the bridge. For damage feature we have to filter the lower frequencies, as damage shows higher value of frequency which are usually 100Hz and more. When we have mixed signals we can see for normal state frequency is 20 Hz, for overloading it is upto 60 Hz, so when we draw the transfer function for this frequency we will see that it will be constant at one. So when there is no change in transfer function there will be any change in residual and hence there will be no change in covariance. So it will be same upto 60Hz in our example. As we are using FFT which is only sensitive to frequency or ratio of frequency, when we have the signal of 100 Hz there will be change in Transfer function and we will get residual which will result the damage feature. Damage index or feature which is the covariance between the autoregressive coefficients and residual is given by the following relation

$$D_t = \text{cov}[|e_t|,|\theta_t^1|] = E[|e_t - \mu_{e_t}|,|\theta_t^1 - \mu_{\theta_t^1}|] \tag{3.9}$$

3.4 Damage Detection Algorithm

For damage detection we have developed algorithm which can detect the damage and it's instant. As discussed before there are two parameters which are to be determined to develop the algorithm. One is AR coefficients and the other is residual. AR Coefficient changing to have filter pass band that exists around dominant frequency, while residual error occurs out of filter pass band. Covariance of residual and coefficients is taken is damage feature, which is activated while the frequency out of filter pass band of non- damage state occurs. An ideal transfer function is that which include non damaged frequency and must not include damage frequency in its filter pass band of non damage transfer function. If the length of the filter pass band is small it will show the truck loading as damage, which is the wrong signal. While if the length of the filter passes band is more and include the damage frequency, it will not give the damaged signal which is very dangerous situation. So appropriate length of filter pass band should be selected. I deal transfer function is one which filters the white noise and detect the damage feature. So damage algorithm should be developed in such a way that it is activated by damage, not by overloading or other noises.

3.5 Sampling Rate and AR Order

These are two very important parameters to get real time or close to real time health monitoring of the structure. We are using simulated acceleration data for this study. When the sampling rate is increased the range of filter pass band is increased. As discussed before that for the damaged structure frequency reaches upto 180 Hz , so let we take it as 200 which is said Nyquist

Frequency. But for setting the sampling rate Aliasing effect should be considered, i.e. the sampling frequency should have to be two times the Nyquist frequency, otherwise we will not be able to detect the damage in the given structure. While determining sampling rate ($1/\Delta t$) aliasing is important because too low sampling rate cannot draw a line between damage and non damage signal. So in this study we are using 400 sampling rate.

$$f_{sample} \geq 2 \times [nyquist \;\; frequency] = 2 \times [\, frequency \;\, when \;\, damage \;\, occurs]$$

Similarly AR Order is important for damage detection, if we have higher AR Order the range of filter pass band is increased. So optimal order should be selected for damage detection algorithm. In this study it is determined as 8 and can be seen in the following figure 3.11.

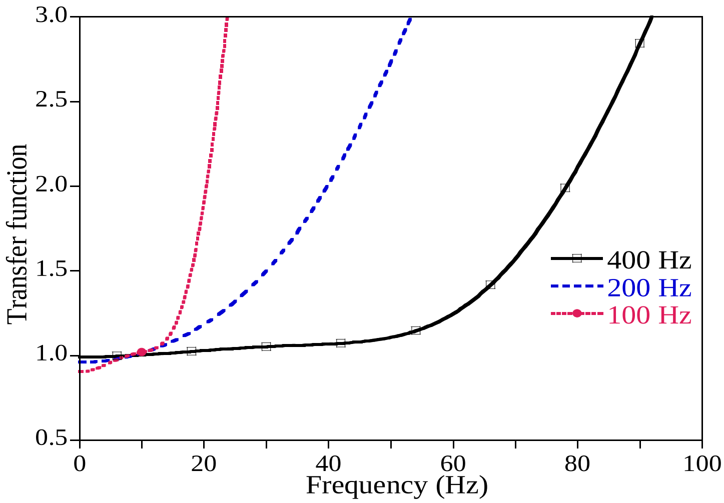


Figure 3.10 Sampling Rate

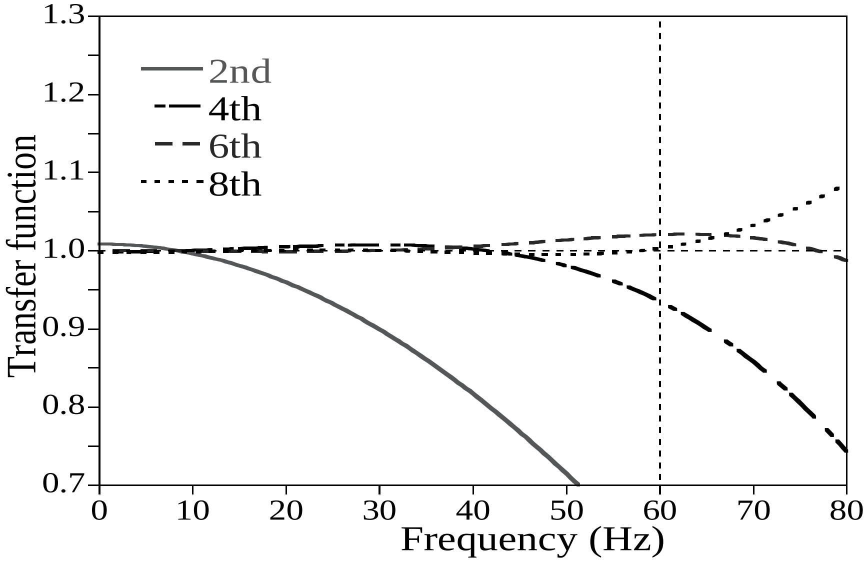


Figure 3.11 AR Order Selection

4 Extreme Value Theory

4.1 Extreme Value Analysis

Extreme value theory has emerged as one of the most important statistical disciplines for the applied sciences over the last 50 years. This theory deals the statistical distributions which deviates from the median of probability distribution. Extreme value theory is important to study the highly unusual events, or extreme events such as earthquakes, extreme floods, high wind speeds, rain fall, extreme temperatures, extreme sea waves, etc.

The distinguishing feature of an extreme value analysis is the objective to quantify the stochastic behavior of a process at unusually large – or small scale. In particular, extreme value analysis usually requires estimation of probability of events that are more extreme than any that have already been observed. This theory is based on the extremal type's theorem, also called three types theorem, stating that there are only three types of distributions that are needed to model the maximum or minimum of the collection of random variables from the same distribution. In other words if you generate N data sets from the same distribution, and create a new data set that includes the maximum values from theses N data sets, the resulting data set that includes the maximum values from these data sets, the resulting data set can only be described by one of the three models- specifically, the Gumbel, Frechet, and Weibull distributions.

4.2 Extreme value distribution

Extreme value distribution is the limiting distributions for the minimum and maximum of a very large collection of independent and identically distributed random variables from the same arbitrary distribution. Damage values as shown in the Fig 4.1 are the extreme values which lies at the tail of distribution , which are called outliers, so this prompts us to study the extreme value theory for this damage detection scheme.

In this study we used extreme value distribution and selected optimal sample size. One extreme value was extracted from each sample size and number of total extreme values were called as observation time. We used 4000 and 10,000 sec. observation time for our threshold value determination and for higher values of sample size and pseudo significance level.

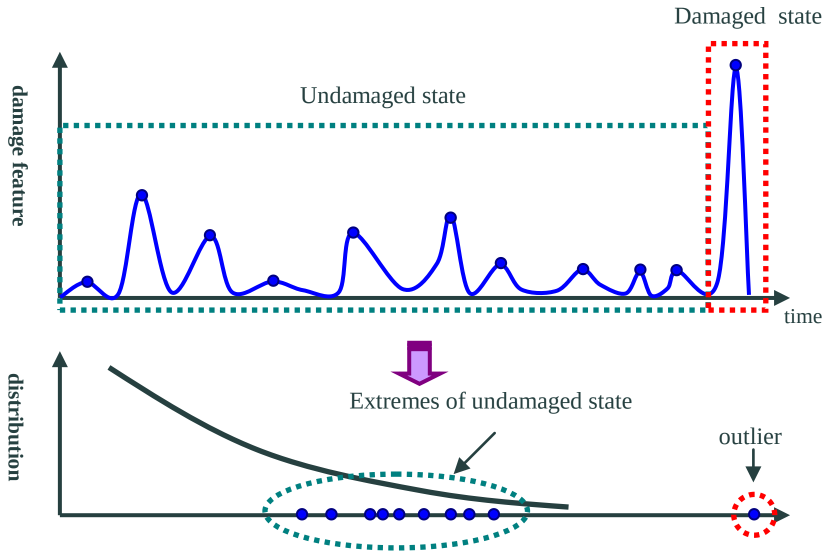


Figure 4.1 showing the outlier in damage feature

4.2.1 Extracting extreme Values and Sampling Technique

As we are treating the extreme values as random variables, so to get the extreme value distribution it is important that random variables should be identically and independently distributed. As original signal is in the form of continuous time series. So we divide the signal (Damage Feature) into sub sets which is called sample size. And from each sub set we will get one extreme value as shown in Fig 4.2

In this study we are using the sample size 700 and to get the reliable threshold value we are using 4000, 10000 extreme values. In practical sense total number of extreme values are called observation time which is equal to 5 hour for 10,000 extreme values.. So as shown in the fig 4.2 w have to randomly mix the continuous time series data.

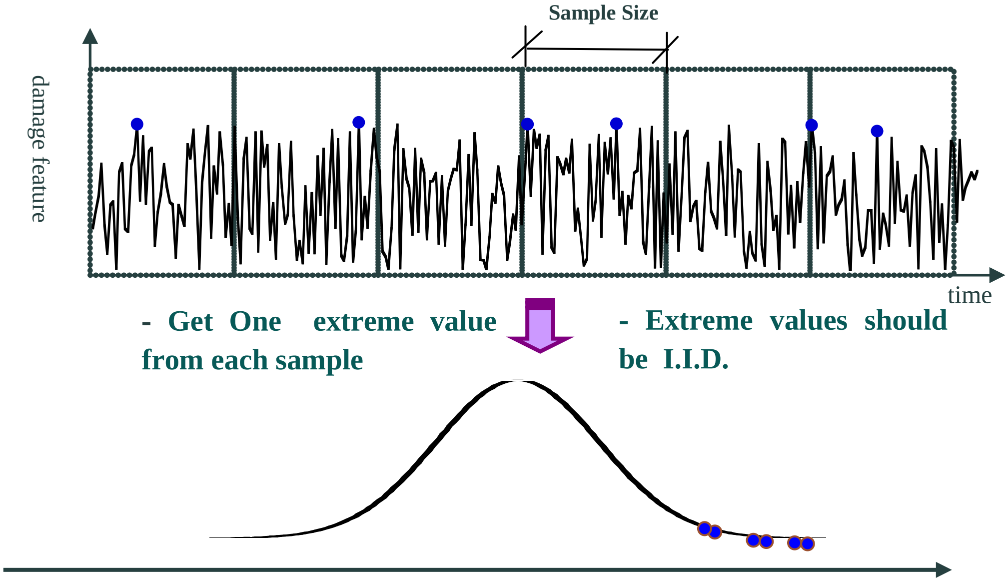


Figure 4.2 I.I.D Random Variables

Table 4.1 shows the results of i.i.d condition for extreme value distribution, and continuous time series signal for different values of significance level.

Table 4.1 Significance level and Threshold values

$\hat{\alpha}$	Time Dependent Dynamic Data		I.I.D Condition	
	Threshold	Percent over the Threshold	Threshold	Percent over the Threshold
0.1%	7.38×10^{-8}	8.5336%	1.35×10^{-6}	0.1075%
0.05%	1.67×10^{-7}	4.7704%	1.58×10^{-6}	0.0504%
0.01%	7.79×10^{-7}	0.4089%	2.05×10^{-6}	0.0096%

4.2.2 Origional and Extreme Value Distribution

A relationship between original (or parent) distribution and independent and identically distributed random variables (extreme values) can be established as given by equation 4.1

$$H_n(X) = F^n(X) \tag{4.1}$$

$F(X)$: CDF for original distribution

$H_n(X)$: CDF for EV distribution

X : I.I.D random variable

n : Sample size

As we can see from the fig 4.3, that when we want to use the extreme value distribution we have raise the power of CDF of original distribution to the power of sample size (n). To get the threshold value we will use extreme value distribution curve $H_n(X)$. Usually the extreme values lie at the extreme positions of the distribution, which are not clearly known. So in this study we propose to use the data of those extreme values in CDF, which are closer to the mean position and will give us more reliable threshold value from EV curve as shown in the fig 4.2.

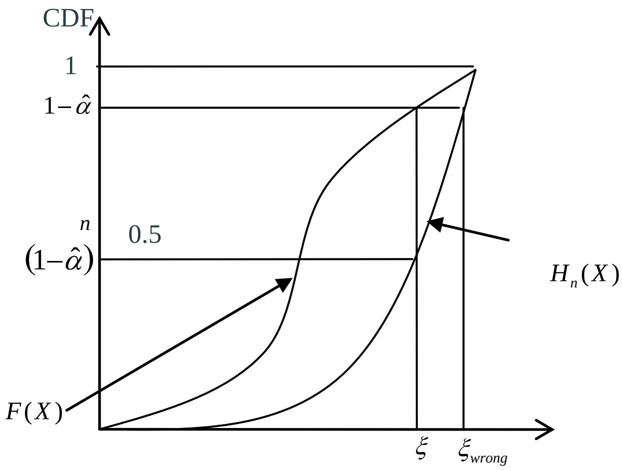


Figure 4.3 CDF and Threshold Value

4.2.3 Optimal sample size for threshold value

Optimal sample size to determine threshold value is also important , as shown in the above equation that abundant data lies near the mid of extreme value distribution curve so we use that part to get the optimal sample size, which is given by the following relation

$$n = \log_{(1-\hat{\alpha})} 0.5 \qquad (4.2)$$

In Fig 4.4 it is shown that for different values of sample size we have threshold values, which becomes stable near the optimal sample size 700. So in our study we are using 700 sample sizes to determine threshold value.

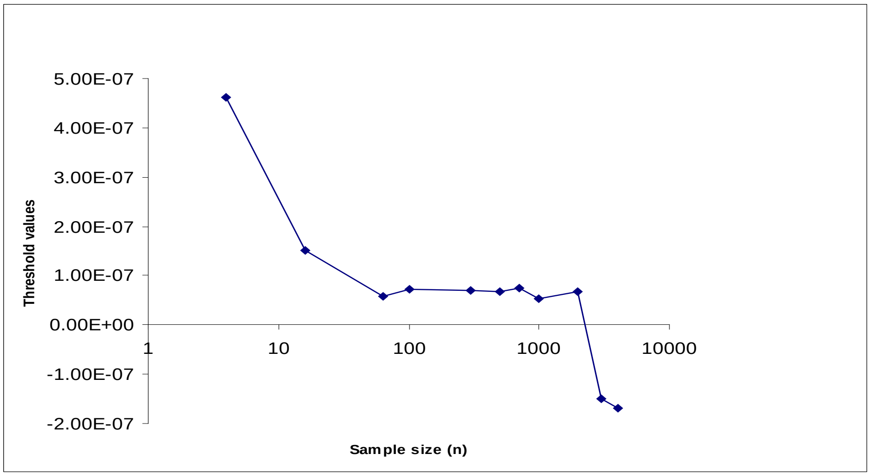


Figure 4.4 Optimal sample size for threshold value

4.3 Threshold Value

Threshold value is the level which separates or divides the feature into damage and undamaged parts. We know that when the structure is damaged, the distribution of the data for those damaged point lies at the $H_n(X)$ right extreme location, those values are called outliers and monitoring system should find those outliers or damage points at real time or close to real time to assure the safety of the human life. And when there is damage the feature crosses the threshold value, so we can easily find the damage instant.

Extreme value theory is well established to find those outliers. Extreme value distribution is utilized in our study, sampling technique which imposes i.i.d condition while extracting extreme values from continuous time series signal. Those damage feature values are treated as random variables and are randomly mixed to get extreme values from each sample size. So sampling is very important which is optimized in our study and then comes the decision making about those extreme values to decide for damage or undamaged situation of the structure. For that purpose original Hypothesis was modeled in different way and a pseudo significance level is determined. Because we don't know about the actual damage area of the structure. Above methodology is verified for generated acceleration data for 2 span continuous truss. Based on extreme value theory and optimal size of sample threshold values are derived which shows the results according to the theory.

4.4 Algorithm for Threshold Value

An algorithm is developed to determine the threshold value which uses the cunane distribution for the extreme values and by interpolation we can get the threshold value, from that distribution.

Cunane Distribution = $\left(x_i , \frac{i-0.40}{n+0.2} \right)$ (4.3)

In Threshold value algorithm we have three parameters as input which includes sample size, significance level and observation time or number of extreme values.

$$D_{cr} = \left[H_n((1-\hat{\alpha})^n) \right]^{-1} = \left[F^n(1-\hat{\alpha}) \right]^{-1} \quad (4.4)$$

We used randomly mixed damage feature data to get threshold value, in this study we have fixed sample size and we determine threshold value for different values of significance level.

4.5 Decision making and Hypothesis Testing

To decide whether the considered structure is sound or not using estimated results from prediction model is also very important. No matter how prediction model may work perfectly, it is useless without support of rigorous decision making algorithm. It is unreasonable to decide health of the structure by merely the magnitude of residual errors. For more reliable decision making of structural health monitoring, statistical approach is inevitable. Distribution of the residual errors must be found statistically from sparse residual and pick up outliers from the distribution in a given significance level. Outliers almost lie in the tail of distribution of residual errors extreme value distribution is utilized for more accurate selection of outliers because extreme value distribution is well established for tail distribution.

Hypothesis testing is a method of making decision using experimental data it is also sometimes called confirmatory data analysis. In frequency probability , these decisions are almost always made using Null hypothesis tests i.e. the test that answer the question assuming that the null hypothesis is true, what is the probability of observing a value for the test statistic that is at least as extreme as the value that was actually observed. Statistical hypothesis testing is a key technique of frequentist statistical inference, and is widely used.

While hypothesis testing and decision making we have to use the significance level. We have defined the hypothesis in our study that if there is damage there is alarm and null hypothesis as if there is damage there is no alarm. Now significance level which is defined as incorrectly rejecting the null hypothesis or measure of the probability of Type 1 error is defined from the fig 4.5 by the relation

$$\alpha = \frac{b}{b+c} \tag{4.5}$$

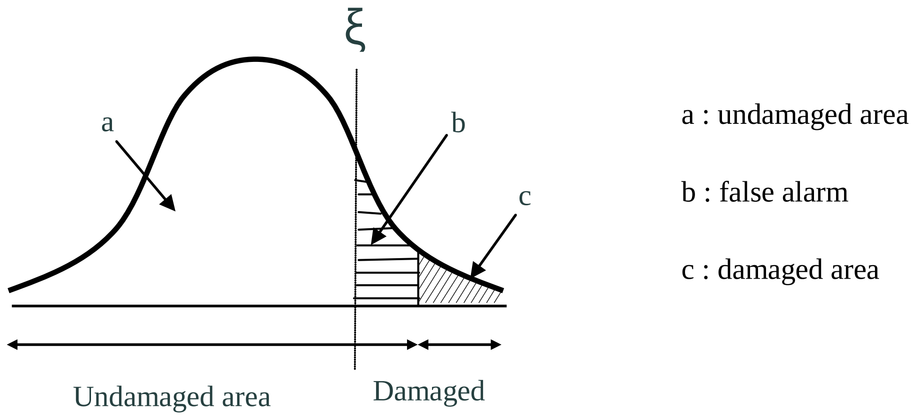


Figure 4.5 Hypothesis testing

But the problem is that in our original structure we don't know the damage area, so we cannot define the significance level as the measure of type one error because we don't know the actual damage in the structure. We propose another significance level called as pseudo significance level and defined by the relation

$$\hat{\alpha} = \frac{b}{a+b} \tag{4.5}$$

Which is the measure of the probability when there is no alarm so in this study this pseudo significance level is used

5 Numerical Example

5.1 Example 2 Span Truss Bridge

The validity of the proposed structural health monitoring is verified through a two span continuous truss shown in the fig 5.1. The sensors are cross located each other in order to avoid loss of information because of symmetry. Typical material properties of steel (young modulus=210 Gpa, specific mass=7.85kg/m3) are used for all truss members. The cross sectional area of top, bottom, vertical and diagonal members is 1122.5 m^2 , 93.6 m^2 , 62.5 m^2 respectively.

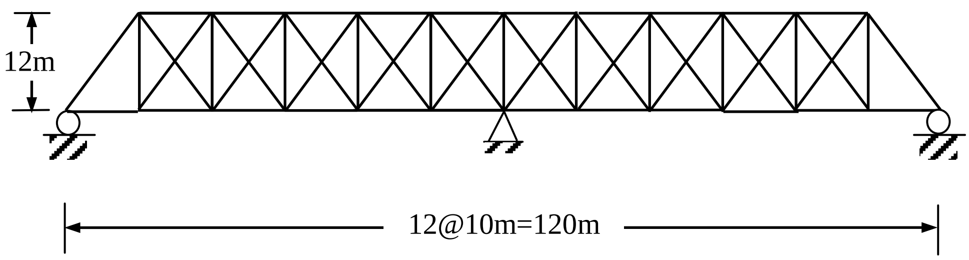


Figure 5.1 two span Truss Bridge

The natural frequencies of the truss range from 6.6Hz to 114.7Hz. The damping characteristic are simulated by 5% Rayleigh damping. Sampling rate 400 Hz and the duration of simulated acceleration is 1 hour (3600 seconds). 5% proportional noise is put on measured data to consider measurement noise.

5.2 Loading Scenario and Damage Scenario

It is assumed that accelerations are measured under normal operational condition. The moving vehicles are classified into three types' car, bus and truck and assumed that distribution of weight follows normal distribution. The car, bus and truck follows N $(2.3,0.2^2)$, N $(13.5,3.2^2)$ and N $(33.8,2.9^2)$ respectively. The limit speed of universal road 60km/h is applied for speed of vehicle load and 20% reduction of speed for truck is applied. The car, bus and truck vehicle loads are generated by 77%, 15% and 7% respectively. The overloading condition is also generated three times at 2244 second, 2474 second and 3401 second in order to compare with changes due to damage.

It is assumed that sudden damage occurs twice in the considered structure at 2730 second and 3002 second. Damage is implemented as reduction of cross sectional area. The cross sectional area of upper member "B" and lower member A are reduced at first damage instant 40%, 50 % respectively and lower member "A" are reduced additionally at second damage instant by 20%. Damaged members are represented by dotted lines in the fig 5.2.

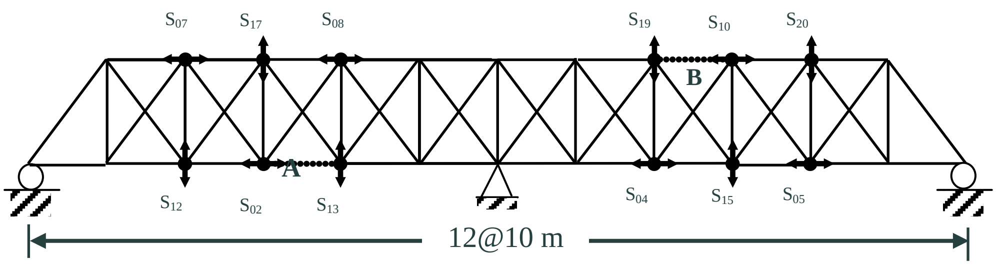
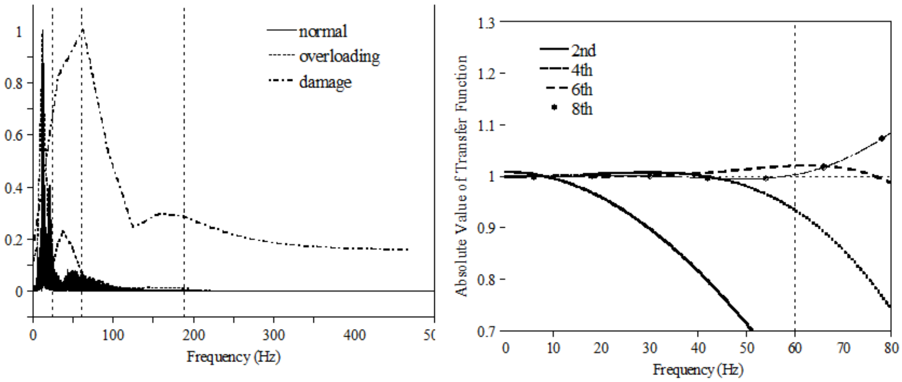


Figure 5.2 Structure with the location of sensors and damage

Minimum sampling rate should be fixed as 400Hz because 200Hz signal comes from damage state (figure5.3). AR model must draw a line on 60Hz because normal and overloading condition must be considered as non-damage and 8th order AR model is employed because it can draw a line on 60Hz.

Pseudo Significance Level $\hat{\alpha}$ is 0.0001 and the best sample size (n) for $\hat{\alpha} = 0.0001$ is 7000, which is determined by equation 4.2.

The results for the damage detection and threshold values are shown for the sensors which are deployed in Fig 5.2



frequency spectrum   change of $|H(\omega)|$ with change of AR order

Figure 5.3 Frequency spectrum and AR Order

In the following pages we are showing the results for damage feature which are taken for sample size 100, and AR Order 8 which is determined by the filter design. And for threshold value we have used sample size 7000 and observation time 5 hours with significance level of 0.01%. Which means that there will be 0.01 % of values above Threshold. Those values can give false alarm and also contains damage feature. To differentiate these two signals, sensitivity level of sensors should be adjusted such that it ignores false alarms.

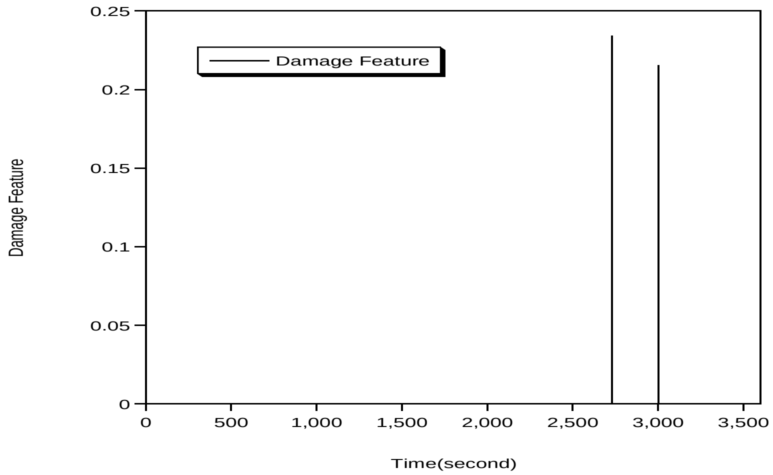


Figure 5.4 Damage Feature for Sensor S02

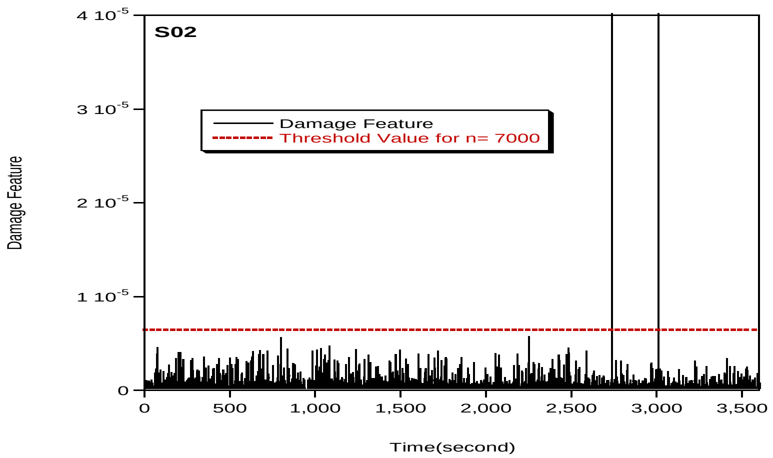


Figure 5.5 Threshold Value for Sensor S02

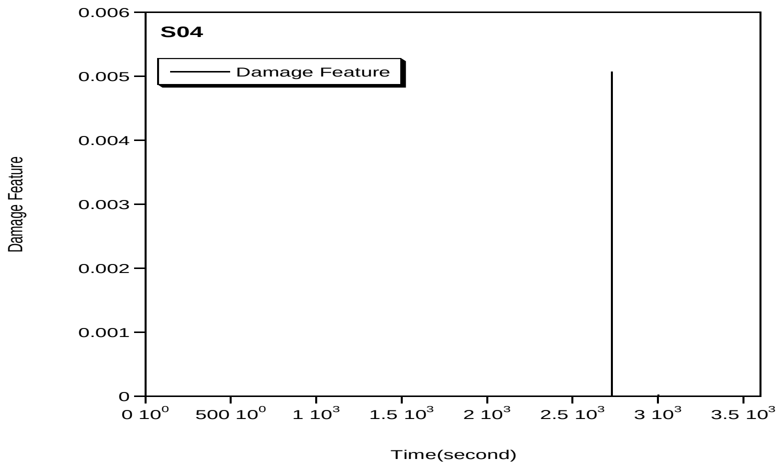


Figure 5.6 Damage Feature for Sensor S04

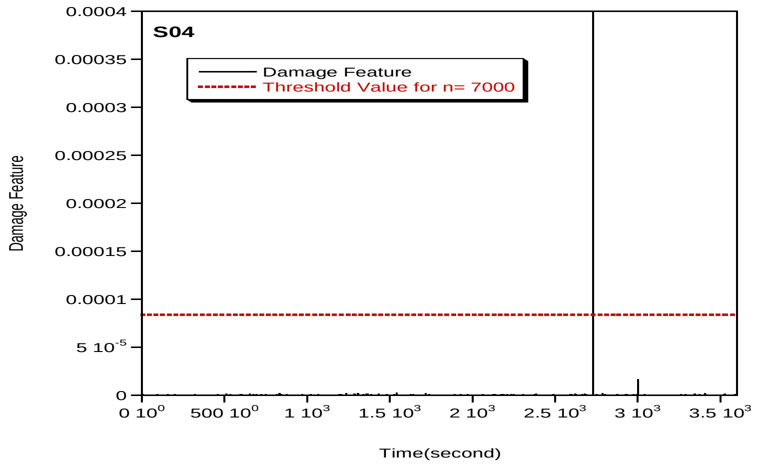


Figure 5.7 Damage Feature for Sensor S04

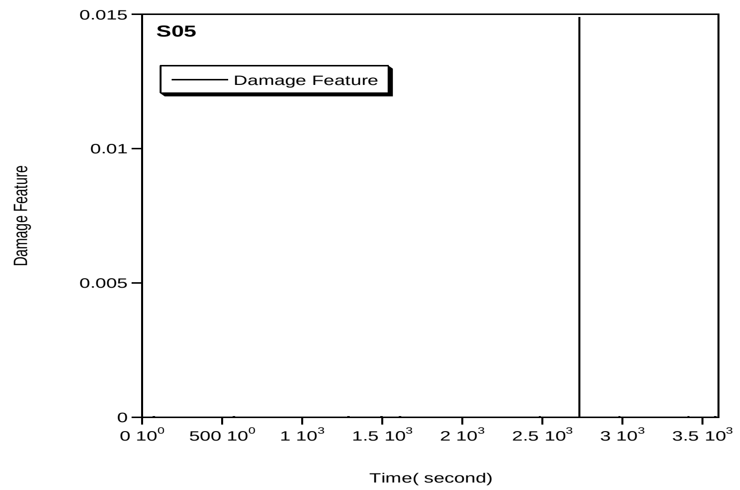


Figure 5.8 Damage Feature for Sensor S05

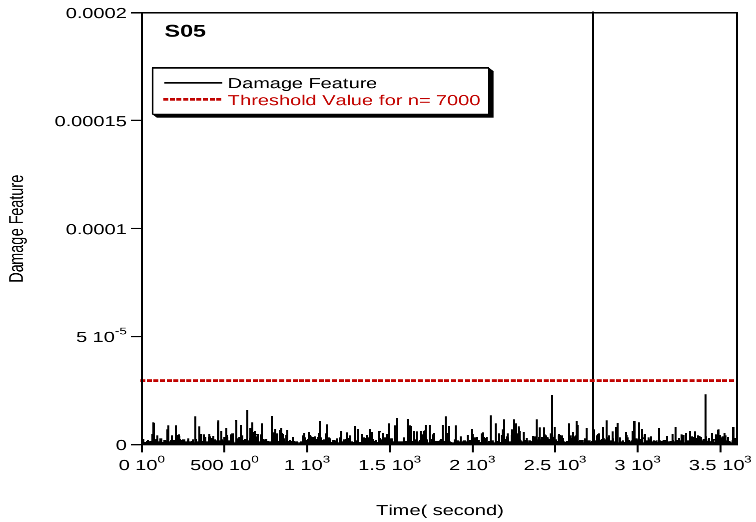


Figure 5.9 Threshold Value for Sensor S05

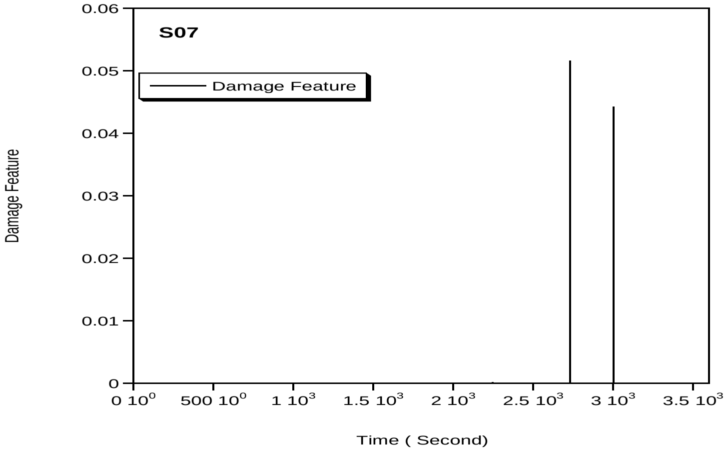


Figure 5.10 Damage Feature for Sensor S07

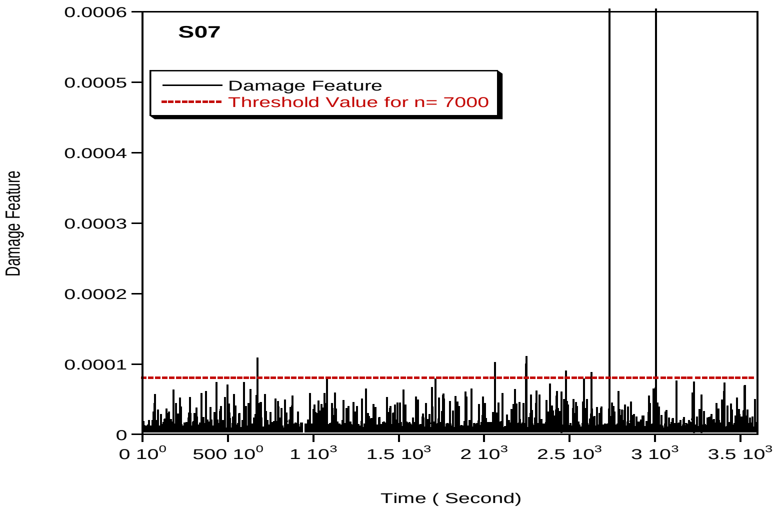


Figure 5.11 Threshold Value for Sensor S07

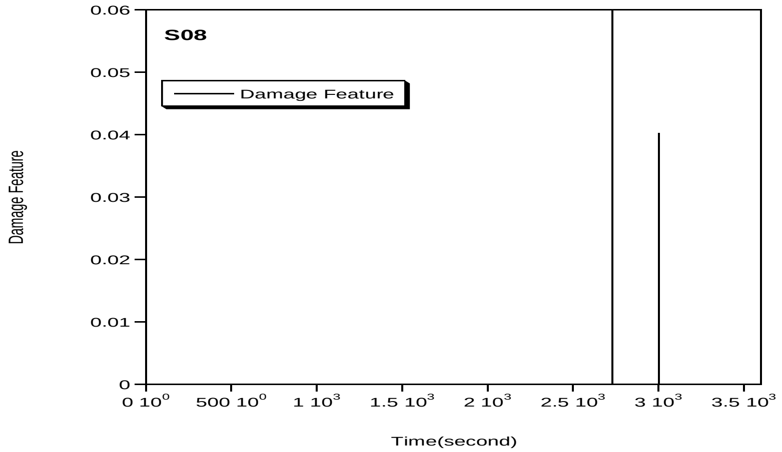


Figure 5.12 Damage Feature for Sensor S08

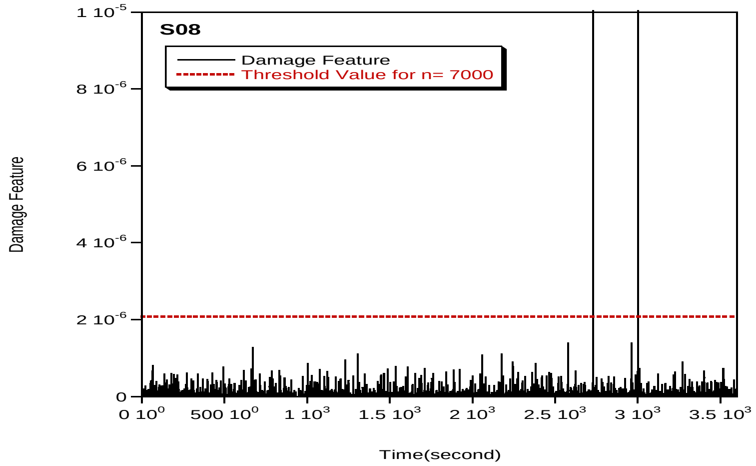


Figure 5.13 Threshold Value for Sensor S08

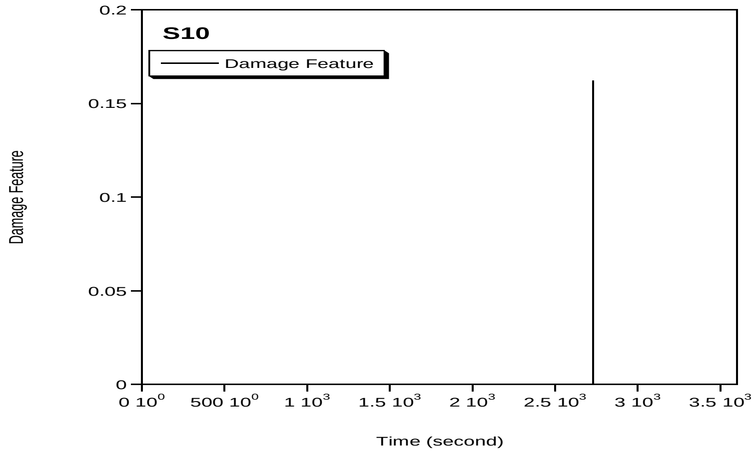


Figure 5.14 Damage Feature for Sensor S10

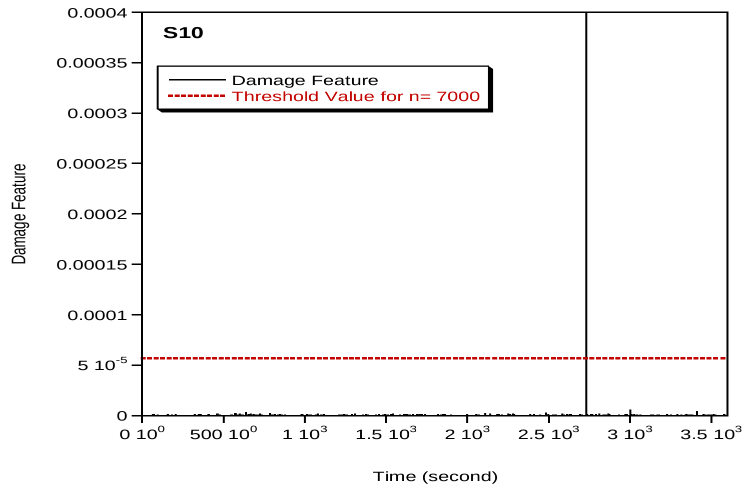


Figure 5.15 Threshold Value for Sensor S10

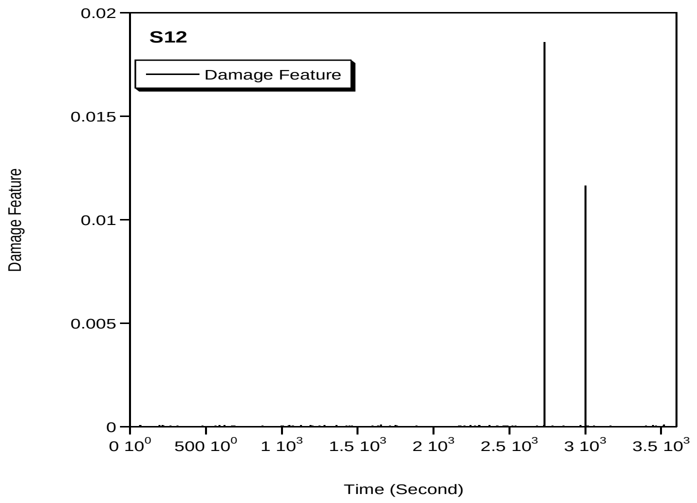


Figure 5.16 Damage Feature for Sensor S12

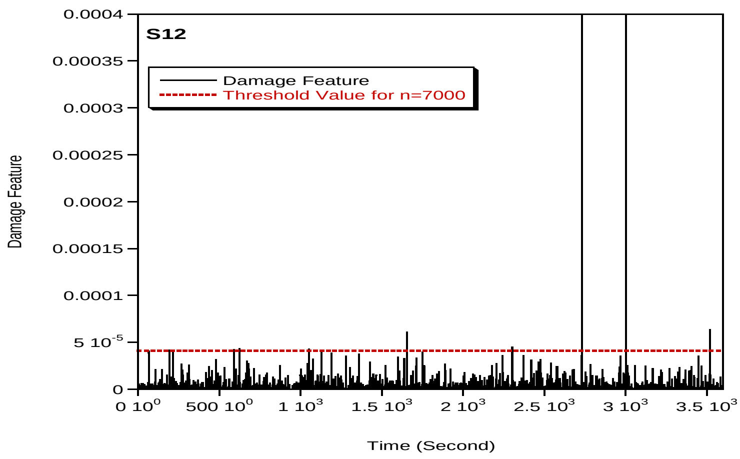


Figure 5.17 Threshold Value for Sensor S12

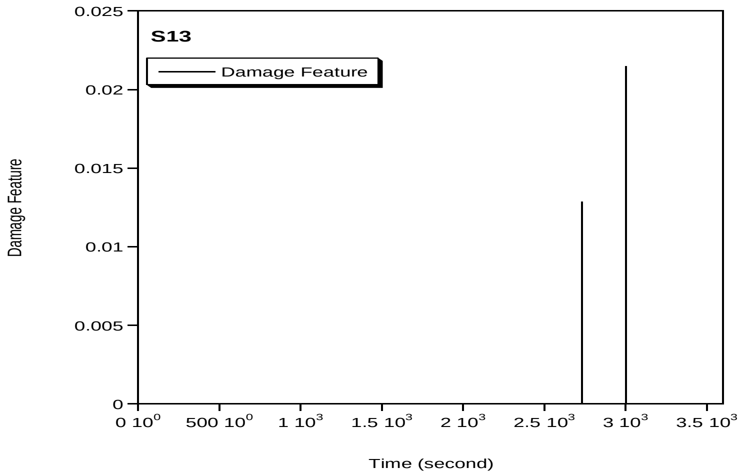


Figure 5.18 Damage Feature for Sensor S13

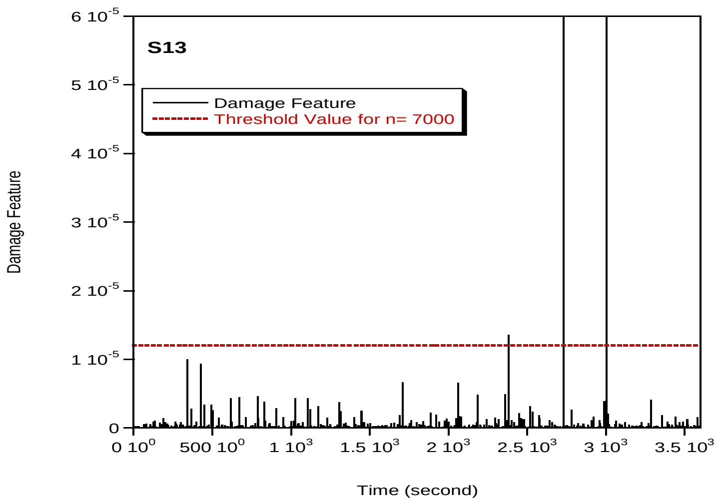


Figure 5.19 Threshold Value for Sensor S13

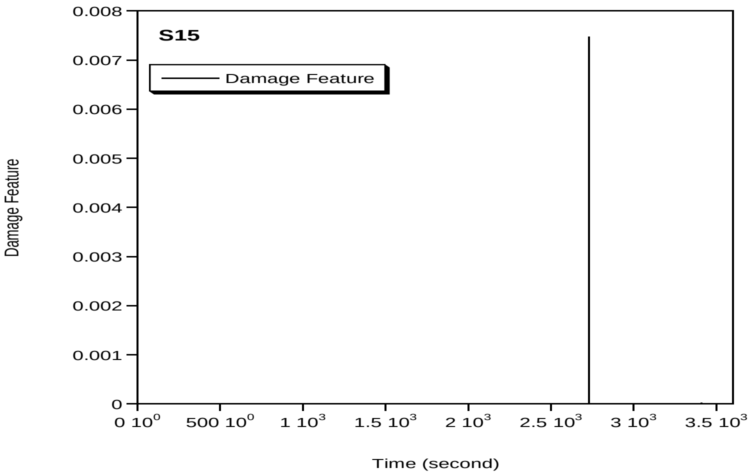


Figure 5.20 Damage Feature for Sensor S15

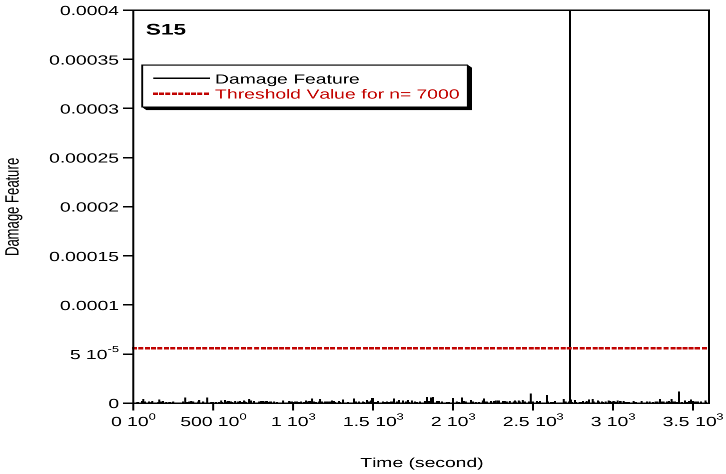


Figure 5.21 Threshold Value for Sensor S15

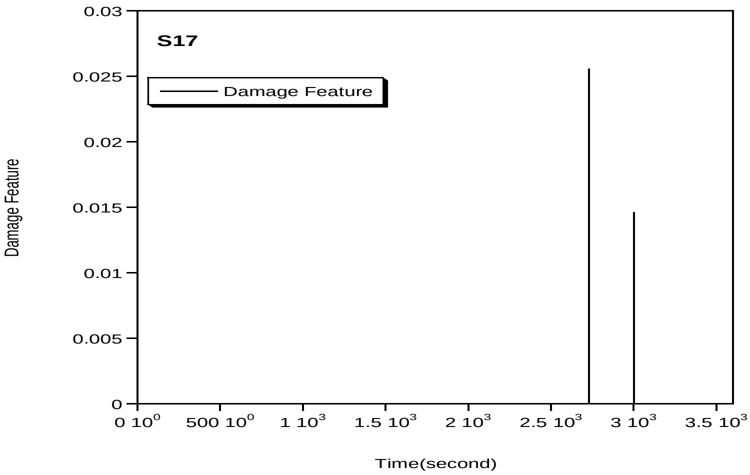


Figure 5.22 Damage Feature for Sensor S17

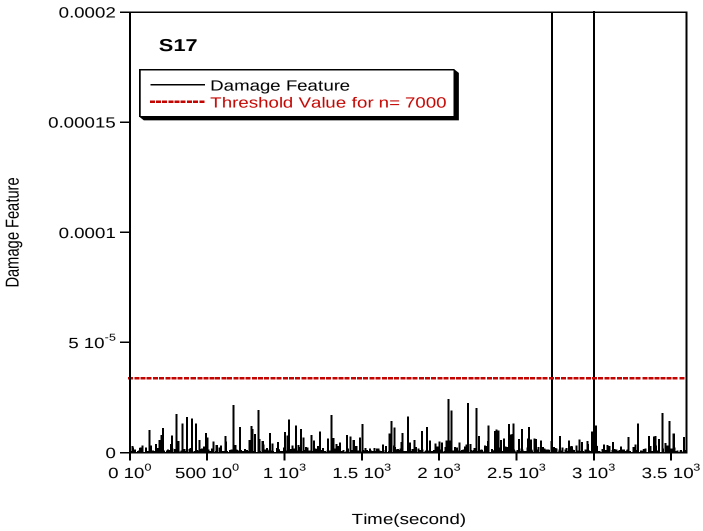


Figure 5.23 Threshold Value for Sensor S17

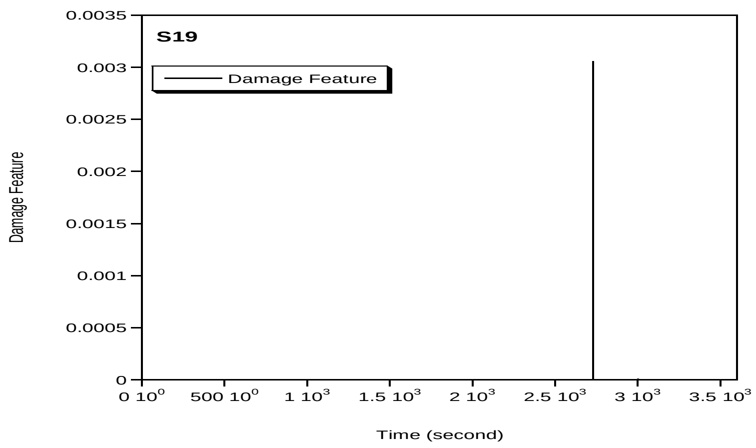


Figure 5.24 Damage Feature for Sensor S19

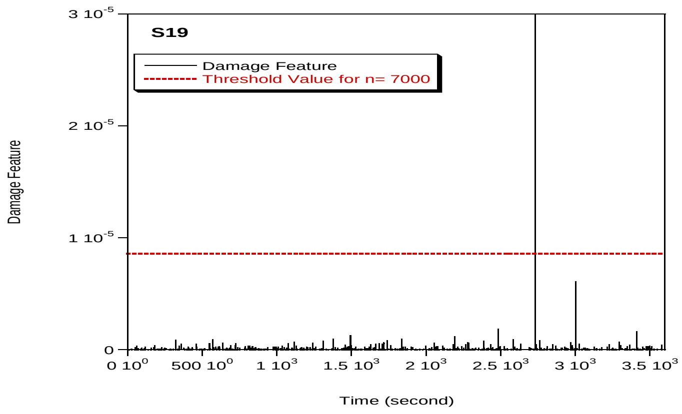


Figure 5.25 Threshold Value for Sensor S19

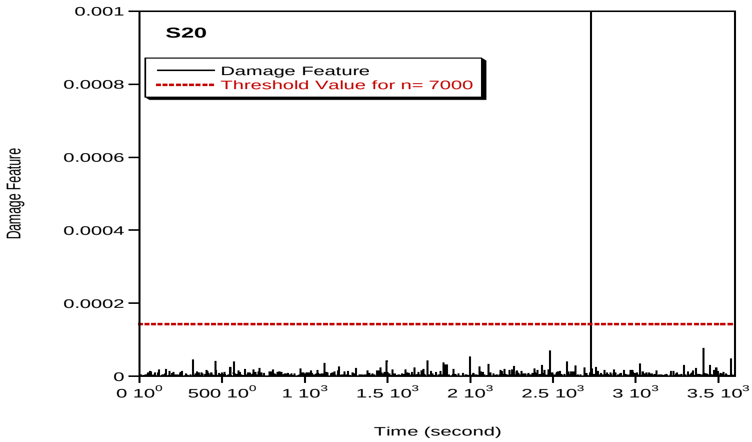


Figure 5.26 Damage Feature for Sensor S20

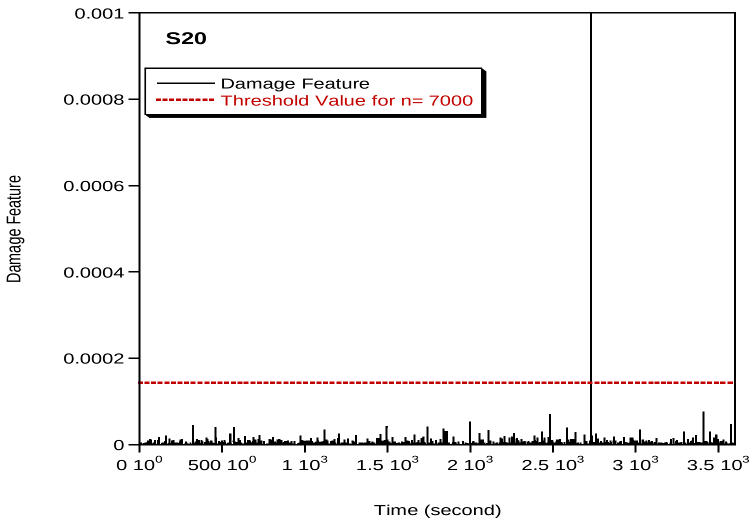


Figure 5.27 Threshold Value for Sensor S2

6 Results and Discussion

6.1 Conclusion

For any structural health monitoring system primary requirement is that, it should ascertain with confidence that whether the damage is present or not. So damage detection is most important for any structural health monitoring system. Then comes the decision making about the damage. This study was mainly related to get the threshold value and decide about the damage. The conclusion of the study is as follows

1- In the previous study generalized extreme value distribution was used and some threshold value was determined. Which could not match the actual situation. For extreme value distribution it is necessary that the random variables should have to be independently and identically distributed, which were not considered in previous study. So In this study extreme values were randomly mixed and i.i.d condition was employed which provided reliable results for threshold value.

2- Also it was found that while getting the extreme value distribution from original distribution; it should have to be raised to the power equal to sample size. Which was no considered in the previous study?

$$H_n(X) = F^n(X)$$

Due to this mistake extremely large threshold values were determined and while selecting the significance level for example 1%, it is statistically accurate that 1% values should have to be above threshold value. But when a very high threshold value is employed it cannot give 1% values above it. In that case extremely small significance level resembles the above situation.

3- Optimal sample size for the threshold value is determined from the mean value of the extreme value distribution as abundant data is available around the mean.
4- A Pseudo significance level for threshold value is proposed alternatively. As we don't know the actual damage area in probability distribution. So we cannot define the type 1 error.
5- it was also concluded in this study that to have a reliable structural health monitoring algorithm , monitoring time and observation time Should be given importance, as monitoring acceleration data of 5 Hours cannot represent the behavior of the actual structure. So to develop reliable SHM scheme, about one year acceleration data of the bridge should be utilized which is designed usually for 100 years.
6- It was studied that while increasing the observation time we can get better threshold value. When we increased our observation period to the monitoring period for the available data, we had better results for threshold value.

6.1.1 Further Studies

Proper decision making and hypothesis testing technique should be developed, because when we have very low threshold value, we may have many false alarms for the heavy vehicles entering the bridge. As we have already determined the frequency of those heavy loading cases in this study it is required to look into for avoiding large number of false alarms. Simulated acceleration data was used for this study, to get more reliable results practical acceleration data should be used and long monitoring duration should be employed. If it is possible the information about damaged state should be applied to this scheme with model based SI and/or with engineering experiences.

초 록

구조물은 그 수명 동안 여러 가지 형태의 하중에 노출된다. 특히 일반적인 상태의 하중상태와 비정상적인 조건의 하중 상태를 연구 하는 것이 매우 중요하다. 이 연구에는 여러 가지 손상 지표를 적용 할 수 있는데 이 논문에서는 자기회귀 모형을 이용해 손상 탐지를 수행하며 자기회귀 모형을 통한 잔차와 자기회귀의 계수 사이의 공분산을 손상지표로 적용 하였다.

비 구조모형 기반의 기법(non-model based scheme) 구조물에 부착된 가속도 측정기에서 얻은 정보를 토대로 손상을 탐지한다. 이 정보는 시간영역에 존재 하고 있는데 정보의 속성을 정확히 규명하고 이해하기 위해서는 전달함수라는 주파수영역 분석이 필수적이다.

또한 구조물로부터 얻은 신호에는 환경적 요인의 변화로 인한 잡음이 포함되어 있다. 이를 제거하기 위해 시간창 기법(time-windowing technique)을 도입한다. 환경적 요인은 매우 긴 주기로 변화 하므로 시간창 기법을 통해 잡음을 걸러낼 수 있다.

이러한 알고리즘을 2 경간 트러스 교량 예제에 적용하여 손상 탐지를 수행 하였다. 구조물의 손상은 트러스 부재의 단면적 감소로 표현 한다.

또한 승용차, 버스, 트럭등의 일반적인 하중 조합을 생성하여 구조물에 적용 시켰다.

극치분포는 손상지표의 외치들을 걸러내기 위해 도입 하였다. 정상치와 외치를 구분하는 임계값은 유의수준과 sample size 를 통해 결정 된다. 이 논문에서는 극치분포를 이용함에 있어 기존 연구에서 놓쳤던 부분을 보완하는 내용이 포함 되어 있다.

<u>주요어</u>

손상탐지, 시간창 기법, 자기회귀 모형, 손상지표, 비 인과필터, 임계값, 극치분포

학번 : 2008-23538

Bibliography

1- Kang, PhD Thesis (2008) SNU, Korea. “ Detection of Abrupt Changes of Structure using Regularized autoregressive Model”

2- Hoon Sohn , David W. Allen and et al, “Structural Damage Classification Using Extreme Value statistics” j, of dynamic systems, measurement and control march 2005, vol 127/125

3- Sturat Coles springer series in statistics “An introduction to statistical modeling of extreme values”

4- Gumbel,E.J., 1958, statistics of extremes, Columbia University Press , New York.

5- Castilo , E, 1998 , Extreme value distribution Theory and Application, Academic Press series in statistical modeling and decision Sciences, San Diego, CA.

6- K Worden and G. Manson “Damage detection Using Outlier Analysis” Journal of Sound and Vibration (2000) 229(3), 647-667

7- Seung –k park, Hyun – moo Koh, Hae sung Lee, “ structural damage detection using time windowing technique from measured acceleration during earthquake

8- J.S Kang, H.S.Lee IABMAs (2006) , “ Structural Health Monitoring Using Dynamic responses with regularized autoregressive model”

9- J.S.Kang ,S.K Park, H.S.Lee, 2005” Structural system identification in time domain using measured acceleration” j. of sound and vibration 2008,pp 215-234.

10- Akaike, H. (1969). “Power Spectrum Estimation through Autoregression Model Fitting”, *Ann. Inst. Stat. Math.*, vol. 21, pp. 407-419

11- Aktan, A. E., Catbas, F. N., Grimmelsman, K. A., and Tsikos, C. J. (2000). “Issues in infrastructure health monitoring for management”. *Journal of Engineering Mechanics*, ASCE, 126(7), pp. 711-724

12- Oppenheim, A. V., Schafer, R. W. (1999). *Discrete-Time Signal Processing – Second Edition.* Prentice-Hall, Inc. Upper Saddle River, New Jersey.

13- Oppenheim, A. V., Schafer, R. W. (1999). *Discrete-Time Signal Processing – Second Edition.* Prentice-Hall, Inc. Upper Saddle River, New Jersey

Acknowledgment

First of all, I am thankful to Almighty Allah (God) for giving me strength, determination and passion for accomplishing this work.

I am grateful to my Advisor , Professor Hae Sung Lee , for his support and guidance not only in my studies but also helping me in my domestic problems. His interest and enthusiasm for my reserarcch work were unparralled and his innovation and command over different subjects of civil engineering will continue to influence me throught my career. It has truly been a privilage to learn under his guidance and supervision.

I express my profound gratitude to the Higher Education Commission (HEC), Government of Pakistan, for the award of MS-Engineering scholarship under the HEC project titled: MS Level Training in Korean Universities/Industry. This opportunity enabled me to pursue my higher education in a technological challenging atmosphere of a reputed university of South Korea. The professional management of our scholarship program was a great source of motivation for us to do research in a stimulating yet conducive environment. The efforts made by HEC helped me in expanding my frontiers of knowledge to the next level.

I am very thankful to my seniors S.K Park, Y.H Hong, K.J Jung, K.W Lyu, S.H Lee, J.S Kim and my mentor Mr K.Y Park for their support in my studies and taking care of me and making my stay comfortable and memorable in Korea. I also want to thank H.S Kim , J.H Baek , K.H Kim and J.H Lee. I wish good luck for all of the members of Structural Analysis Lab. It wouldn’t have been so simple to complete my studies without their help and support. I had unforgettable memories and great time with all the members of my lab. I also want to thank My senior from Pakistan Mr Zia for his support and encouragement.

Last but not least, I am great full to my family my brothers and sisters, specially my father and mother for their continuous support, encouragement and prayers. I am also thankful to my wife who supported me and prays for me in difficult times.

Muhammad Tahir Bashir.