

DAMAGE IDENTIFICATION USING INDUCTIVE LEARNING

by

Peter M. Tappert

Thesis submitted to the Faculty of

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

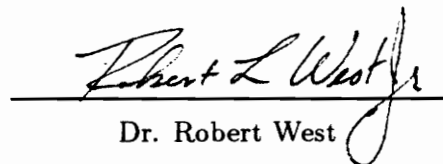
MASTER OF SCIENCE

in Mechanical Engineering

APPROVED:



Dr. Harry H. Robertshaw, Chairman


Dr. Zaffir Chaudhry
Dr. Robert West

November 16, 1994

Blacksburg, Virginia

c.2

LD
5655
V855
1994
T375
c.2

Damage Identification Using Inductive Learning

by

Peter M. Tappert

Committee Chairman: Harry H. Robertshaw

Mechanical Engineering

(ABSTRACT)

A damage identification method incorporating the use of inductive learning is presented. Inductive learning is the process of learning from examples. The method utilizes as much dynamic-response data as is available, ordering this information to find the best data with which to discriminate from among a set of damage states available for dynamic testing. This method takes into account the inherent variabilities in the damage identification problem. These inherent variabilities include but are not restricted to sensor noise, changes in environmental conditions, slight changes in boundary conditions, and manufacturing differences. The method statistically isolates changes in the dynamic-response characteristics due to damage from these inherent variances. This method is model independent and can be used to accommodate any sensors, actuators, and data type.

In order to demonstrate the method, an experiment was performed on a $12'' \times 12'' \times \frac{1}{8}''$ aluminum plate hung horizontally from the corners to simulate free-free boundary conditions. The plate was sensed and actuated by two piezoelectric patches mounted diagonally symmetric from one another. A small test mass (2% of the mass of the plate) was placed at four discrete locations, changing the physical properties of the structure. The structural impedance-responses were measured for all of the damage cases for both sensors. This information was processed by the damage identification algorithm to generate rules to which a small amount of data, extracted from a single set of structural impedance-response information, can be applied. The method was able to successfully discriminate all of the damage states from one another as well as detect the existence of a change in physical properties due to a damage state of which there was no prior knowledge.

Acknowledgements

This year and a half has been one of the most rewarding experiences of my life. A great deal of this satisfaction I owe to Dr. Robertshaw, my advisor and committee chairman. He has taught me a lot more than Engineering. I would like to thank him for giving me an opportunity to prove myself. I would also like to thank Dr. Chaudhry and Dr. West for being on my committee and for their suggestions.

I would like to acknowledge the support of the Army Research Office's University Research Initiative Center Program, Grant No. 03-92-6-0181; Dr. Gary Anderson, Program Manager. The same goes to the Center for Intelligent Material Systems and Structures for the use of their equipment.

I would especially like to thank Tom Snyder for his assistance in this work. His insight, suggestions, and time invested were invaluable. I hope I am able to return the favor. Special thanks also go to Dan Cole and Will Saunders, who were with me from day one. They gave me the confidence to make the decision to pursue an advanced degree and have also given me engineering insight that you cannot find in any textbook. I would also like to thank everyone in lab and the office, who have made this road a lot easier and enjoyable. They are (in no particular order): Tom, Darryl, Ben, Dan, Will, Gary, Jim, Paul, Steve, Randy, Bradley, EE Dan, Leslie, and all of Cudney's Freaks. I hope I did not leave anyone out.

Most importantly, I would like to give credit to my parents, where it is really due. I could not have accomplished anything without their moral and financial support. I want to let them know that their efforts and sacrifices are not unrecognized and unappreciated. . . Thanks!

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Literature Review	3
1.3	Objective	9
1.4	Scope	11
2	Statistical Approach	14
2.1	Variability	14
2.2	Distributions	15
2.3	Data representation	16
2.4	Distribution of a Sample Mean	18
2.5	Confidence Intervals for Means	21
2.6	Dealing with Unknown Variances	23
2.7	Statistical Test of Hypothesis	26
2.7.1	Verification of Equal Variance Assumption	28
2.7.2	Verification of Normality Assumption	30
2.8	Rule Generation	31
3	Experimental Verification	33
3.1	Testbed	33
3.2	Data Acquisition	34
3.3	Data Manipulation	36
4	Results	46

4.1	Rule Application Considerations	46
4.1.1	Method 1	46
4.1.2	Method 2	48
5	Conclusions	51
5.1	Significance	51
5.2	Future Work	52

List of Figures

2.1	Probability density function $f(x)$; shaded area represents $P(a \leq X \leq b)$	16
2.2	Random samples of imaginary part of impedance-response.	19
2.3	Sample means and variances of imaginary part of impedance-response.	19
2.4	Standard Normal Distribution	22
2.5	Student's t-distribution	24
2.6	Confidence Intervals for Means	25
2.7	Hypothesis test of two means, $H_0 : \mu_X = \mu_Y$	28
2.8	Graphical description of the test $H_0 : \sigma_X^2 = \sigma_Y^2$ against $H_1 : \sigma_X^2 > \sigma_Y^2$.	29
3.1	Piezoelectric and test mass locations	34
3.2	Ten examples of the 'no damage' case (real)	37
3.3	Ten examples of the 'no damage' case (imaginary)	37
3.4	Sample mean and variance of real part of the 'no damage' case	39
3.5	Sample mean and variance of imaginary part of the 'no damage' case	39
3.6	Predictor variable of damage cases A and B using imaginary part and red sensor.	42
3.7	Predictor variable of damage cases A and B using real part and red sensor.	42
3.8	Confidence intervals of real part of impedance-response.	44
3.9	Confidence intervals of imaginary part of impedance-response.	44
3.10	Discrimination between damage states A and B.	45

List of Tables

4.1	Results of Method 1.	48
4.2	Results of Method 2.	49

Chapter 1

Introduction

“Example is the school of mankind, and they will learn at no other.”

–Edmund Burke

1.1 Motivation

The structural damage identification problem has evolved from the desire to lower maintenance costs, lengthen the life of the structure, and reduce the risk of a catastrophic failure. Often the structure of interest is costly and time consuming to access for conventional visual inspection. These visual inspection techniques are often inadequate in identifying damage states of a structure invisible to the human eye, such as delaminations in composite materials. For this reason, non-destructive evaluation techniques such as ultrasonic and eddy current scanning, acoustic emission, and X-ray inspection have been developed. These methods, although useful in many instances, are very expensive, cause a great amount of down-time for the structure, and are impractical in many cases such as the in-service testing of aircraft and offshore structure foundations. These shortcomings of current non-destructive evaluation techniques indicate the need for damage identification methods which do not require direct human accessibility of the structure.

In order to develop a method of damage identification which does not require human accessibility of the structure, we make the claim that the response characteristics of a structure are functions of physical properties and boundary conditions. Therefore any change in physical properties (e.g. delamination, crack, hole, etc. . .)

or in boundary conditions (loose bolts, missing rivets, etc. . .) will cause changes in response characteristics. *Response characteristics* in this case refer to any kind of manipulation of the dynamic-response of a structure, be it an impulse-response, frequency-response, impedance-response, etc. The problem at hand becomes how to quantify these changes in response characteristics to indicate the location and extent of damage incurred. This damage identification problem is just one facet of researchers' attempts to develop a 'smart structure'.

In theory, a 'smart structure' contains an array of sensors and actuators which allow the structure to adapt in order to function more efficiently when influenced by perturbations in the operating conditions or in the structure itself, as when damage is incurred. The quest for a 'smart structure' has led to the development of lighter and stronger composite materials as well as a wide variety of new materials used for sensing and actuating. These include ferroelectric materials such as piezoelectric polymers (PVDF) and piezoceramics (PZT), magnetostrictive materials such as Terfenol-D, and shape memory alloys (SMA) such as Nitinol. Optical fibers have also shown to be an effective sensor. These developments of materials, sensors, and actuators have enabled researchers to more effectively implement structural control and damage identification algorithms.

As researchers attempt to control more complex structures and structures whose physical parameters are not too well known, they try to develop controllers which are more robust and which adapt to variations in the mathematical model of the plant. From this research, concepts such as 'colocation' and adaptive algorithms such as the least mean squares algorithm (LMS) have been developed. These developments have allowed researchers to free themselves of their dependence on very accurate plant models.

In order to improve a controller's effectiveness in the face of changes to the plant induced by damage, a damage identification process would be incorporated into the 'smart structure' scheme. Ideally, this damage identification process would be able to detect damage as it is incurred by the structure, determine the location and extent of the damage, predict if and when catastrophic failure of the structure will occur, and alert the operator as to how the performance of the structure is affected

so that appropriate steps can be made to remedy the situation. Obviously, this is no easy task.

Because the task of developing an effective damage identification process is so overwhelming, it is absolutely crucial that one carefully defines the aspect of the process they are addressing. As a specific aspect of the process is addressed, we must keep in sight how it fits into the overall process defined above.

The first step in defining the problem is to address what kind of system we are dealing with. We must know the number, location, and type of sensors and actuators available. We must know if the structure is unique, or produced in quantity. It is also important to be aware of the function of the structure, so that we know what types of damage we might expect. If there are damage cases that we expect, we must determine whether these cases are available for testing. We must also determine what variations we might encounter when the dynamic response of the structure is gathered. We must determine whether these variations are going to cause changes in the response characteristics of the structure. We must determine whether the type of damage expected is sensitive to the chosen mode of excitation. Finally, and most importantly, we must address the question of whether and how the solution of this problem fits into the complete damage identification process as defined above. It is on these criteria that we judge the usefulness of research in the field of damage identification.

1.2 Literature Review

Virtually all of the literature published regarding the development of a damage identification method that does not require human accessibility to the structure is based on the same claim we presented earlier: *Response characteristics* of a structure are functions of physical properties and boundary conditions. The problem at hand becomes how to quantify changes in response characteristics in order to indicate damage.

Kudva (1993), Marantidis (1993), and Gentry (1993) presented a series of papers concerning smart structures concepts for aircraft structural health monitoring. Kudva's paper assesses the state-of-the-art technologies associated with on-board

structural health monitoring systems (SHMS). These technologies include sensors, computers, analysis algorithms, hardware architecture, system integration, and technology demonstration. Marantidis' paper addresses in more detail the sensor and sensing technologies of structural health monitoring of aircraft. These sensing technologies include acoustic emission, fiber-optics, micro-electromechanical sensors (MEMS), corrosion sensors, computer vision, and neural-network-based systems. Gentry's paper presents a computer program entitled SHARP (structural health assessment and review program). SHARP involves a data acquisition system, a structural health library, a life analysis algorithm, and an evaluating expert system. SHARP, however, is in its basic stages and, when implemented, will be subject to many of the problems which will be discussed in the upcoming paragraphs.

The maturity of sensor and sensing technologies as Marantidis discussed is essential to the fruition of an effective damage identification method. There have been researchers in recent years who have relied solely on the maturity of a specific sensor in order to identify damage. By far the most prominent of these is the fiber optic sensor. Hofer (1987) presented the application of fiber optic sensors to the task of damage identification, outlining what he calls a fiber optic nervous system (FONS) for aircraft. He discusses how fiber optic sensors can be woven into the mesh of composite materials. When damage occurs in the structure, the fiber is broken and does not transmit light anymore, indicating damage. Bennett et al. (1988) discusses how vibration causes interference patterns in the output of a fiber optic sensor. He presents how this fact could be used to detect and measure damage due to impact. Glossop et al. (1990) discusses selection of fiber optic sensor parameters and configurations used to detect damage. Fiber optics, although a very effective sensor, have several drawbacks as well. While their sensitivity is their strength, it is also their weakness. Perturbations in operating conditions greatly affect the signal from the sensor as Bennett showed.

Saravanos et al. (1994) discussed another damage identification method relying solely on a specific type of sensor, in this case, a piezoelectric sensor. He discusses adding a layer of piezoelectric material to a composite structure. The changes in the voltage signature of the piezoelectric material would be indicative of delaminations in

the composite. Saravanos found that adding a layer of piezoelectric greatly increases the weight of structure which could cause problems for certain applications.

The vast majority of the damage identification methods which are not dependent upon the maturity of a specific sensor are dependent upon a mathematical model of the structure. These model-dependent methods usually employ finite-element analyses, modal analysis techniques, or both. Cawley and Adams (1979) created optimism when they attempted to predict the shifts in natural frequencies due to various types of damage. Their predictions were based on a finite-element analysis of an aluminum plate. They showed that the frequency shifts due to significant damage were small, but the prediction of the mathematical model was correct to within 50% of the actual frequency shifts.

This paper spawned a great deal of research based on the use of finite-element models as a base to look for changes in the response characteristics of a structure. Some of this work has included creating a refined finite-element model and fitting a set of experimental data to that model in an effort to identify changes in the mass, stiffness, or damping matrices of the model. Kung et al. (1989) presented a method of developing a finite-element model of a structure and fitting experimentally determined mass, stiffness and damping matrices to it, identifying the minima in these matrices as sites of damage. Kaouk and Zimmerman (1993) discuss using a generalized minimum rank perturbation theory to identify changes in the inertial properties of the finite-element model. They were able to detect the removal of a discrete mass (42% of the mass of the structure) from a cantilevered beam.

Lindner and Kirby (1994) and Lindner et al. (1993) discuss a similar method in which one begins with a refined finite-element model and fits a set of experimental data to it. They define estimated modal stiffness damage coefficients, which are metrics used to measure element-by-element changes in the stiffness matrix of the model. The method was shown to be successful in a numerical example of a cantilevered beam in which the stiffness of an element is reduced by 5%, but has not yet been experimentally verified. Similarly, Tavares et al. (1993) discussed comparing experimental data to a finite-element model using modal assurance criteria (MAC) which are based on the calculation of residual forces. They experimentally analyzed

the effect of different damage states on a truss with respect to the MAC. Campanile (1993) defined an integrity coefficient based on a finite-element model and experimentally determined modal matrices. With this method, he was able to detect a slot (40 - 50% of the thickness of the structure) in a cantilevered beam.

Artificial neural networks have come into use as a computational model consisting of simple processors called nodes with numerous connections between them. Neural networks have primarily been associated with pattern recognition applications. In recent years, neural networks have been applied in conjunction with finite-element models to the task of damage identification. Kudva et al. (1992) used finite-element analysis to generate a set of training data for a neural network. The finite-element model yielded a set of strain measurements on which the neural network was trained. This research consisted of a numerical study only, and indicated some of the problems associated with neural networks. For example, the set of training data must be carefully chosen because too much information and too little information both lead to inaccuracies and even lack of convergence of the network. Elkordy et al. (1993) used finite-element models to generate mode shape data. Changes in the mode shape data due to damage were used as training data for the neural network. When implemented experimentally on a truss, changes were detected but discrepancies between the predicted and actual response (a recurring theme with model-dependence) proved fatal to the method. Rhim and Lee (1994) used finite-element analysis to produce the poles and zeros of a transfer function as training data for use in a neural network. The method is designed to be analytically trained and experimentally implemented. They too discovered the convergence problems of neural networks and presented a numerical simulation only. Most of the literature presented up to this point has had a dependence on a finite-element model to a certain degree. This finite-element model-dependence seems to be inadequate for the task of damage identification. The only successes have come with very detailed models of very simple structures. When determining whether these methods are practical in a realistic setting, we recognize that we will need to perform damage identification on much more complex structures than a cantilevered beam.

Steering clear of finite-element analysis, Wolff and Richardson (1989) and

Richardson and Mannan (1992) have discussed how a model of a structure, consisting of mass, stiffness, and damping matrices can be created by curve fitting the frequency-response function of the structure using modal-analysis techniques. With this, they attempted to determine the existence of damage via changes in the mode shapes or stiffness matrices. Once again, creating a model limits the amount of complexity allowable in order for the method to work. Research has been performed to verify that changes in mode shapes and natural frequencies due to cracks and delaminations are relatively small. Collins et al. (1990) addressed the modelling of a crack as part of a continuum rather than a discrete model and showed that changes in natural frequencies were small (less than 3% change in frequency) for a significant crack. Chen et al. (1993) came to a similar conclusion regarding delaminations. This work shows that any inaccuracies in a model, as when a finite-element model is used, makes it increasingly difficult to predict changes in the dynamic-response of a structure. Because of this, some researchers have concentrated on updating their models until reasonable resolution is achieved.

Glass and Hanagud (1990) and Glass and Macalou (1991) presented a search-based method of model-update using modal-analysis techniques. The method entails defining a model space of differential operators, boundary conditions, and initial conditions. Beginning with a simple model, the method increases the complexity of the model by searching for the best fit of different families of models on the basis of modal-analysis techniques. In order for this method to be successful though, one must have nearly an infinite number of models to search through in order to give the precision necessary to identify damage. Doebling et al. (1993) discusses a method of updating a model by measuring the amount of strain energy contained in the modelled modes. With complex structures however, it becomes increasingly difficult to model enough modes to account for an adequate amount of strain energy. Sensburg and Tomlinson (1993) attempted to update a finite-element model of a structure based on a mathematical optimization method. They found that even for the simplest of structures, the finite-element model lacked precision enough to detect damage as large as notches one quarter of the thickness of a cantilevered beam.

It is becoming apparent that it is unreasonable to assume that one will be able

to model a complex structure with enough precision to be able to discern between changes in the response-characteristics of a structure due to damage from those due to model inaccuracies. For this reason, some researchers have attempted to develop a damage identification method independent of models.

Gomes and Silva (1990) looked at shifts in natural frequencies due to slots milled in beams. They compared these shifts to those from actual cracks and determined that it is acceptable to use slots as a means of simulating cracks. Roitman et al. (1992) used only experimentally measured damping ratios and natural frequencies to attempt to detect damage in a scaled model of an offshore oil rig. They determined that it is difficult to discern changes in these parameters due to damage. Peterson et al. (1993) discuss using an eigensystem realization algorithm which does not implement the use of finite-element methods. However, the end result of this method is a realization of the structure in mass, damping, and stiffness matrices which is, in effect, a model. Yen and Kwak (1993) applied neural networks to a model-independent method of damage identification using differences in sensor output as the training data for the network. They presented a numerical study which showed that with access to a damaged structure, the method was able to discern between damage states of a truss. Manning (1994) proposed a similar method which used experimentally determined poles and zeros of the damaged and undamaged structure to train the network. Once again, a numerical study showed that the method was able to discern between damage states with prior knowledge of damage.

Povich and Lim (1994) actually conducted an experiment using a similar neural network approach as presented by Manning. Using frequency-response function data to train the network, the method was able to determine when a strut was removed from a truss. The authors mention several problems as far as convergence issues of the network are concerned. Tappert and Robertshaw (1994) used numerically simulated correlation coefficients of impulse-response functions in an inductive learning program to detect the existence of a point mass on a plate. This method had difficulty detecting damage on which it was not trained. From the literature presented, it is apparent that attempts to develop a model-independent damage identification method have fallen short of the ultimate goal. The most successful of the model-independent methods

require an a priori knowledge of all possible damage states in order to be effective.

1.3 Objective

There exists a dichotomy in damage identification circles. When research is presented which entails a dependence upon models, inevitably questions will be asked about correlation to actual experimental data and expansion to more complex structures. When research is presented which is independent of models, questions will be asked about the availability of the damaged structure to dynamic testing. Although these questions are legitimate, they have been unproductive and have caused dissension among researchers. The one constant that holds true with any damage identification method is that sooner or later, a dynamic response of a real structure must be gathered experimentally. The questions that need to be asked are: Under what conditions will the dynamic response be acquired? Will variations in these conditions cause variations in the dynamic response as well? The answer is inevitably yes. Actually different parts of the dynamic response are going to be affected more than others. When models or modal parameters are extracted from dynamic-response data, the data are fit to a predetermined model. Whenever data is fit to a model, there will be inaccuracies. With these inaccuracies, crucial information may be lost.

Regardless of the method used to identify damage, there are certain variances inherent to the dynamic response of a structure which must not be ignored. There are variances due to sensor noise, variances due to slight changes in boundary conditions, variances due to manufacturing differences, and variances due to environmental conditions such as temperature, pressure, and humidity fluctuations. *In order to successfully identify damage, we must be able to statistically isolate changes in the dynamic response due to damage from the inherent variances. To this end, we develop a technique which sorts through dynamic-response information, selecting the data in which the changes due to damage are statistically distinguishable from the inherent variances.*

We begin this approach by claiming an a priori ignorance to the manifestations of damage in the dynamic-response characteristics of a structure. Realizing that the

dynamic-response characteristics of a structure are comprised of spatial and temporal information, we gather as much temporal and spatial data as possible and discover empirically how damage manifests itself in the dynamic-response characteristics of the structure. The tool we use to accomplish this is *inductive learning*.

Inductive learning is the process of formulating general rules to represent the relationship between a dependent variable (a.k.a. class) and independent variables (a.k.a. attributes, predictor variables) from a specific set of examples. In other words, inductive learning methods turn a set of examples into an ordered set of rules or code via the use of statistical processes. Inductive learning methods have been used from everything from heart disease prediction to chemical process optimization (Leech, 1986). Essential to the success of inductive learning methods is the definition of an *operational specification* of the task at hand. The operational specification entails definitions of independent variables, dependent variables and examples. The key is to select examples which provide a good representation of the range of the independent variables which in turn, significantly categorize the dependent variable.

The use of inductive learning methods has its strength in determining, from a large collection of independent (predictor) variables, which predictors are the most significant. This allows one to present the inductive learning algorithm with a great amount of data in order to determine the most statistically significant way in which to categorize the dependent variable. This feature is key in applying inductive learning methods to damage identification. As mentioned before, the first task is to determine how damage manifests itself in the response characteristics of a structure. Since the response characteristics of a structure consist of spatial and temporal information, we sift through as much spatial and temporal information as possible with the inductive learning methods in order to determine how the damage manifests itself.

The temporal and spatial information is presented to the inductive learning algorithm via the predictor variables which are functions of the dynamic-response characteristics. As mentioned before, we would like to gather examples which provide a good representation of the range of these predictor variables. The range of the predictor values is determined by the inherent variances of the damage identification problem. That is, we must present the algorithm with any variances that we may

encounter in the testing process. Among these are manufacturing variances, material characteristic variances due to temperature, humidity and pressure, and variances due to sensor noise. Once adequate information is presented to the inductive learning algorithm, it will be able to develop a set of *if...then...* rules which are able to categorize the dependent variable with the most significant predictor variables. The dependent variable, in this case, defines the damage state. It determines whether or not the structure is healthy, and possibly where the damage is, how large the damage is, or what type of damage is present.

1.4 Scope

This thesis describes how a damage identification algorithm is configured using inductive learning. It presents an experimental example which identifies various locations of a small mass (2% of the mass of the structure) on an aluminum plate with free-free boundary conditions, demonstrating the proposed method. Ideally, this method would be applied to a structure for which there are a finite number of expected damage states which are available for dynamic testing. Although this availability to dynamic testing is ideal, it is by no means necessary. However, the more information available, the more versatile and successful the method will be.

It was mentioned earlier that a damage identification method incorporating the use of inductive learning statistically isolated changes in the dynamic response due to damage from the inherent variances of the problem. In the example presented in this thesis, we consider only inherent variances due to sensor noise and slight changes in boundary conditions. The testbed for this experiment is a $12'' \times 12'' \times \frac{1}{8}''$ aluminum plate hung horizontally from the corners with fishing line to simulate free-free boundary conditions. The plate is sensed and actuated by two piezoelectric patches mounted diagonally symmetric from one another.

While the method presented is independent of sensor type and dynamic-response type gathered, a Hewlett-Packard impedance analyzer is used in the example to measure the structural impedance-response for a range of frequencies, using each piezoelectric as a sensor/actuator. Changes in the physical properties are induced on

the plate via the addition of the small test mass. The test mass was placed at one of four different locations on the plate which are referred to henceforth as damage states A through D. Each of these damage states can be thought of as an expected damage state that is available for dynamic testing.

The next step in the process was to define the ‘operational specification’ of the problem. The dependent variable is defined as the damage state. The dependent variable has six possible outcomes: No damage, damage state A-D, or damage state of unknown type. The independent or predictor variables are functions of the impedance-response of the plate. The examples gathered should give a good representation of the range of the independent variables based on the assumptions of the inherent variances involved. For this reason, for each damage state, the impedance-response was gathered only after the testing configuration was disassembled and re-assembled, to take into account the variation in boundary conditions.

We mentioned the importance of gathering as much temporal and spatial data as possible. Because of this, we make use of all of the spatial data collection possibilities we have, which is the use of both sensor/actuators to measure the real and imaginary parts of the impedance-response for a range of 4200-5000 Hz, in 1 Hz intervals. We then gather ten examples of each damage state as defined above.

In order to form the independent (predictor) variables, we perform statistical manipulations of the impedance-response information. This entails binarily comparing each damage state against every other damage state on the basis of a statistical test of hypothesis. This statistical test of hypothesis yields a metric as to how well we can distinguish the damage states from one another. From this information, we are able to determine with which sensor, frequency, and data type (real/imaginary) we need to measure in order to distinguish between damage states. These statistical manipulations are detailed in chapter 2.

Once we have determined what information we need in order to categorize the dependent variable, we extract only this information from the impedance-response, and apply it to a set of statistically generated rules. This enables us to greatly reduce the computation time in order to make the method effective in an on-line setting. Chapter 3 discusses the experimental process in more detail. Chapter 4 gives

a summary of the results of the experiment and chapter 5 discusses the significance of what has been accomplished by this research.

Chapter 2

Statistical Approach

“Variability in measurements and processes is a fact of life.”

- Hogg and Ledolter

2.1 Variability

Variability is a dispersion in measurements. Hogg and Ledolter (1992) discuss variability as either being *measurement variability* or *process variability*. Given that the measuring device is sufficiently accurate, process variability results from any number of things. For a manufacturing process, this process variability results from differences in raw materials, differences in the machines used to create the product, differences in the habits of the operators creating the product, and differences in the environmental conditions when the product is made. Measurement variability results from inaccuracy of the measuring instrument, the environmental conditions while the measurement is made, and the way the measurement is set up.

When using a statistical approach to identify damage, we similarly can describe the variability of a dynamic response of a structure as being composed of process variability and measurement variability. Process variability results from the manufacturing differences discussed above. Measurement variability results from the method of acquiring the dynamic response of the structure; When a dynamic-response is acquired, be it an impulse-response, step-response, frequency-response, impedance-response, etc..., there are certain noises in the hardware and inaccuracies caused by the algorithms which calculate the responses. Measurement variability also results

from differences in environmental conditions when the measurement is made; Humidity, temperature, and pressure all can have an effect on the response characteristics of the structure and sensors. Finally, measurement variability comes from differences in the set up of the measurement; This fact is predominantly apparent in the effect that slight changes in boundary conditions have on the response characteristics of a structure.

2.2 Distributions

Statistical methods help us deal with the forementioned variabilities, inducing information about a process from an incomplete set of data. To this end, graphical and numerical methods of representing data have been developed. Consider a measurement. This measurement is subject to the measurement variability and the process variability we discussed above. We call this measurement a *random variable*, X . When a single measurement is made, it is denoted as a *random sample*, x , of the random variable, X . Statisticians have developed methods of describing the dispersion of the random samples due to the variability. For continuous random variables (i.e. the random samples fall in a continuum), we make use of a *probability distribution function* (p.d.f.), $f(x)$. The area beneath the p.d.f. is unity; that is $\int_R f(x)dx = 1$, where R is the outcome space of the random variable, X . The probability that a random sample, x , will be on the interval $[a, b]$ is defined by:

$$P(a \leq X \leq b) = \int_a^b f(x)dx \quad (2.1)$$

Fig. 2.1 shows a graphical representation of a normal or gaussian p.d.f.

The mathematical expectation of a random variable is expressed as:

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x)dx \quad (2.2)$$

This is what is normally referred to as the *mean*, μ , of the random variable, X . There are a number of ways to describe how the random samples are dispersed about this mean. The most common way to accomplish this is to define the *variance*, σ^2 , of a random variable, X , expressed mathematically as:

$$\sigma^2 = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx \quad (2.3)$$

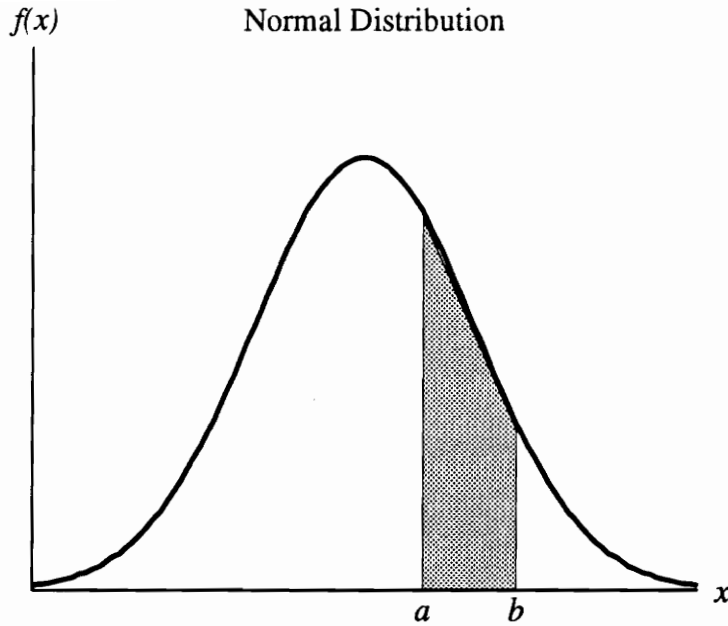


Figure 2.1: Probability density function $f(x)$; shaded area represents $P(a \leq X \leq b)$

$$= \int_{-\infty}^{\infty} x^2 f(x) dx - \mu^2 = E(X^2) - \mu^2 \quad (2.4)$$

The square root of the variance is commonly referred to as the *standard deviation*, σ . Once we have definitions of mean and variance, we can express the p.d.f. of a random variable in these terms. The distribution that we will be predominantly dealing with is the *normal* or *gaussian distribution*. Given a random variable, X , which has a normal distribution with mean, μ and variance, σ^2 , we can describe its p.d.f with the following equation:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x - \mu)^2}{2\sigma^2}\right], \quad -\infty < x < \infty. \quad (2.5)$$

2.3 Data representation

Although it would be nice to have precise knowledge of the mean, variance and p.d.f. of a random variable, in practice, this is not the case. Almost without exception, we

are forced to assume a p.d.f. and estimate values of mean and variance in order to be able to induce information about the random variable.

We begin by making an estimate of the mean, μ , of a random variable, X . Consider once again a measurement which we denote a random variable, X , with a mean, μ , and a variance, σ^2 . If we take n measurements, we get the data set $\{x_1, x_2, \dots, x_n\}$. The estimate of the mean of the random variable is called the *sample mean*, \bar{x} , and is calculated as follows:

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n) = \frac{1}{n} \sum_{i=1}^n x_i. \quad (2.6)$$

It is apparent that as we increase the number of samples (observations or measurements), n , the value of the sample mean, \bar{x} , approaches that of the true mean, μ .

$$\lim_{n \rightarrow \infty} \bar{x} = \mu. \quad (2.7)$$

In a similar fashion, we can make an estimate of the variance, σ^2 of a random variable, X . This estimate is called the *sample variance*, s^2 , and it is calculated as follows:

$$s^2 = \frac{1}{n-1}[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2] = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2.8)$$

Once again, it is apparent that as the number of samples, n , is increased, the value of the sample variance, s^2 , approaches that of the true variance, σ^2 .

$$\lim_{n \rightarrow \infty} s^2 = \sigma^2. \quad (2.9)$$

As an example, let us consider measuring the dynamic response of a structure. This dynamic response can be an impulse-response, step-response, frequency-response, impedance-response, etc. . . In this case, we will consider an impedance-response. This impedance-response contains an imaginary part and a real part. For each of these parts of the response, there are data points corresponding to the magnitude of the impedance response at specific frequencies. We can treat each of these data points as a random sample, x , of a random variable, X . For instance, if we measure the impedance response of a structure from 4600 to 5000 Hz with a single data

point every 1 Hz, we have 401 data points for the real part and 401 data points for the imaginary part. This corresponds to a single random sample of 802 independent random variables.

As mentioned before, this measurement of the impedance response is subject to variability. For this initial example, let us assume that variability is dominated by the sensitivity to boundary conditions. This means that the dynamic response varies slightly depending on the set up of the testing apparatus. In order to develop an understanding of this process, we disassemble and reassemble the testing apparatus, measuring the impedance response each time. With this, we gather random samples of the 802 independent random variables we defined above. We can better approximate the means and the variances of our 802 random variables as we gather more and more random samples.

Fig. 2.2 shows the 10 random samples we gathered, representing the imaginary part of the impedance-response of our structure. Each dot represents one random sample of one random variable. As sections of Fig. 2.2 are blown up as in Fig. 2.3, we can show more clearly the distribution of the random samples. The x's in the figure represent the sample means calculated from Eq. 2.6 and the asterisks represent the sample variances calculated from Eq. 2.8. It is very important to note the differences in the scales of the subplots. We can clearly see that the random variables representing the impedance-response of the structure at this particular resonance have a much higher sample variance than those representing a relatively inactive area of the impedance response. This fact will play a crucial role in damage identification using inductive learning methods. We now have 802 random variables, each with a sample mean and a sample variance calculated from 10 random samples. In order to go any further, we need to know how reliable these estimates of the mean and variance are to the true values.

2.4 Distribution of a Sample Mean

Assume we have a linear function of n mutually independent random variables, $Y = \sum_{i=1}^n a_i X_i$. If μ_i is the mean and σ_i^2 of the random variable, X_i , for $i = 1, \dots, n$, we

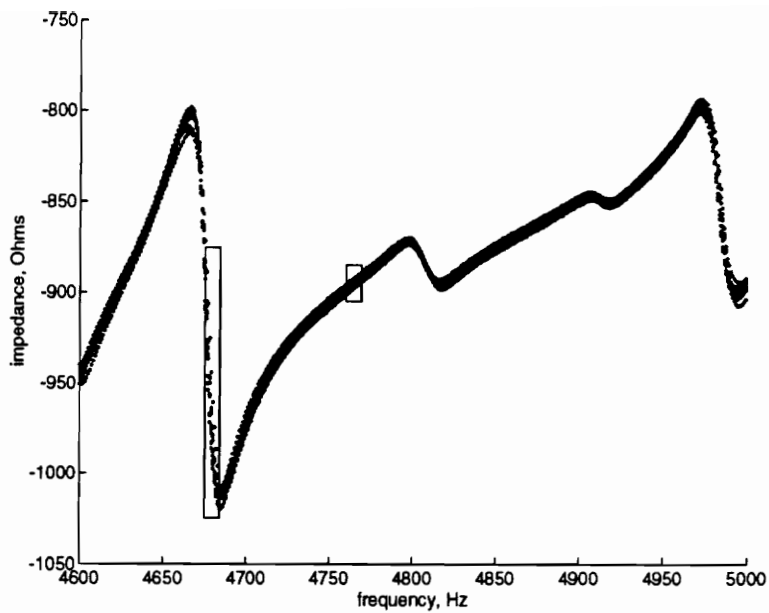


Figure 2.2: Random samples of imaginary part of impedance-response.

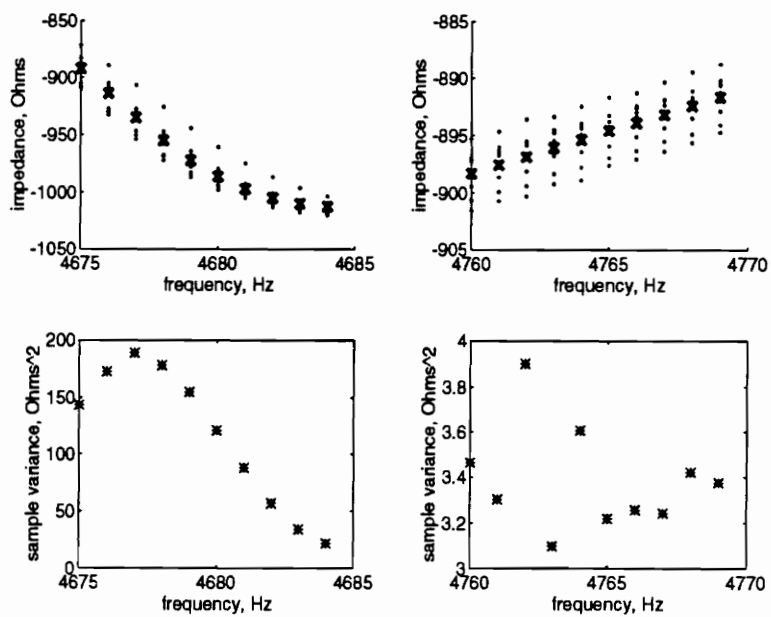


Figure 2.3: Sample means and variances of imaginary part of impedance-response.

can show that the mean and variance of Y can be calculated as follows:

$$\mu_Y = \sum_{i=1}^n a_i \mu_i \quad (2.10)$$

$$\sigma_Y^2 = \sum_{i=1}^n a_i^2 \sigma_i^2. \quad (2.11)$$

Referring back to our impedance-response example, we had calculated a sample mean of 802 random variables. The sample mean can be thought of as a random variable which is a linear combination of random samples. That is for n random samples, the sample mean is calculated as follows:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \left(\frac{1}{n}\right) X_1 + \left(\frac{1}{n}\right) X_2 + \dots + \left(\frac{1}{n}\right) X_n \quad (2.12)$$

This corresponds to a linear combination of X_1, X_2, \dots, X_n , all with weights, $a_i = 1/n$. We have already shown how to calculate the mean and variance of a linear combination of random variables. Performing this operation, we can show:

$$\mu_{\bar{X}} = \sum_{i=1}^n \left(\frac{1}{n}\right) \mu_X = \mu_X \quad (2.13)$$

$$\sigma_{\bar{X}}^2 = \sum_{i=1}^n \left(\frac{1}{n}\right)^2 \sigma_X^2 = \frac{\sigma_X^2}{n} \quad (2.14)$$

From these equations, we know how to calculate the mean and the variance of \bar{X} . In order to get a feel for how the random variable, \bar{X} is distributed, that is, determine the p.d.f of \bar{X} , we invoke the central limit theorem.

Central Limit Theorem 1 *If \bar{X} is the mean of a random sample X_1, X_2, \dots, X_n from a distribution with mean μ and finite variance $\sigma^2 > 0$, then the distribution of*

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{\sum X_i - n\mu}{\sigma\sqrt{n}}$$

approaches a distribution that is normal with a mean of 0 and a variance of 1, $N(0, 1)$, as n becomes large.

This theorem states that for a sufficiently large sample size, n , we can approximate the distribution of \bar{X} to be normal with mean, μ , and variance, σ^2/n , that is $N(\mu, \sigma^2/n)$.

2.5 Confidence Intervals for Means

Now that we have an approximation for the distribution of \bar{X} , we can make statements as to how reliable an estimator \bar{X} is of the mean, μ of the random variable.

In stating the central limit theorem, we defined a new random variable, Z . The purpose in doing this is to be able to represent all normal distributions by the same function. Z is a random variable with a standard normal distribution; that is, it has a normal distribution with mean of 0 and a variance of 1. Fig. 2.4 graphically shows the standard normal distribution. The integral function, $\Phi(z)$, represents the probability that a random sample of Z will be less than a specified value, z . The following equations hold true:

$$P(Z \leq z) = \Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-w^2/2} dw \quad (2.15)$$

$$\Phi(-z) = 1 - \Phi(z) \quad (2.16)$$

With this knowledge of the standard normal distribution function, we can take a probability defined as $1 - \alpha$ and find a value $z(\alpha/2)$ that satisfies the following equation:

$$P \left[-z(\alpha/2) \leq Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq z(\alpha/2) \right] = 1 - \alpha \quad (2.17)$$

In other words, we can specify an α , say 0.05, and use the standard normal distribution function to find values of $z(\alpha/2)$ with which there is a probability of $1 - \alpha = 0.95$ that a random sample of Z falls between $-z(\alpha/2)$ and $z(\alpha/2)$.

This probability equation simplifies by recognizing the fact:

$$P[Z \leq -z(\alpha/2)] = P[Z \geq z(\alpha/2)] = 1 - \alpha \quad (2.18)$$

Performing a few algebraic manipulations, the equations yield an inequality involving μ :

$$\bar{X} - z(\alpha/2) \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z(\alpha/2) \frac{\sigma}{\sqrt{n}} \quad (2.19)$$

Referring back to our impedance response example, we gathered 10 random samples of 802 random variables and calculated a sample mean, \bar{x} , for each. These equations

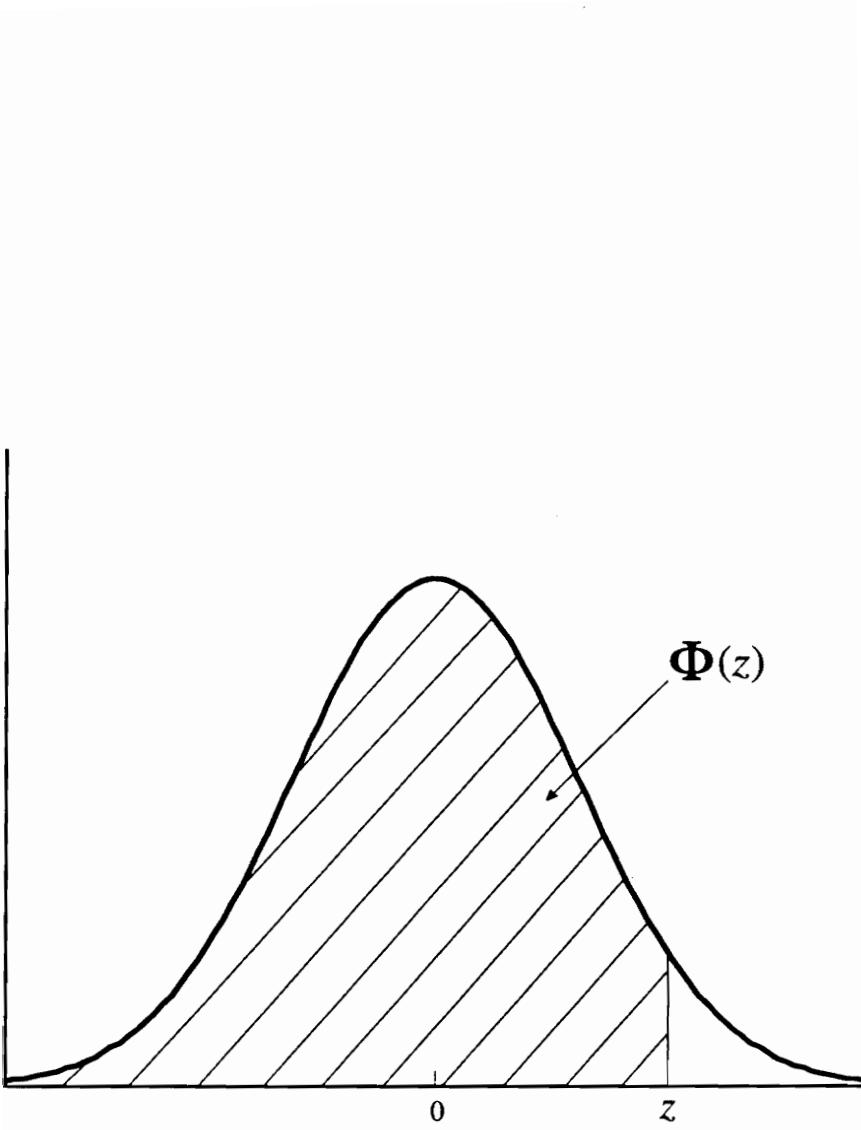


Figure 2.4: Standard Normal Distribution

state that if we knew the true variance, σ^2 , we could say that we were $100(1 - \alpha)$ percent confident that the true mean, μ , lies on the interval:

$$\left[\bar{x} - z(\alpha/2) \frac{\sigma}{\sqrt{n}}, \bar{x} + z(\alpha/2) \frac{\sigma}{\sqrt{n}} \right]. \quad (2.20)$$

The problem is that we do not have a precise knowledge of the true variance, and are forced to make an estimate.

2.6 Dealing with Unknown Variances

When we discussed the distribution of the sample mean, \bar{X} , we defined a new random variable, Z :

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \quad (2.21)$$

We similarly define a new random variable, T , in order to deal with the distribution of the sample variance, S^2 , which can be treated as a random variable as well:

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \quad (2.22)$$

This random variable, T , has its own probability distribution function called the *Student's t-distribution*, $h(t)$:

$$h(t) = \frac{c}{(1 + t^2/r)^{(r+1)/2}}, \quad -\infty < t < \infty \quad (2.23)$$

In this equation, $r = n - 1$, indicates the number of degrees of freedom, and c is selected in order to make the area under $h(t)$ be 1. Fig. 2.5 shows the p.d.f. of a Student's t-distribution. We can see that the shape is very similar to that of the normal distribution, but the number of degrees of freedom, r , affects the 'skinniness' of the p.d.f.

In a similar fashion to the previous section, we can set up a probability function involving the true mean, μ :

$$P \left[\bar{X} - t(\alpha/2; n - 1) \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + z(\alpha/2; n - 1) \frac{S}{\sqrt{n}} \right] = 1 - \alpha. \quad (2.24)$$

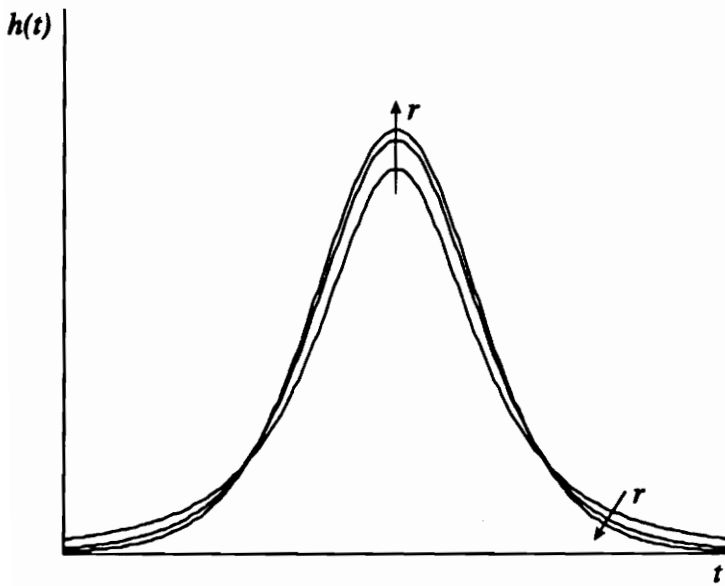


Figure 2.5: Student's t-distribution

Once again, we can say that if we take n random samples of a random variable, X , calculate a sample mean, \bar{x} , and a sample variance, s^2 , we are $100(1 - \alpha)$ percent confident that the true mean, μ lies on the interval:

$$\left[\bar{x} - t(\alpha/2; n - 1) \frac{s}{\sqrt{n}}, \bar{x} + t(\alpha/2; n - 1) \frac{s}{\sqrt{n}} \right]. \quad (2.25)$$

We finally have an equation that we can apply to our impedance-response example. We have already calculated the sample mean and sample variance of the 802 random variables. Applying Eq. 2.25, we set $\alpha = .005$ and find the 99% confidence interval for each true mean.

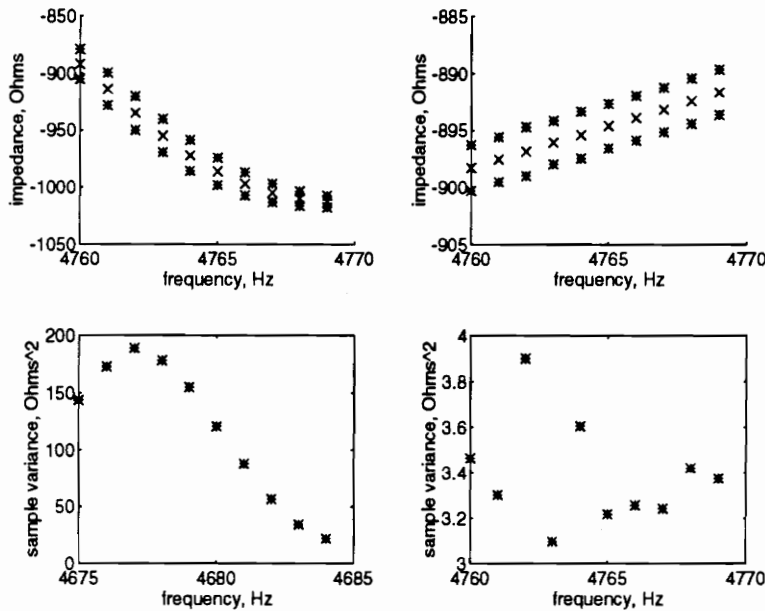


Figure 2.6: Confidence Intervals for Means

Fig. 2.6 shows the 99% confidence intervals for the true means of a select few random variables in the impedance-response example. Once again, it is important to note the scales of the subplots. We can see that as the sample variance decreases, the *confidence bounds* narrow. This fact plays a crucial role in the monitoring of the

health of a structure. If we are to detect changes in the response characteristics of a structure due to damage, we need to introduce a method of quantifying differences in the dynamic-response data. To this end, we employ a statistical test of hypothesis.

2.7 Statistical Test of Hypothesis

Consider that we have two independent random variables, X and Y , with respective means, μ_X, μ_Y , and variances, σ_X^2, σ_Y^2 . Referring back to our impedance-response example, these two random variables could be the magnitude of the imaginary part of the impedance-response at a specific frequency; one random variable is of a healthy structure while the other is from a damaged structure. With the boundary condition dominance assumption, we gather n_X random samples from the healthy structure and n_Y random samples from the damaged structure. We have already shown that the distribution of \bar{X} is normal with mean, μ_X and variance, σ_X^2/n_X , and that the distribution of \bar{Y} is normal with mean, μ_Y and variance, σ_Y^2/n_Y . At this point, we make a null hypothesis, H_0 ; that is we assume a statement about the relative values of the means of our two random variables. In this case, we assume that the means are equal to one another, $H_0 : \mu_X = \mu_Y$ or $\mu_X - \mu_Y = 0$. With this null hypothesis and assuming that we know the variances, we can define a new random variable which is the difference between the two sample means:

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\sigma_X^2/n_X + \sigma_Y^2/n_Y}} \quad (2.26)$$

This new random variable, Z , has a normal distribution with mean, $\mu = 0$, and variance, $\sigma^2 = 1$. If we were conducting a typical hypothesis test, we would compare the value of this random variable to a prespecified value, $z(\alpha)$. The α in this case is what we call a significance level of the hypothesis test. We would typically test the null hypothesis, H_0 against an alternate hypothesis, H_1 . For example, if we thought that μ_X may be greater than μ_Y , if $Z \geq z(\alpha)$, we could say with $100(1 - \alpha)\%$ confidence that $\mu_X > \mu_Y$. If we just wanted to know if the two means were not equal, if $|Z| \geq z(\alpha/2)$, we could say with $100(1 - \alpha)\%$ confidence that $\mu_X \neq \mu_Y$. It is a variation on this approach that we will use to discriminate between damage cases.

This approach, however, assumes that we have a precise knowledge of the variances of the two random variables.

In order to perform a hypothesis test with sample variances in lieu of true variances, we must make assumptions about the distributions of the random variables. We assume that the distributions of the random variables are normal and that the variances are about equal. With the assumption that the variances are about equal, we can obtain an estimate of this common variance, S_p^2 , which we call our pooled variance:

$$S_p^2 = \frac{(n_X - 1)S_X^2 + (n_Y - 1)S_Y^2}{n_X + n_Y - 2} \quad (2.27)$$

Stating the null hypothesis to be $H_0 : \mu_X = \mu_Y$, we define a new random variable:

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{S_p^2 \left(\frac{1}{n_X} + \frac{1}{n_Y} \right)}} = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{(n_X - 1)S_X^2 + (n_Y - 1)S_Y^2}{n_X + n_Y - 2} \left(\frac{1}{n_X} + \frac{1}{n_Y} \right)}} \quad (2.28)$$

This random variable, T , has a Student's t -distribution with $n_X + n_Y - 2$ degrees of freedom. Once again, if our alternate hypothesis were, $H_1 : \mu_X \neq \mu_Y$, we could say with $100(1 - \alpha)\%$ confidence that the means are not equal if $|T| \geq t(\alpha/2; n_X + n_Y - 2)$.

We already expect the means of our two random variables to be different. This is the assumption that we based this approach on: Changes in the physical properties of a structure cause changes in the dynamic response characteristics of the structure. What we need is a metric to indicate how well we can discriminate between dynamic responses of different damage states. The answer is in the definition of the new random variable, T . Instead of comparing the value of T to a cutoff value of an alternate hypothesis, we take the inverse t -distribution function of T , that is $t^{-1}(T)$. The result of this is the probability that the means of the two random variables are equal. This can be shown graphically as in Fig. 2.7. We have already discussed that the p.d.f. of the true mean with a given sample mean and sample variance has a Student's t -distribution centered about the sample mean. If we are comparing two random variables, as in Fig. 2.7, taking the inverse of the t -distribution function yields a number, $\alpha = t^{-1}(T)$, which turns out to be the intersection of the two probability distribution functions of the means of the two random variables. Once we have solved for the significance level, α , we can say with $100(1 - \alpha)\%$ confidence that the means

of the two random variables are not equal. Therefore, α is a measure of how well we can distinguish the random variables from one another. As the significance level decreases, the random variables become more easily distinguishable.

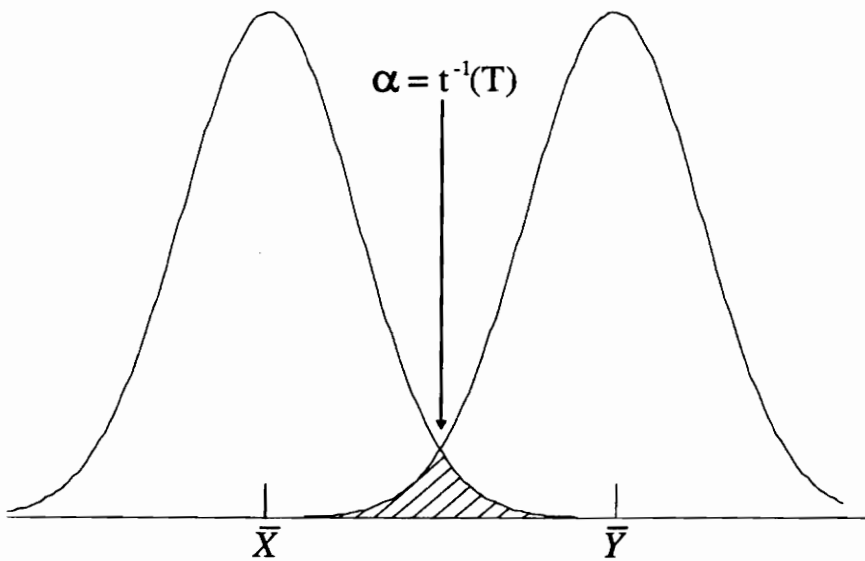


Figure 2.7: Hypothesis test of two means, $H_0 : \mu_X = \mu_Y$

2.7.1 Verification of Equal Variance Assumption

Before we can calculate this significance level, α , we must verify the assumptions we made in defining the new random variable, T . To this end, we define a new random variable:

$$F = \frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2} \quad (2.29)$$

This random variable, F , can be shown to have an F -distribution. The F -distribution has two parameters which represent the number of degrees of freedom of the statistical model. These degrees of freedom are based on the number of random samples of the two compared random variables. Therefore F has an F distribution with $n_X - 1$ and

$n_Y - 1$ degrees of freedom. In order to determine whether the variances are about equal, we perform a hypothesis test. If we claim $H_0 : \sigma_X^2 = \sigma_Y^2$, as our null hypothesis, then $F = S_X^2/S_Y^2$ has an F -distribution with $n_X - 1$ and $n_Y - 1$ degrees of freedom. Considering the alternate hypothesis, $H_1 : \sigma_X^2 > \sigma_Y^2$, if $S_X^2/S_Y^2 \geq F(\alpha; n_X - 1, n_Y - 1)$, we can say with $100(1 - \alpha)\%$ confidence that $\sigma_X^2 > \sigma_Y^2$.

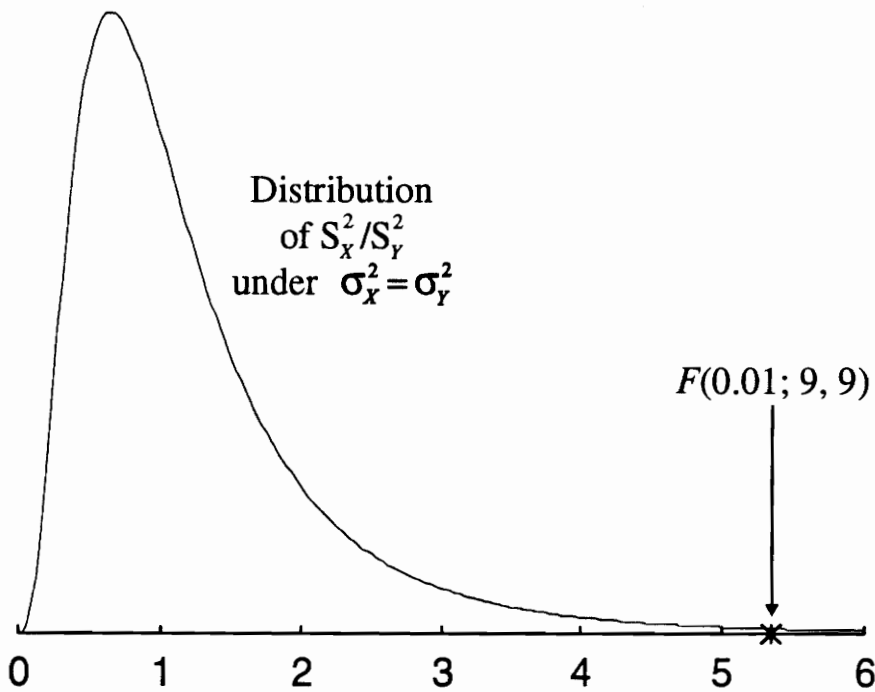


Figure 2.8: Graphical description of the test $H_0 : \sigma_X^2 = \sigma_Y^2$ against $H_1 : \sigma_X^2 > \sigma_Y^2$

Referring back to our impedance-response example, we took n_X random samples of the healthy structure and n_Y random samples of the damaged structure. Under the null hypothesis, $H_0 : \sigma_X^2 = \sigma_Y^2$, S_X^2/S_Y^2 has an F distribution with $n_X - 1$ and $n_Y - 1$ degrees of freedom. Supposing $n_X = n_Y = 10$, we can set a maximum limit for this ratio by selecting a significance level, α , of our hypothesis test. In this case, we select $\alpha = .01$ and find $F(0.01; 9, 9) = 5.25$. As long as the ratio of the sample variances does not exceed this number, we accept the null hypothesis, that is $H_0 : \sigma_X^2 = \sigma_Y^2$. This is shown graphically in Fig. 2.8.

2.7.2 Verification of Normality Assumption

The other assumption that we must verify in order to use Eq. 2.28 is that the random variables in question have normal distributions. We once again use a statistical test of hypothesis in order to determine whether a random variable has a certain distribution, in this case, a normal distribution.

The first step is to divide the outcome space, R , of the random variable, X into k exhaustive and exclusive cells. Each one of these cells has a probability associated with it, based on the normality assumption, that a random sample will land there. We call these probabilities, p_1, p_2, \dots, p_k , where $\sum_{i=1}^k p_i = 1$. Considering that we have n random samples of our random variable, X , the number of observations that land in the cells are denoted as Y_1, Y_2, \dots, Y_k , where $\sum_{i=1}^k Y_i = n$. We can also talk about the expected number of observations in a particular cell based on the probability in each cell and the total number of observations, that is $E(Y_i) = np_i$.

With this, we can define a new random variable:

$$Q_{k-1} = \sum_{i=1}^k \frac{(Y_i - np_i)^2}{np_i} \quad (2.30)$$

This new random variable, Q_{k-1} has what is called a chi-square, χ^2 distribution with $k-1$ degrees of freedom. In determining the probabilities within the cells, we calculate a sample mean, \bar{x} , and a sample variance, σ^2 . The calculation of these two parameters takes away two degrees of freedom from the χ^2 distribution, therefore, Q_{k-1} has a χ^2 distribution with $k-3$ degrees of freedom.

Just as before, we have a null hypothesis, H_0 , which is that the distribution of X is normal. The alternate hypothesis, obviously, is that X is not normally distributed. We once again select a significance level, α of our test. If the calculated random variable, Q_{k-1} , exceeds $\chi^2(\alpha; k-3)$, then we reject the null hypothesis.

Referring back to the impedance-response example, we had $n = 10$ observations of the random variable, X . We could split the outcome space into $k = 5$ different cells. We then calculate the probability of a random sample occurring in the cells by using the cumulative distribution function of a normal distribution with mean, \bar{x} and variance, s^2 . The random variable, Q_4 is then calculated according to Eq. 2.30. As

long as $Q_4 \leq \chi^2(.01; 2) = 9.2$, we cannot reject the normal distribution assumption.

2.8 Rule Generation

In this chapter, we have discussed the statistical principles necessary to understand how variability, inherent in measurement processes, is dealt with. More importantly, we have presented a metric to measure the degree to which we can discriminate between random samples of two independent distributions. The final step in the process is to develop a set of rules which would be applied to random samples in order to discriminate between two independent distributions.

It is common knowledge that with a normal distribution, 99% of the random samples will fall within the interval $\mu \pm (2.576)\sigma$. Unfortunately, in practice, we seldom know μ and σ and must approximate them with \bar{x} and s . However, $\bar{x} \pm (2.576)s$ does not necessarily contain 99% of the random samples. We need to find a value, k , such that the probability of the random interval $\bar{X} \pm kS$ containing 99% of the random samples is large. Therefore, we select a significance level, α , with which we can say we are $100(1 - \alpha)\%$ confident that 99% of the random samples will be contained in the interval $\bar{x} \pm ks$. Statisticians have worked on this problem and they have developed tables which cross-index a percentage of samples, p , which is .99 in our case; a significance level, α , and a number of observations. The output of the table is a value, k .

Referring once again to our impedance-response example, we had $n = 10$ observations. Suppose we wanted to be $100(1 - \alpha) = 99\%$ confident that 99% of the random samples are contained in the interval $\bar{x} \pm ks$; we look up on the chart for values of $p = .99$, $\alpha = .01$, and $n = 10$, to find that $k = 5.594$.

If the hypothesis test had determined that we could easily distinguish between a specific pair of random variables, we would apply this value of k and create intervals which represent rules for random variables. With this, the groundwork is laid for damage identification using inductive learning. We present the inductive learning algorithm with as many independent random variables as possible. The algorithm uses the hypothesis test to determine with which random variable we can most easily

discriminate between damage states. We then make use of the k value to create rules to which we apply a random sample of an impedance-response to identify damage.

Chapter 3

Experimental Verification

“Every experiment is like a weapon which must be used in its particular way—a spear to thrust, a club to strike. Experimenting requires a man who knows when to thrust and when to strike, each according to need and fashion.”

– Philippus Aureolus Paracelsus

3.1 Testbed

The testbed for the experiment used to verify the method is a $12'' \times 12'' \times \frac{1}{8}''$ aluminum plate. Each corner of the plate has a $\frac{1}{32}''$ hole through which fishing line is tied. Each fishing line strand is 4'' long and tied to a paper clip. The paper clips are hooked on eye screws mounted on a wooden frame. With this, the plate hangs in a horizontal position from the four fishing line strands, simulating free-free boundary conditions.

The plate is actuated and sensed with two piezoelectric ceramics which we denote ‘green’ and ‘red’ according to the color of the lead attached to its electrode. The sensoriactuators are located diagonally symmetric from one another. In order to change the physical properties of the plate, we place a small test mass on the plate. We designate four possible locations of a small test mass to be placed on the plate which define ‘damage’ cases. This test mass is 2% of the mass of the plate itself. The possible locations of the test mass and the locations of the piezoelectrics are shown in Fig. 3.1.

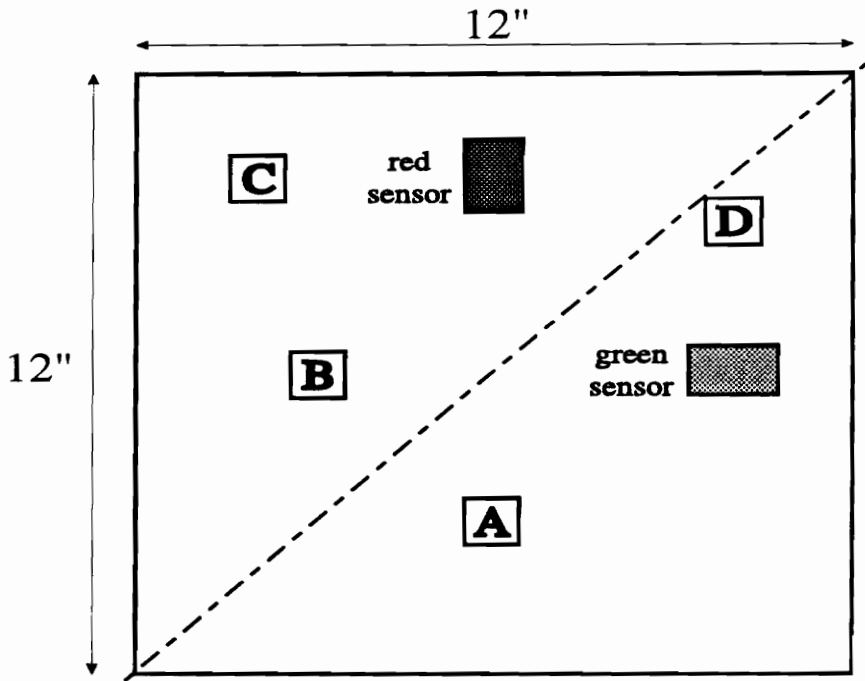


Figure 3.1: Piezoelectric and test mass locations

3.2 Data Acquisition

While the proposed method is independent of sensor type and dynamic-response type, we choose to use a frequency swept structural impedance-response for this experiment because of its reputation for repeatability. The equipment used to measure the structural impedance is a Hewlett-Packard 4194-A Impedance Gain/Phase Analyzer. In making this measurement, one lead is hooked up to the electrode on the sensor/actuator and one lead is grounded to the plate itself with copper tape. In measuring the impedance-response, the analyzer sends a one volt peak-to-peak sinusoidal signal to the piezoelectric and measures the current across the piezoelectric, yielding the impedance response. With this impedance analyzer, one can control the integration time as well as the number of averages. These two parameters were selected in order to minimize a phenomenon known as 'ring down'. Ring down happens when frequencies are swept too fast, causing the vibratory effect at one frequency to bleed into the next, corrupting the impedance data which is being measured. We select a 'medium'

integration time and 16 averages in order to sufficiently minimize 'ring down'.

The impedance analyzer can give output in either real/imaginary or magnitude/phase. We choose to use real/imaginary data, although there is no reason why the method would not work for magnitude/phase data as well. The next step is to select a frequency range and spacing. The analyzer only has the capability of measuring 401 frequency points per sweep. This limits the ease in acquiring the great amount of data needed for this method to be successful. However, we decided to select a range of 800 Hz with 1 Hz spacing. Using this range and spacing, we strike a balance of gathering enough data for the method to work and not gathering too much data as to make the data acquisition cumbersome. The selection of which frequency range to implement came from preliminary tests of the method in which a much broader frequency range was used to distinguish between two damage states. The frequency range was narrowed down to 4200 Hz to 5000 Hz after it was shown to be a region which yielded several significant predictor variables.

The selection of the number of examples needed is a very nebulous subject. Obviously, the more examples we have, the more confident we are about our estimates of the mean and variance of a measurement. The question is how do we know we have enough examples for the method to work? There really is no easy answer to this question. Hogg and Ledolter (1992) discuss a number of ways to determine the number of examples needed for certain confidences on means and variances but these are usually based on the assumption that one knows either the mean or the variance. For this experiment, we made estimations based on measurements made in previous experiments and determined that ten examples should be satisfactory to get the accuracy we desired.

In order to determine what comprises an example, we must make assumptions as to what variances in the acquisition of the dynamic-response we anticipate. In this experiment, we ignore effects of manufacturing differences because we are using a single structure. We also assume that the environmental conditions are constant so that we need not formulate them into our problem. What we have left is variances due to sensor noise and variances due to changes in boundary conditions which come from slight differences in the testing setup. The examples should yield a reasonably

accurate representation of the range of the independent variables, which are functions of the dynamic-response. Therefore, before each impedance-response is measured, the testing apparatus is disassembled and reassembled, to give us a good representation of the variance due to changes in boundary conditions. The disassembly requires the leads being removed from the impedance analyzer and the plate being removed from its wooden frame. Sensor noise is inherent in the collection of multiple examples.

So for each damage state ('no damage' and damage states A-D), we disassemble and reassemble the testing apparatus ten times, measuring the impedance response from 4200 Hz to 5000 Hz in 1 Hz increments, using both the red and the green sensors after each reassembly. From this data gathering scheme, for each damage state we have ten random samples of $801 \text{ (data points)} \times 2 \text{ (sensors)} \times 2 \text{ (real/imaginary)} = 3204$ random variables. It is via these 3204 random variables that we will distinguish the damage states from one another.

3.3 Data Manipulation

The data sets are exported from the impedance analyzer to a PC and stored in `-.dat` files. These `-.dat` files are 401 lines long and have three columns: frequency, real part of the impedance-response, and imaginary part of the impedance-response. These data sets are then imported into MATLAB and saved as `-.mat` files using `loadit.m`, which is located in Appendix A as are all `-.m` files mentioned here.

For each combination of damage state, sensor, and data type (real/imaginary), we create a matrix. That is we create $5 \text{ (damage states)} \times 2 \text{ (sensors)} \times 2 \text{ (data types)} = 20$ matrices. The rows in these matrices indicate single examples and the columns indicate frequency. Therefore, we create 20 $[10 \times 801]$ matrices using `pairit.m`. As an example, Fig. 3.2 and 3.3 plot two of these matrices, showing the ten random samples for the random variables comprising the 'no damage' case using the green sensor for both the real and the imaginary parts of the impedance-response.

Each column of each matrix represents a single random variable. For each column, we can calculate a sample mean and a sample variance. This is also accomplished using `pairit.m`. Fig. 3.4 and Fig. 3.5 show the sample means and sample

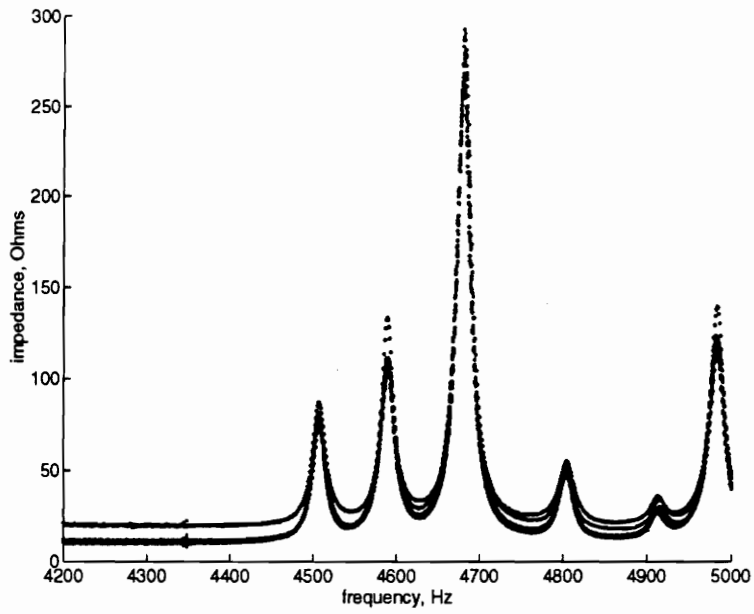


Figure 3.2: Ten examples of the 'no damage' case (real)

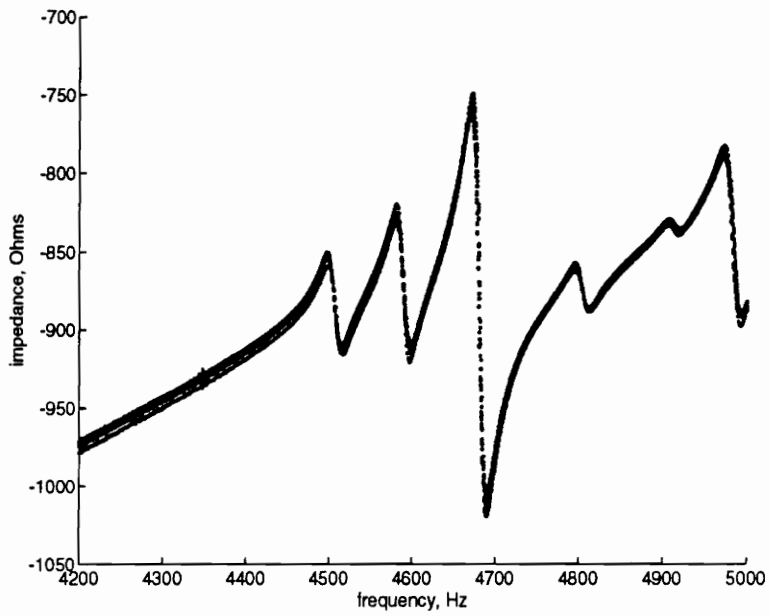


Figure 3.3: Ten examples of the 'no damage' case (imaginary)

variances of the random variables shown in Fig. 3.2 and Fig. 3.3.

For each damage state, there are 3204 random variables with which we can compare to the corresponding random variables of the other damage states. At this point, we employ the use of a statistical test of hypothesis in order to get a measure of how confident we are that the means of the two random variables we are comparing are different. With this, we compare all of the random variables of each damage state with their corresponding random variables in the other damage states. The number of comparisons is given by the following equation:

$$N = r \sum_{i=1}^{n-1} i \quad (3.1)$$

In this equation, r is the number of random variables and n is the number of damage states. For this example, we need to make $3204 \sum_{i=1}^4 i = 32,040$ comparisons.

Before performing these statistical tests of hypothesis, we must verify the underlying assumptions we discussed in chapter 2. If the sample distribution of a random variable is shown not to be normal, any statistical test of hypothesis with using this random variable is not valid. Therefore, we employ a technique which makes use of a χ^2 distribution as outlined in chapter 2 to verify that the distributions of the random variables we are comparing are normal. This process entails a specification of a significance level, α . This significance level means that if the distribution were truly normal, $100(1-\alpha)\%$ of sample distributions gathered would pass the test of normality. Therefore, as α decreases, the leniency of the test increases. This process is coded in `pairit.m` and for this experiment, α is set to be 0.01. If a random variable fails the test of normality, its samples are redefined as infinite, so that when the statistical test of hypothesis is performed, the significance level will be undefined.

The assumption of equal variances must also be verified. Assuming that the two random variables of comparison have passed the test of normality, the ratio of their sample variances must not exceed a certain number. This number is based on an F -distribution as outlined in chapter 2. We once again must specify a significance level of the test. We state that if we measure sample variances of two distributions with the same variance, there is a probability of $(1-\alpha)$ that the ratio of the two sample variances will be less than the number generated by the F -test. This process is coded

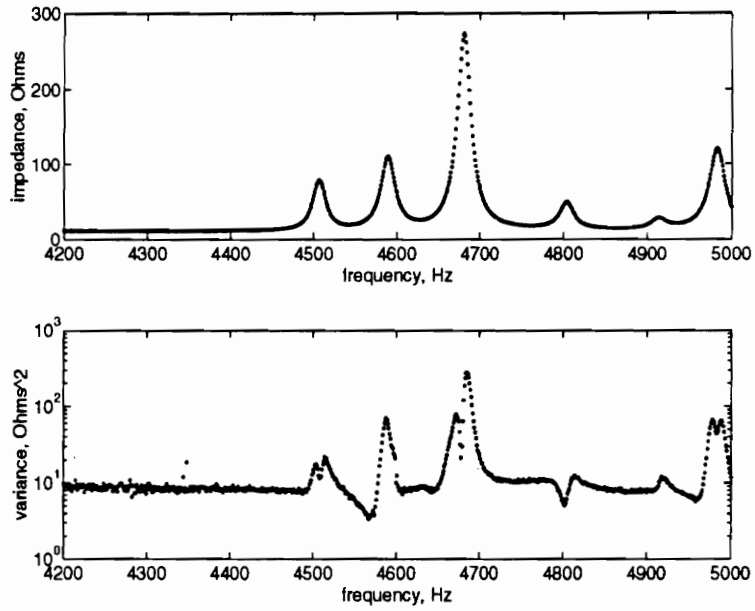


Figure 3.4: Sample mean and variance of real part of the 'no damage' case

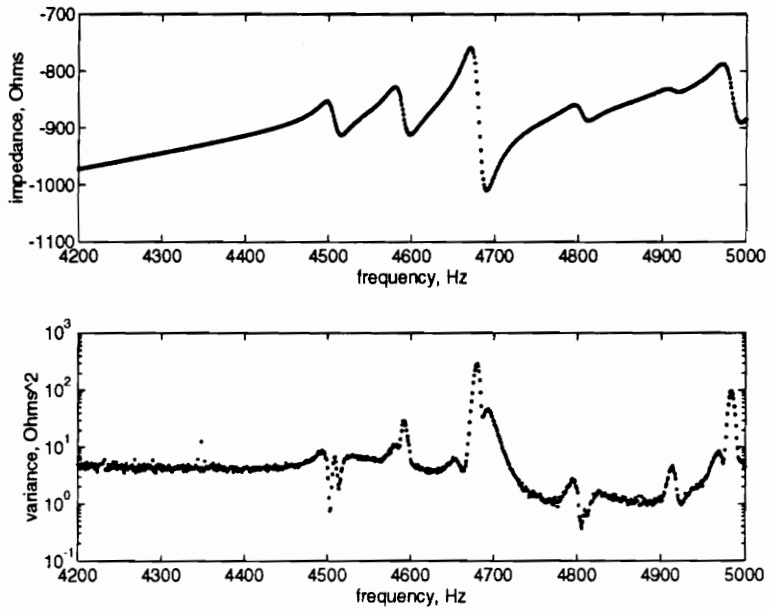


Figure 3.5: Sample mean and variance of imaginary part of the 'no damage' case

in `doit.m`. Once the two random variables of comparison have passed the normality test and the test of equal variance, we are prepared to perform the statistical test of hypothesis.

Assuming that two random variables of comparison have passed the test of normality and the test of equal assumption, a statistical test of hypothesis is performed, yielding a significance level, α . With this significance level, we state that given the sample means and sample variances of the two random variables of comparison, we are $100(1 - \alpha)\%$ confident that the true means of the two random variables are not equal. Therefore, this significance level, α is a metric as to how well we can discriminate between two random variables. The smaller the significance level, the better we can discriminate between random variables. We therefore make the 32,040 comparisons we defined above. The results of these comparisons are what we call *predictor variables*.

We sort the predictor variables on the basis of data type (real/imaginary) and sensor (green/red). For example, when comparing damage states A and B, we have four predictor variables which are $[801 \times 1]$ vectors. The values in each of these predictor variables come from the statistical test of hypothesis performed on ‘like’ random variables (i.e. the same frequency, data type, and sensor).

These predictor variables are generated using `doit.m`. This code steps through possible combinations of damage states, performing a statistical test of hypothesis for all ‘like’ random variables. If one of the random variables of comparison has failed the test of normality, the result of the hypothesis test is undefined. Likewise, if the random variables of comparison fail the test of equal variance, the predictor variable value is undefined. In performing the statistical test of hypothesis, `doit.m` calls the function `hypstest.m` which inputs sample means, sample variances, and number of samples, and outputs a significance level, α , as outlined in chapter 2.

The predictor variables for the comparison of damage states A and B are graphically depicted in Fig. 3.6 and Fig. 3.7. The comparison of the imaginary parts of the impedance-response using the red sensor is shown in Fig. 3.6. The comparison of the real parts of the impedance-response using the red sensor is shown in Fig. 3.7. In both of these figures, one can see differences in the sample means of

the responses as well as differences in the sample variances. It is a combination of these two parameters upon which the significance level is calculated. The minima of these predictor variables are the ‘best places to look’, that is the statistically most distinguishable points of the given frequency spectra for a given data type and sensor.

By selecting the minima of the predictor variables for a specified data type and sensor, we know ‘where’ to look for damage. For a given data type, sensor, and pair of damage states, we select the N best frequency points using `sigpv.m`. The next step is to determine ‘how’ to classify an impedance-response to be indicative of damage. This is accomplished by generating confidence bounds on random samples of the random variables corresponding to the minima of the specified predictor variable. In generating the confidence bounds of random samples, as outlined in chapter 2, we must specify a percentage of samples, p , and a significance level, α , with which we say we are $100(1 - \alpha)\%$ confident that $p\%$ of random samples will be contained in the interval $\bar{x} \pm ks$. In this example, we set p to be .99 and α to be .01. With these two parameters specified, and considering that there were ten examples gathered, k is determined to be 5.594. The routine `boundit.m` takes the significant predictor variables selected by `sigpv.m` and generates the confidence bounds for random samples of the random variables. The actual confidence bound generation is contained in `toler.m`.

The crux of this method is that when the confidence bounds are generated for random variables corresponding to the minima of the predictor variables for a pair of damage states, the confidence intervals for each damage state are the most mutually exclusive. In order to show this, we generate confidence bounds for a range of frequencies for comparison of damage states A and B using the red sensor. Fig. 3.8 shows the bounds for a 99% confidence interval for 99% of random samples of the real part of the impedance-response using the red sensor, for both damage states A and B. The same is shown for the imaginary part of the impedance-response in Fig. 3.9. Several important aspects of these figures should be mentioned. When there is a discontinuity in the confidence bounds, it means that the random variable failed the test of normality. When there are confidence bounds for ‘like’ random variables, but no significance level, this means that the pair of random variables failed the test of equal variance. For the rest of the compared random variables, we see that the more

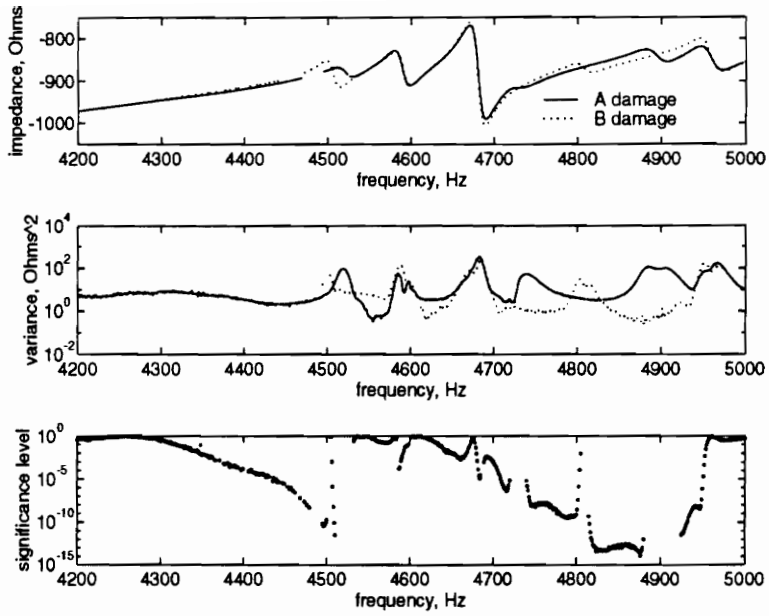


Figure 3.6: Predictor variable of damage cases A and B using imaginary part and red sensor.

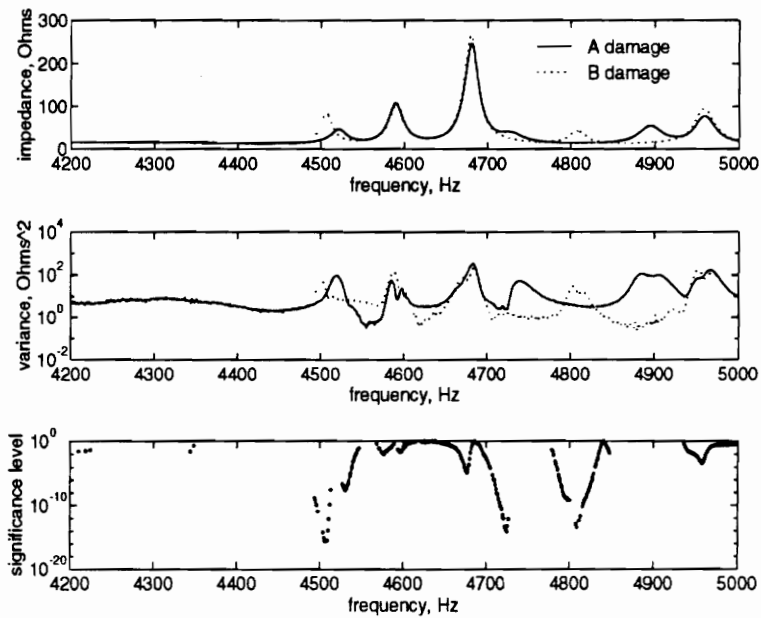


Figure 3.7: Predictor variable of damage cases A and B using real part and red sensor.

exclusive the confidence bounds, the lower the significance level. This is precisely the result we expected.

Using `sigpv.m`, we select the four best frequency points for each pair of damage states, sensor, and data type (i.e. the four lowest significance levels). We then generate confidence bounds on each of the corresponding random variables. These confidence bounds define 2 (sensors) \times 2 (data types) \times $4 = 16$ ‘rules’ by which we compare new data in order to identify the damage state. For example, when discriminating between damage states A and B, we can use any of the rules shown in Fig. 3.10. The circles represent the confidence bounds for damage state A and the x’s represent the confidence bounds for damage case B. We would measure the impedance-response of a structure, extract the corresponding frequency points and see where these data points fell relative to the confidence bounds. The generation of the rules is the vital aspect of this work. However, the rules can be applied in a number of ways, depending on desired computation time and sensitivity to false alarms. Two possible scenarios of rule applications are discussed in the next chapter.

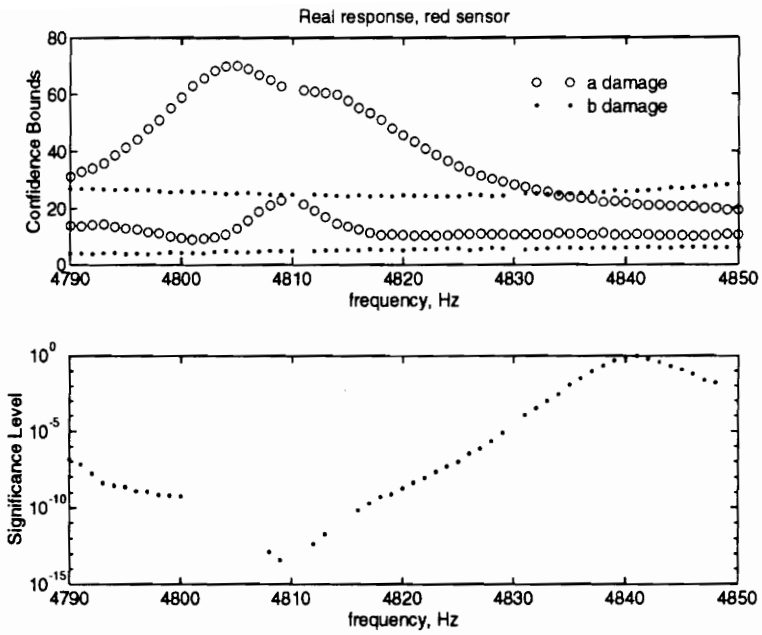


Figure 3.8: Confidence intervals of real part of impedance-response.

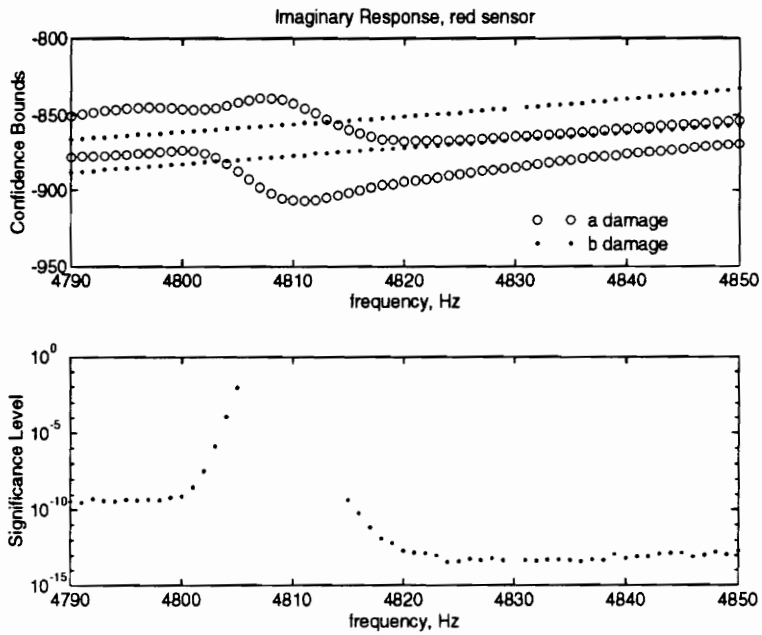


Figure 3.9: Confidence intervals of imaginary part of impedance-response.

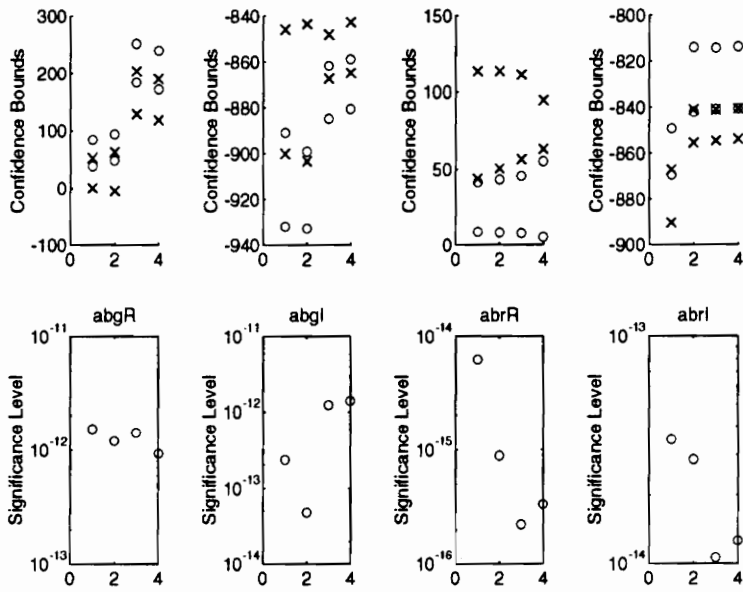


Figure 3.10: Discrimination between damage states A and B.

Chapter 4

Results

“I pass with relief from the tossing sea of Cause and Theory to the firm ground of Result and Fact.”

–Sir Winston Churchill

4.1 Rule Application Considerations

Chapters 2 and 3 discussed how rules for discriminating between damage states are generated. The implementation of these rules is not nearly as cut and dried. The rules can be successfully implemented in a number of ways. When a specific rule application algorithm is configured, certain attributes must be considered, such as running time of the algorithm, sensitivity to unknown damage states, and sensitivity to false alarms. Any of these aspects can be improved at the cost of the others. This chapter presents two proposed rule application algorithms, discussing the parameters which alter the algorithms attributes.

4.1.1 Method 1

In the previous chapter, we had five damage states and we generated $4 \text{ (rules)} \times 2 \text{ (data types)} \times 2 \text{ (sensors)} = 16$ ways of discriminating between any two damage states. The first method presented which is the most simple as far as programming is concerned, is to simply apply all of the rules to a set of test data. This may seem redundant, but the redundancy is crucial when dealing with unknown damage states.

If we were only going to encounter the damage states we trained for, there would be no need for redundant rule applications.

The process entails importing the test data and extracting the information corresponding to the rules being applied. If this data point falls within the bounds of a certain damage state for a rule, an index for that damage state is increased by one. There are 16 (different discriminators) \times 4 (other damage states) = 64 rules applied for each damage state. If we are testing a certain damage case, we would expect to see the index for that damage case very close to 64. We could test for each damage state and assume the highest index corresponds to the damage state which is present. However, if we were testing an unknown damage state, this would give us a false answer.

In order to give this algorithm sensitivity to unknown damage states, we define a threshold level. If the largest index is less than the threshold level, we determine that we are dealing with an unknown damage state. We have determined that there is a probability of .99 that 99% of random samples will be in the given confidence interval. Therefore, if we were testing a certain damage state, we would expect a value of $.99(.99)(64) = 62.7$ for the index level. We would typically set the threshold a little lower to account for variances we may not have considered in the original rule generation. The lower the threshold is set, however, the more susceptible to false alarms we are. For this example, we set the threshold at 60. If the largest index of a test is lower than the threshold, we determine that it is an unknown damage state. When testing against a null hypothesis, we are susceptible to what statisticians call *Type I* and *Type II* errors. A Type I error is when we reject the null hypothesis when it is actually true. A Type II error is when we accept the null hypothesis when it is actually false. Increasing the threshold defined here reduces the probability of Type I errors and increases the probability of Type II errors.

This method was applied to a set of test data. It was a blind test, in that the researcher applying the rules had no knowledge of the actual damage state of the test data. Test data was gathered from seven different damage states and the algorithm was applied using `rule.m`. The results of these tests are shown in Table 4.1.

The final two damage states tested consisted of a mass being placed on the

Table 4.1: Results of Method 1.

	Damage Indices					
Damage State	A	B	C	D	No damage	Damage Prediction
C	6	10	63	42	30	C
A	61	39	19	0	38	A
D	16	10	33	64	33	D
D	9	12	29	64	34	D
No damage	36	30	17	5	64	No damage
3% mass at center	38	41	13	5	39	Damage Unknown
1% mass at center	17	35	30	9	52	Damage Unknown

center of the plate (3% and a 1% of the total mass of the plate respectively). We can see from these results, that the algorithm was successful in discriminating between damage states to which we had prior knowledge as well as identifying the existence of damage to which we had no prior knowledge. There are certain aspects of these results that should be noted. Of the five tests of expected damage states, the average of the highest damage indices is 63.2, which is very close to the 62.7 we calculated for an expected value. We determine that the threshold of 60 we selected is sufficient for the cases tested.

The gathering of the actual data takes approximately two minutes. The application of the rules only takes about six or seven seconds. The time for data acquisition could be greatly reduced by developing a method to extract the necessary information for the rule application without gathering insignificant information. This is a topic of future research. Although the actual rule application takes only a few seconds, as damage states and number of sensors are increased, this elapsed time greatly increases. The next method to be presented addresses this problem.

4.1.2 Method 2

Instead of testing for all damage states simultaneously, this method tests for a single damage state before it tests for another. Initially, the method assumes a damage state. It tests this damage against all other damage states, one by one. If at any

point the test data does not fall in the bounds of the assumed damage state, the algorithm assumes another damage state and repeats the process until there are no damage states left, in which case the algorithm determines that the damage state is of unknown origin.

In testing for a particular damage state, we apply the rules as we did in the first method. Instead of applying all 16 discriminators, we select a certain number of these sixteen at random. This once again incorporates redundancy into the algorithm so that we can detect unknown damage states. The redundancy can be adjusted by selecting a number of the sixteen discriminators. The larger the number, the more confident we are with the answer, but this is at the cost of running time of the program. Using the identical data from the first method, we apply the rules using this method via `newtest.m` and `ruleit.m`. We selected values of 1, 3 and 5 for the number of rules applied for each discriminating test. The results are presented in Table 4.2.

Table 4.2: Results of Method 2.

	Damage Prediction		
	Number of Tests		
Damage State	1	2	3
C	C	C	C
A	A	A	A
D	D	D	D
D	D	D	D
No damage	No damage	No damage	No damage
3% mass at center	Unknown	Unknown	Unknown
1% mass at center	No damage	Unknown	Unknown

We can see from this table, that this method was also able to successfully discriminate between damage states to which we had prior knowledge as well as detect the existence of states to which we had no prior knowledge. It is also important to note that when the method was applied to data from a 1% mass placed at the center, we get a false answer when the number of tests is 1. When the number of tests is

increased, we get the correct prediction. This result coincides with what we expect; as the number of tests increase, we increase sensitivity to unknown damage states.

Chapter 5

Conclusions

Finis coronat opus.

–Anonymous

5.1 Significance

This thesis has presented a novel approach to the damage identification problem. This approach makes use of statistically-based inductive learning methods. When dynamic-response data are acquired, we must account for any variances inherent to the data acquisition process. We have shown that different parts of the dynamic-response are affected by these inherent variances more than others. The approach presented here takes this fact into account, selecting the parts of the dynamic-response with which we can statistically isolate the changes due to damage from the inherent variances in the data acquisition process.

Many researchers have hypothesized that the best way to identify damage via changes in a dynamic-response was to look for changes in natural frequencies and mode shapes. We have shown that this is not necessarily the case. By generating predictor variables, whose values are generated via a statistical test of hypothesis, the ‘best places to look’ may vary depending upon how the dynamic-response is affected by the inherent variances.

In order to demonstrate this approach, we performed an experimental test in which damage states consisting of the addition of a small test mass in various locations were discriminated from one another. This was accomplished with great

success. We were also able to detect the existence of damage states of unknown origin. The inherent variances in this test were assumed to be slight variations in boundary conditions and sensor noise.

Once the initial computations are performed, the running time of the algorithms is very small. This allows for the method to be incorporated in an on-line setting as opposed to several pattern recognition techniques which require a lot of computing time. As complexity increases and more information is gathered, these pattern recognition techniques have been shown to have trouble converging to acceptable answers. By contrast, the inductive learning approach thrives as more information is gathered.

Ideally, this method will be applied to a structure on which there are a finite number of damage states which are available for dynamic testing. Generally, this will seldom be the case. Even with knowledge of the healthy structure only, we would be able to generate confidence bounds on the dynamic-response based upon information gathered to measure the inherent variances in the problem. As more examples are gathered and as more damage states are available for testing, the more versatile and effective the method becomes.

Independent of the damage identification method itself, this research has raised issues which must be considered by anyone attempting to identify damage via changes in a dynamic-response. If the concepts of inherent variances are considered and incorporated into already existing damage identification methods, their effectiveness will be greatly improved.

5.2 Future Work

This research has laid the groundwork and opened many doors for future work. The most obvious extension of this work would be to take into account other variances inherent to the problem of damage identification. For this work, we had only assumed sensor noise and variations in boundary conditions to be the dominant inherent variances in our problem. Future work will entail the consideration of environmental effects on the dynamic-response such as temperature, humidity and atmospheric con-

ditions. Manufacturing variability is something that needs to be addressed as well.

Extensions of this method could also be made with other types of sensors and dynamic-response data. Since this method is not dependent on any specific type of sensor or data type, it is applicable to virtually any structure. These limitless possibilities should be investigated.

Finally, a more comprehensive investigation into rule application algorithms should be performed. This research has merely touched on the the issues of running time, sensitivity to false alarms, and sensitivity to unknown damage states. If one could develop a rule application algorithm in which these issues were quantitative as opposed to qualitative, the entire damage identification using inductive learning package becomes more marketable and attractive.

References

- Bennett, K.D., R.O. Claus and M.D. Dumais. (1988). "Impact Characterization by Optical Fiber Ultrasonic Sensors". *Ultrasonics Symposium Proceedings*. Vol. 2, pp. 1049-1052.
- Campanile, L.F. (1993). "A Modal Method for Structural Health Monitoring". *Proceedings of the 4th International Conference on Adaptive Structures*, November 1993, Cologne, Germany, pp. 205-219.
- Cawley, P. and R.D. Adams. (1979). "The Location of Defects in Structures from Measurements of Natural Frequencies". *Journal of Strain Analysis*. Vol 14, No. 2, 1979, pp. 49-57.
- Chen, Hsin-Piao, J.J. Tracy, and R. Nonato. (1993). "Natural Frequencies of Delaminated Composite Laminates Under Compressive Loads". AIAA-93-1480-CP, pp. 1493-1505.
- Collins, K.R., R.H. Plaut, and J. Wauer. (1992). "Free and Forced Longitudinal Vibrations of a Cantilevered Bar with a Crack". *Journal of Vibration and Acoustics*. April 1992, Vol. 114, pp. 171-177.
- DiPasquale, Edmondo and Ahmet S. Cakmak. (1989). "Detection of Seismic Structural Damage". *Seismic Engineering: Research and Practice Proceedings of the Sessions Related to Seismic Engineering at Structures Congress '89*. San Francisco, California, May 1989, pp. 468-477.
- Doebbling, S.W. et al. (1993). "Selection of Experimental Modal Data Sets for Damage Detection Via Model Update". AIAA-93-1481-CP, pp.1506-1517.

- Elkordy, M.F., K.C. Chang, and G.C. Lee. (1993). "Neural Networks Trained by Analytically Simulated Damage States". *Journal of Computing in Civil Engineering*, Vol. 7, no.2, pp. 130-145.
- Gentry, Jeffrey D. et al. (1993). "Structural Health Assessment and Review Program (SHARP)-Prototype of an On-board Structural Health Monitoring System". *SPIE Vol. 1917 Smart Structures and Intelligent Systems*. 1993, pp.1020-1027.
- Glass, B.J. and S. Hanagud.(1990). "An AI-Based Approach to Structural Damage Identification by Modal Analysis". AIAA-90-1170-CP, pp. 2390-2397.
- Glass, B.J. and A. Macalou. (1991). "Search-based Model Identification of Smart Structure Damage". *AD-Vol.24/AMD-Vol. 123, Smart Structures and Materials*, pp. 33-38.
- Glossop, N.D.W. et al. (1990). "Optical Fibre Damage Detection for an Aircraft Composite Leading Edge". *Composites*. Vol. 21, No. 1, January 1990, pp. 71-80.
- Gomes, A.J.M. Araujo and J.M. Montalvao E. Silva. (1990). "On the Use of Modal Analysis for Crack Identification". *Proceedings of 8th International Modal Analysis Conference*. Kissimmee, Florida, 1990. pp. 1108-1115.
- Hickman, G.A., J.J. Gerardi, and Y. Feng. (1990). "Application of Smart Structures to Aircraft Health Monitoring". *International U.S./Japan Conference on Adaptive Structures*. Maui, Hawaii, November 1990, pp. 966-984.
- Hofer, B. (1987). "Fibre Optic Damage Detection in Composite Structures". *Composites*. Vol 18, No. 4, September 1987, pp. 309-316.
- Hogg, Robert V. and Johannes Ledolter. (1992). Applied Statistics for Engineers and Physical Scientists, Macmillan Publishing Company, New York, New York.
- Kaouk M. and D.C. Zimmerman. (1993). "Structural Damage Assessment Using a Generalized Minimum Rank Perturbation Theory". AIAA-93-1483-CP, pp. 1529-1538.

- Kiel, David H. (1993). "Active Damage Control Using Artificial Intelligence: Initial Studies into Identification and Mitigation". Master's Thesis, VPI&SU.
- Kudva, J.N., C. Marantidnis, and J. Gentry. (1993). "Smart Structures Concepts for Aircraft Structural Health Monitoring". *SPIE Vol.1917 Smart Structures and Intelligent Systems*. 1993, pp. 964-971.
- Kudva, J.N., N.Munir, and P.W. Tan. (1992). "Damage Detection in Smart Structures Using Neural Networks and Finite-element Analyses". *Smart Mater. Struct.*, 1 (1992), pp. 108-112.
- Kung, D.N. et al. (1989). "A Time Domain System Identification Technique Based on Impulsive Loading for Damage Detection". *Proceedings of teh 8th International Conference on Offshore Mechanics and Arctic Engineering*. The Hague, Netherlands, March 1989, pp. 307-313.
- Leech, W.J. (1986). "A Rule Based Process Control Method with Feedback". *Proceedings of the ISA/86 International Conference and Exhibit*. October 1986, Houston, Texas, pp.168-175.
- Lindner, Douglas K. and George Kirby. (1994). "Location and Estimation of Damage in a Beam Using Identification Algorithms". AIAA-94-1755-CP, pp. 192-198.
- Lindner, Douglas K., Gregory Twitty, and Richard Goff. (1993). "Damage Detection, Location, and Estimation for Large Truss Structures". AIAA-93-1484-CP, pp.1539-1548.
- Manning, R.A. (1994). "Damage Detection in Adaptive Structures Using Neural Networks". AIAA-94-1752-CP, pp. 160-172.
- Marantidis, Constantinos, J.D. Gentry, and J.N. Kudva. (1993). "Sensor and Sensing Technologies for Structural Health Monitoring of Aircraft". *SPIE Vol. 1917 Smart Structures and Intelligent Systems*. 1993, pp. 972-979.

- Masri, S.F., A.G. Chassiakos and T.K. Caughey. (1992). "Structure-unknown Non-linear Dynamic Systems: Identification Through Neural Networks". *Smart Mater. Struct.*1, 1992, pp. 45-56.
- Maus, Rex, and Jessica Keyes. (1990). Handbook of Expert Systems in Manufacturing, McGraw-Hill, Inc., New York, New York.
- Napolitano, Marcello R., C.I. Chen, and S. Naylor.(1993). "Aircraft Failure Detection and Identification Using Neural Networks". *Journal of Guidance, Control, and Dynamics*, Vol. 16, No. 6, November-December 1993, pp. 999-1009.
- Napolitano, M.R., Ching I. Chen, and R. Nutter. (1992). "Application of a Neural Observer as State Estimator in Active Vibration Control of a Cantilevered Beam". *Smart Mater. Struct.*1, 1992, pp. 69-75.
- Peterson, L.D. et al. (1993). "Damage Detection Using Experimentally Measured Mass and Stiffness Matrices". AIAA-93-1482-CP, pp. 1518-1528.
- Povich, Clinton R. and Tae W. Lim. (1994). "An Artificial Neural Network Approach to Structural Damage Detection Using Frequency Response Functions". AIAA-94-1751-CP, pp. 151-159.
- Rhim, Jaewook and Sung W. Lee. (1994). "A Neural Network Approach for Damage Detection and Identification of Structures". AIAA-94-1753-CP, pp. 173-180.
- Richardson, Mark H. and M.A.Mannan. (1992). "Remote Detection and Location of Structural Faults Using Modal Parameters". *Proceedings of the 10th International Modal Analysis Conference*. San Diego, California, 1992. pp. 502-507.
- Roitman, N. et al. "Identification of Offshore Platform Structural Damage Using Modal Analysis Techniques". *Mechanical Systems and Signal Processing*, (1992), 6(3), pp. 287-295.
- Saravanos, Dimitris A., Victor B. Birman, and Dale A. Hopkins. (1994). "Detection of Delaminations in Composite Beams Using Piezoelectric Sensors". AIAA-94-1754-CP, pp. 181-191.

- Sensburg, O. and G.R. Tomlinson. (1993). "Damage Detection of Aircraft Structures Using Dynamic Analysis and Testing Methods". *Proceedings of the 4th International Conference on Adaptive Structures*, November 1993, Cologne, Germany, pp. 205-219.
- Simon, Solomon Henry. (1993). "An Integrated Smart Structures Wingbox". *SPIE Vol.1917 Smart Structures and Intelligent Systems*. pp. 991-996.
- . (1993). "A Neural Network/Expert System Hybrid which Identifies Untrained Damage". *SPIE Vol. 1918 Smart Sensing, Processing, and Instrumentation*. 1993, pp. 131-137.
- Tappert, Peter M., and H.H. Robertshaw.(1994). "Manipulations of Dynamic-response Data for Use in Inductive Learning Methods to Identify Damage". *Proceedings of the 1994 North American Conference on Smart Structures and Materials*. Orlando, Florida, February 1994.
- Tavares,R. et al. (1993). "Using Experimental Modal Data to Detect Damage in a Space Truss". AIAA-93-1486-CP, pp. 1556-1564.
- Wolff, Tom and Mark Richardson. (1989). "Fault Detection in Structures from Changes in their Modal Parameters". *Proceedings of the 7th International Modal Analysis Conference*. Las Vegas, Nevada. 1989. pp. 87-94.
- Yen, Gary G. (1993). "Neural Network Approach for the Damage Detection of Structures". AIAA-93-1485-CP, pp. 1549-1555.

Appendix A

loadit.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%   This routine imports the experimental data generated from a
%   Hewlett-Packard Impedance Analyzer.  The data is of the form:
%
%
%           ng42_1.dat
%           || | \
%           / | \   example number(1-10)
%   damage   |   frequency range(4200-4600Hz or
%   state(a-d or n) |                               4600-5000Hz)
%
%           sensor(g or r)
%
%   The data is imported into MATLAB and saved as a .mat file
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% These 'for' blocks scroll through the damage state, sensor,
%   frequency range and example number

for k = 'abcd'
for j = 'gr'
for m = [42 46]
for i = 1:10
eval(['load b:\',k,j,num2str(m), '_', num2str(i), '.dat'])
eval(['k,j,num2str(m), '_', num2str(i), '=', k,j,num2str(m), '_',
num2str(i), '(:,2:3);']);
eval(['save c:\pmt\thesis\test2\',k,j,num2str(m), '_',
num2str(i)]);
eval(['clear ',k,j,num2str(m), '_', num2str(i)]);
end
end
end
end
```

pairit.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This routine imports the .mat files generated by loadit.m and
% combines the data into 20 [10 X 801] matrices based on
% combinations of damage state, sensor, and data type
% (real/imaginary). The rows in these matrices represent
% examples and the columns indicate frequency.
%
% The next part of the routine tests each distribution for
% normality.
%
% The next part of the routine calculates sample means and
% variances of the matrices.
%
% The data is saved as mean_var.mat which contains 20 [801 X 1]
% sample mean vectors and 20 [801 X 1] sample variance vectors
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% These 'for' loops import the data
for k = 'nabcd'
    for j = 'gr'
        for m = [42 46]
            for i = 1:10;
                eval(['load c:\pmt\thesis\test2\' ,k,j,num2str(m),'_',
                    num2str(i)])
            end
            eval([k,j,num2str(m),'_2 = ',k,j,num2str(m),'_11;'])
            eval(['clear ',k,j,num2str(m),'_11;'])
        end
    end
end

% Calculation of comparative 'Q' value for the normality test.
% With this, a significance level is set as well as a number of
% degrees of freedom, which is based on number of cells, K,
% defined below.
siglev = .99;
dof = 2;
Qcomp = chi2inv(siglev,dof);

% This combines the two frequency spans and example into matrix.
for k = 'nabcd'
    for j = 'gr'
        for i = 1:10

            eval([k,j,'R(',num2str(i),',,:) = [',k,j,num2str(42),'_',
                num2str(i),'(1:400,1);',k,j,num2str(46),'_',
                num2str(i),'(:,1)]'';'])

            eval([k,j,'I(',num2str(i),',,:) = [',k,j,num2str(42),'_',
```

```

        num2str(i), '(1:400,2);', k, j, num2str(46), '_ ',
        num2str(i), '(:,2)]'';']
    for m = [42 46]
        eval(['clear ', k, j, num2str(m), '_ ', num2str(i)]);
    end
end
for n = 'IR'
    for cc = 1:801
% Define x as the examples of one random sample
        eval(['x = ', k, j, n, '(:,cc);'])

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Test of Normality
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

spread = max(x) - min(x);

% Sample mean of random variable
m = mean(x);

% Sample variance of random variable
s2 = var(x);
s = sqrt(s2);

% Number of cells for test
K=5;

% The r vector indicates the lower bounds of cells
r(1) = -Inf;
for I = 2:K
r(I) = min(x) + (I-1)/K*spread;
end

% Calculation of the probabilities for each cell
for I = 2:K
p(I-1) = normcdf(r(I),m,s)-normcdf(r(I-1),m,s);
end
p(K) = 1-sum(p(1:4));

count = zeros(K-1,1);

% Counting of number of samples in each cell
for J = 2:K
for I = 1:length(x)
    if r(J-1) < x(I) & x(I) < r(J);
        count(J-1) = count(J-1)+1;
    end
end
end

count(K) = length(x) - sum(count);

```

```

% Calculation of number of Q residuals based on number of
% expected samples in each cell and actual number of samples.
for I = 1:K;
    Q(I) = (count(I)-length(x)*p(I))^2 / (length(x)*p(I));
end

% Comparison of Q's based on Qcomp calculated above. If the
% distribution is found not to be normal, the samples are
% redefined as infinite.
Qtot = sum(Q);
if Qtot > Qcomp
    eval([k,j,n,'(:,cc) = Inf*ones(10,1);'])
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% End Normality test
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

    end

% Calculation of sample means and sample variances of matrices
    eval([k,j,n,'m',' = mean(' ,k,j,n,') ;'])
    eval([k,j,n,'s',' = var(' ,k,j,n,') ;'])
    eval(['clear ' ,k,j,n])
end
end
freq = linspace(4200,5000,801);
clear i j k m n
save c:\pmt\thesis\test2\mean_var

```

doit.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%  
% This routine creates predictor variables. The function  
% hypptest.m is called to perform a statistical test of hypothesis  
% on the basis of sample means and sample variances.  
%  
% In this routine, the assumption of equal variances is verified  
%  
% The data are saved as 40 [801 X 1] predictor variable vectors  
% which are the significance levels of the test. If a comparison  
% failed one of the assumption tests, the corresponding value  
% would be NaN.  
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
clear
```

```
% Load sample mean and sample variance data  
load mean_var
```

```
% Define significance level for test of equal variance  
alpha = .01;
```

```
% Number of degrees of freedom are based on number of samples -1  
dofx = 9;  
dofy = 9;
```

```
% The ratio of sample variances cannot be greater than 'test'  
test = finv(1-alpha,dofy,dofx);
```

```
% This loop makes all of the comparisons needed, calling hypptest.m  
% to calculated significance level if it passes equal variance test  
dam = 'abcdn'
```

```
for k = 1:length(dam)-1  
    for i = k+1:length(dam);  
        for n = 'RI';  
            for m = 'rg';  
                for j = 1:length(nrIm);  
                    eval(['if',dam(k),m,n,'s(j)'/',dam(i),m,n,'s(j)>test|',  
dam(i),m,n,'s(j)'/',dam(k),m,n,'s(j)>test, pv',dam(k),  
dam(i),m,n,'(j) = Inf; else pv',dam(k),dam(i),m,n,'(j) =  
hypptest(',dam(k),m,n,'m(j)',',dam(k),m,n,'s(j)',',  
dam(i),m,n,'m(j)',',dam(i),m,n,'s(j)',10,10);end;'])  
                    end;  
                end  
            end  
        end  
    end  
end  
clear k i n m j Num
```

```
save c:\pmt\thesis\test2\numreal
```

hypptest.m

```
function [sig] = hypptest(xnew,sx,ynew,sy,Nx,Ny)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% HYPTEST
%
% This function takes two random variables with sample means
% and sample standard deviations taken from N samples and performs a
% hypothesis test using the Student's t statistics. The function
% outputs two significance factors. These factors are based on two
% separate tests of the null hypothesis. That is that the means of
% the two samples are the same. The first method assumes that xnew
% is the true mean of one random variable and performs a hypothesis
% test to determine to what degree of certainty we can say that the
% second random variable has the same mean. The lower the value of
% Sig, the more confidence we have in saying that the means of the
% two random variables are not equal. The second method, which
% outputs sig, uses the t-statistic to determine to what degree of
% certainty we can say that the means of the two random variables
% have the same mean. Once again the lower the value of sig, the
% more confidence we have in saying that the means of the two
% samples are not equal. The converse is also true, that the higher
% the value of sig, the more confidence we have in saying that the
% means of the two random variables are the same. A lot of this
% function was taken from TTEST.M and TTEST2.M in the STATS toolbox.
% They, however calculate from a given sample whereas this program
% calculates from a given sample mean and sample standard deviation.
% PMT 3/19/94
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% This is the beginning of the two-way hypothesis test
% Degrees of freedom
dfx = Nx-1;
dfy = Ny-1;

% Effective degrees of freedom
dfe = dfx + dfy;

for i = 1:length(xnew)
    msx(i) = dfx*(sx(i));
    msy(i) = dfy*(sy(i));
    Sp(i) = (msx(i)+msy(i))/dfe;
    diff(i) = xnew(i) - ynew(i);
    AA = 1/(dfx+1) + 1/(dfy+1);

% Pooled variance
    pooleds(i) = sqrt(Sp(i)*AA);
    ratio(i) = diff(i)/pooleds(i);
    sig(i) = tcdf(ratio(i),dfe);
    sig(i) = 2*min(sig(i),1-sig(i));
end
```


sigpv.m

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This routine selects the best values of predictor variables
% and saves them in a [N X 2] vector, where the first collumn
% is frequency and the second collumn is signifcance level.
%
% These significant predictor variable values are saved in
% the following form:
%
%           sigabrI
%           ||||
%           ||| \
% damage   // |  data type
% states   // |
% of comparison |
%           sensor
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear
load numreal

num = 4;

dam = 'abcdn'
for k = 1:length(dam)
    for j = k+1:length(dam)
        for m = 'rg'
            for n = 'RI'
eval(['[y,I] = sort(pv',dam(k),dam(j),m,n,'(1:length(pv',dam(k)
        ,dam(j),m,n,')'));']);
eval(['Sig',dam(k),dam(j),m,n,' = [I(1:num)''+ 4199*ones(num,1)
        ,y(1:num)''];']);

eval(['[y,I] = sort(Sig',dam(k),dam(j),m,n,'(:,1));'])
        for i = 1:num
eval(['sig',dam(k),dam(j),m,n,'(i,:) = [y(i), Sig',dam(k),dam(j)
        ,m,n,'(I(i),2)];']);
            end
        end
    end
end

for k = 1:length(dam)
    for j = 1+k:length(dam)
        for m = 'gr'
            for n = 'RI'
                eval(['clear Sig',dam(k),dam(j),m,n])
                eval(['clear pv',dam(k),dam(j),m,n])
            end
        end
    end
end

```

```
    end
  end
end
end
clear k j m n dam freq I i num y
```

toler.m

```
function [lcl,ucl] = toler(xbar,s2,N)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%   This function uses values of k from Table C.9 in Hogg
%   and Ledolter with alpha = .01 and p = .99.  With inputs
%   of sampled mean and sampled variance as well as number
%   of samples, this function outputs confidence bounds for
%   random samples of the random variables.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

    if N == 10
        k = 5.594;
    end

% Generation of confidence bounds
for i = 1:length(xbar)
    del = k*sqrt(s2(i));
    lcl(i) = xbar(i) - del;
    ucl(i) = xbar(i) + del;
end
```

boundit.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This routine takes the significant predictors
% generated by sigpv.m and sets confidence bounds
% on them calling the function toler.m. The data
% are saved as a five column matrix with the first
% column is frequency, the next two columns are the
% lower and upper bound of the first damage state
% and the last two columns are the lower and upper
% bound of the second damage state.
%
% damage states
%
%      \--
%      babgI--data type
%      /
% bound      sensor
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
dam = 'abcdn';

for k = 1:length(dam)
    for j = k+1:length(dam)
        for m = 'rg'
            for n = 'RI'
for i = 1:length(sigbcgR)
                eval(['h = sig',dam(k),dam(j),m,n,'(i,1) - 4199;']);
                eval(['[lcl1,ucl1] = toler(',dam(k),m,n,'m(h)',',',
                    dam(k),m,n,'s(h),10);'])
                eval(['[lcl2,ucl2] = toler(',dam(j),m,n,'m(h)',',',
                    dam(j),m,n,'s(h),10);'])
                eval(['b',dam(k),dam(j),m,n,'(i,:) = [sig',dam(k),
                    dam(j),m,n,'(i,1),lcl1,ucl1,lcl2,ucl2];'])
            end
        end
    end
end

save bounds
```

rule.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This is the first method of rule application.
% The test data are loaded in as testgI,testgR,
% testrI, and testrR. All of the rules generated
% by boundit are applied. This case is set up for
% 4 rules per predictor variable. A threshold is
% set to take into account unknown damage cases.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
load testdata
threshold = 60;

dam = 'abcdn';
a = 0;
b = 0;
c = 0;
d = 0;
n = 0;

for k = 1:length(dam)
    for j = k+1:length(dam)
        for m = 'gr'
            for nn = 'RI'
                for p = 1:4;

eval(['x = b',dam(k),dam(j),m,nn,'(',num2str(p),',1);'])

% Apply first set of bounds
eval(['if b',dam(k),dam(j),m,nn,'(',num2str(p),',2)<test',
    m,nn,'(',num2str(x-4199),',1) & b',dam(k),dam(j),m,
    nn,'(',num2str(p),',3)>test',m,nn,'(',num2str(x-4199)
    ',1);',dam(k),'= ',dam(k),'+1;end;'])

% Apply second set of bounds
eval(['if b',dam(k),dam(j),m,nn,'(',num2str(p),',4)<test',
    m,nn,'(',num2str(x-4199),',1) & b',dam(k),dam(j),m,
    nn,'(',num2str(p),',5)>test',m,nn,'(',num2str(x-4199)
    ',1);',dam(j),'= ',dam(j),'+1;end;'])

        end
    end
end
end
end

if a>b&a>c&a>d&a>n&a>threshold
    damage = 'a'
else if b>a&b>c&b>d&b>n&b>threshold
    damage = 'b'
else if c>a&c>b&c>d&c>n&c>threshold
```

```
    damage = 'c'  
else if d>a&d>c&d>n&d>threshold  
    damage = 'd'  
else if n>a&n>b&n>c&n>d&n>threshold  
    damage = 'no damage'  
else  
    damage = 'damage unknown'  
end  
damage
```

newtest.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This routine selects rules at random and extracts
% the necessary information from the test data
% for use in rulit.m.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Randomly select a test boundary (damage states selected)

j = fix(numpvs*rand)+1;      % selects predictor variable
k = fix(length(sen)*rand)+1; % selects a sensor
m = fix(length(type)*rand)+1; % selects a data type

% Select test bounds and data point to use in ruleit comparison

eval(['testbnds = b',sort([dam(i),dam(h)]),sen(k),
      type(m),'(',num2str(j),'',2:5);'])
eval(['freqpt = b',sort([dam(i),dam(h)]),sen(k),
      type(m),'(',num2str(j),'',1) - lowfreq + 1;'])
eval(['testpt = td',sen(k),'(',num2str(freqpt),'',
      num2str(freqpt),'',num2str(m),'');'])

if dam(h)<dam(i)
    testbnds = [testbnds(3:4) testbnds(1:2)];
end
```

ruleit.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This routine applies the rules for the second
% method discussed in the results section.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

damage = [];

% Performs 2 loops to take into account outliers
for v = 1:2

% If a damage state passes, break out of loop
if damage == dam(i)
    break
end
q = 0;
s = 0;

% Selection of number of tests to pass
numtests = 5;
for t = 1:p
    dam = [dam(2:p) dam(1)];
    for h = 2:p
        for r = 1:numtests
            newtest3;

% If there is an indecisive test, select a new one
            while (testpt > testbnds(1) & testpt < testbnds(2))
                & (testpt > testbnds(3) & testpt < testbnds(4))
                    newtest3;
            end

% If the test data passes first damage state, index q
            if (testpt > testbnds(1) & testpt < testbnds(2)) ~
                (testpt > testbnds(3) & testpt < testbnds(4));
                q = q+1;
            end
            end

% If it passes for the number of tests, index s
            if q == numtests
                s = s+1;
                q = 0;
            else
                s = 0;
                q = 0;
            end
        end
    end

end
if s == p-1
```



```
    damage = dam(i)
    break
else
    q = 0;
    s = 0;
end
end
end
end
if s ~= p-1
    damage = 'unknown'
end
```

Appendix B

This appendix contains all of the predictor variables generated by `doit.m`. There is one predictor variable for each combination of sensor (red or green) and data type (real/imaginary). Since there are 10 possible damage state combinations, there are 40 predictor variables contained here. The figures also show the sample means and sample variances of the random variables.

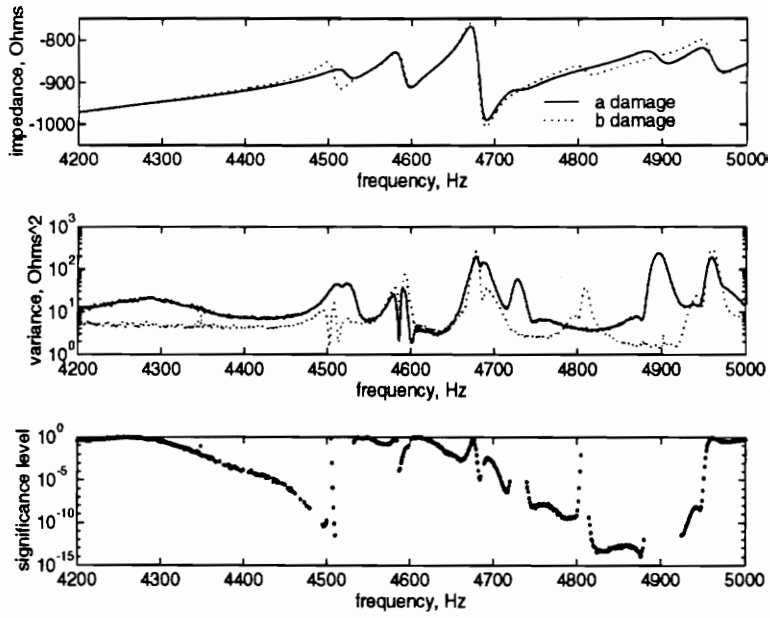


Figure 1: Imaginary impedance response, red sensor.

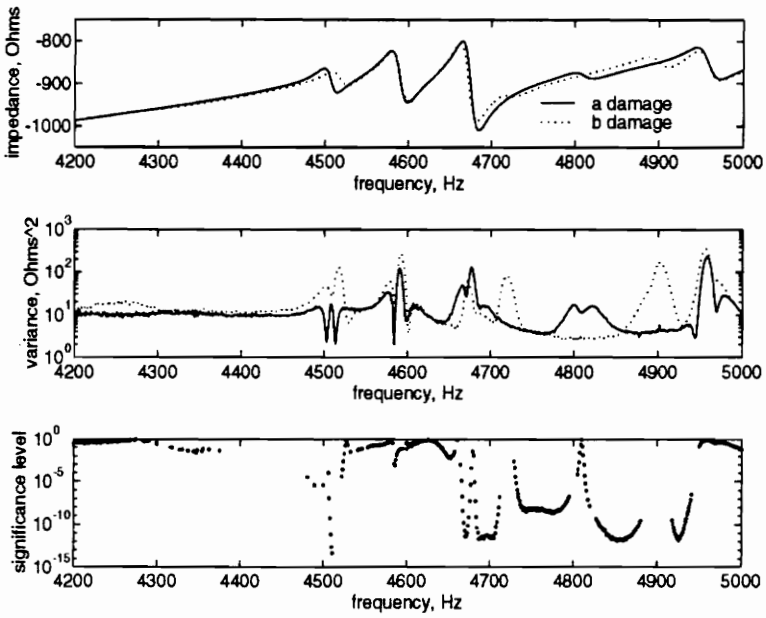


Figure 2: Imaginary impedance response, green sensor.

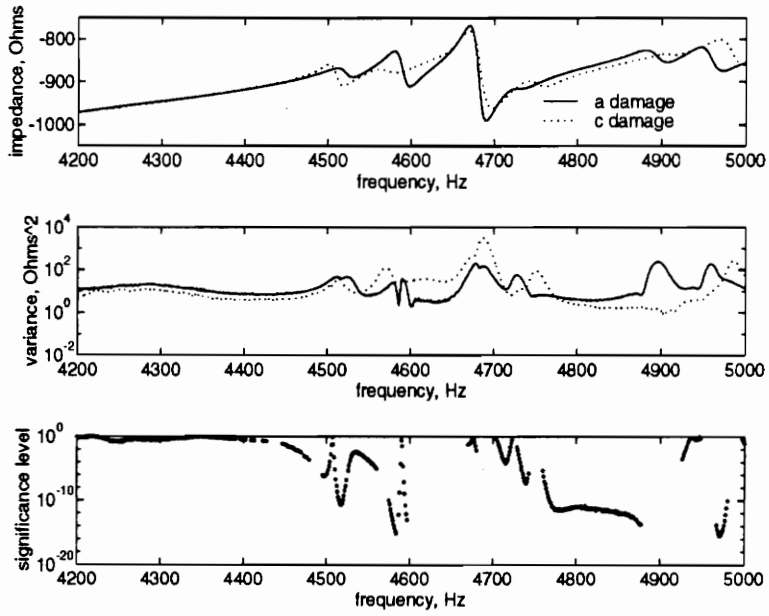


Figure 3: Imaginary impedance response, red sensor.

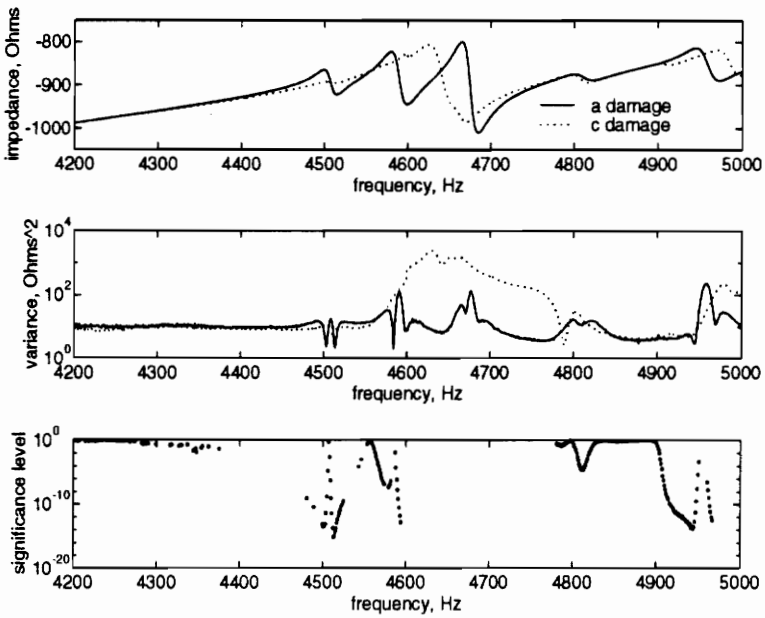


Figure 4: Imaginary impedance response, green sensor.

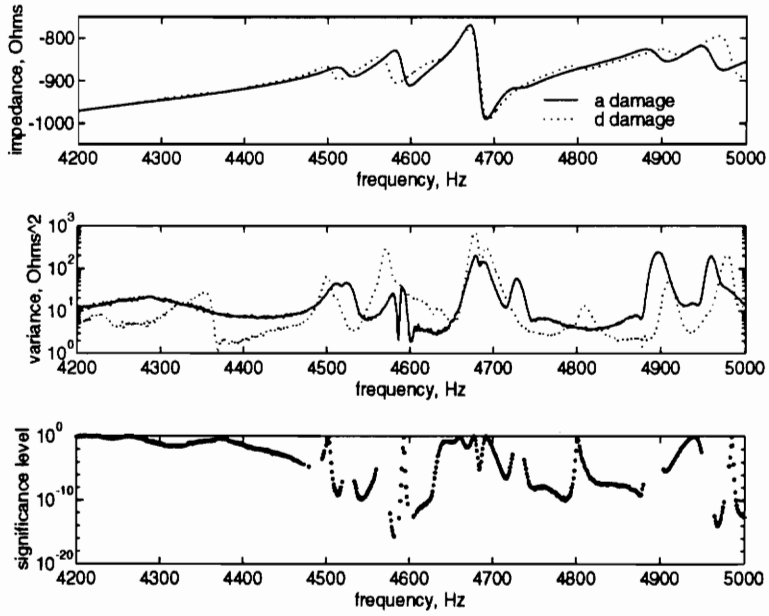


Figure 5: Imaginary impedance response, red sensor.

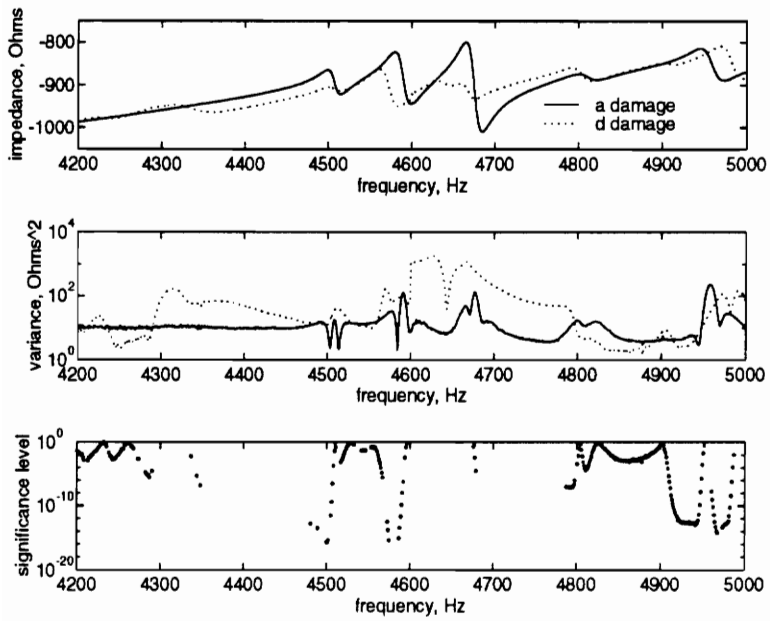


Figure 6: Imaginary impedance response, green sensor.

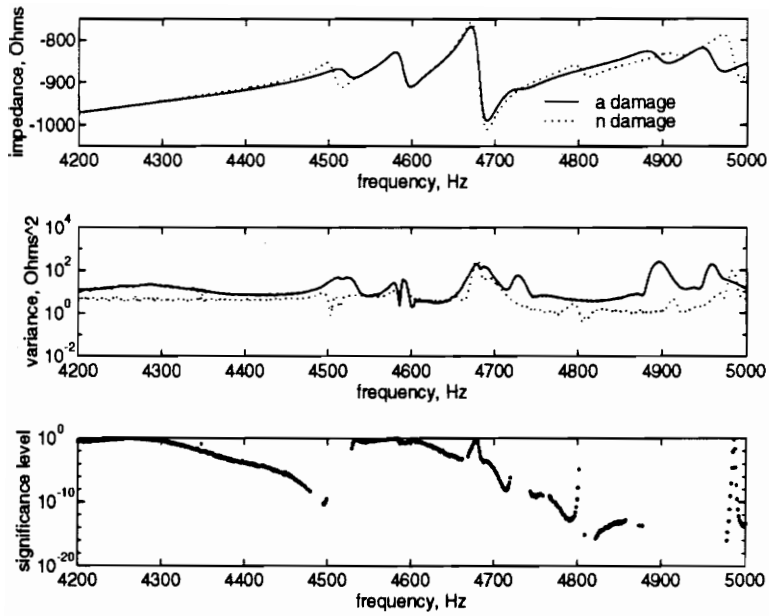


Figure 7: Imaginary impedance response, red sensor.

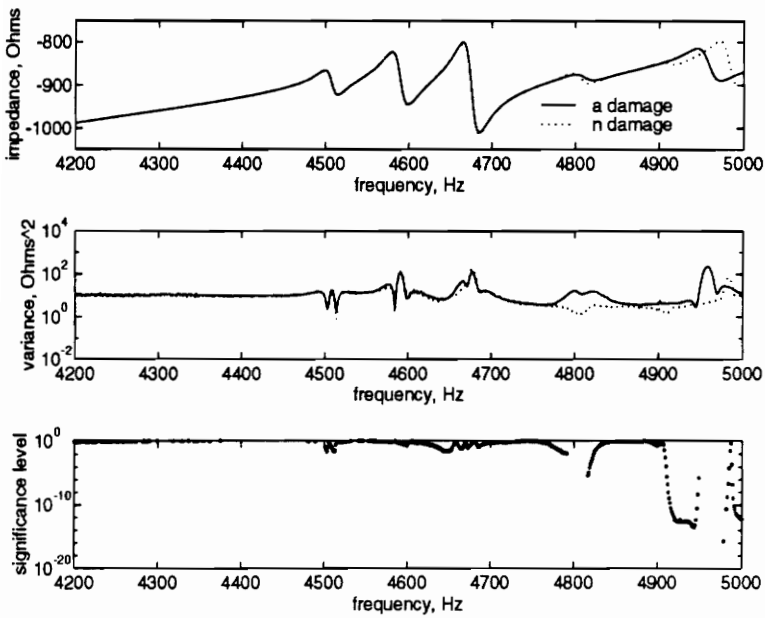


Figure 8: Imaginary impedance response, green sensor.

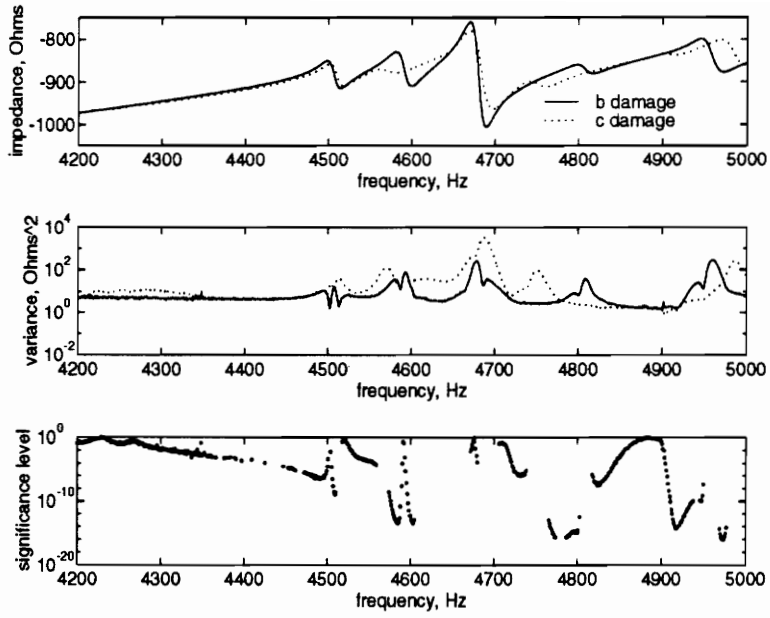


Figure 9: Imaginary impedance response, red sensor.

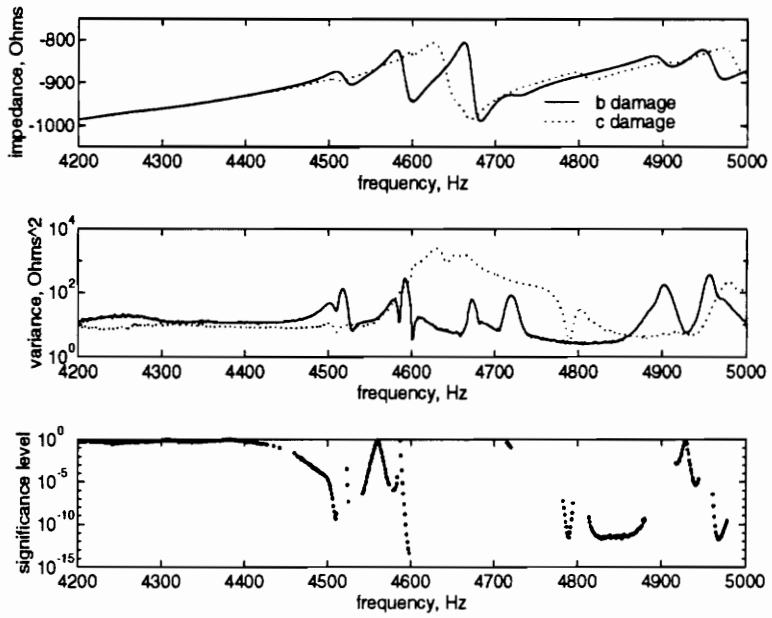


Figure 10: Imaginary impedance response, green sensor.

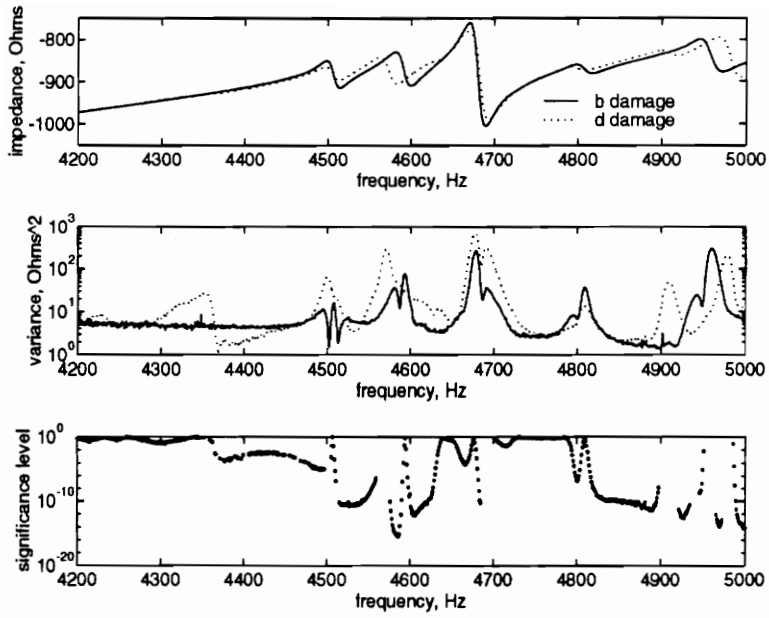


Figure 11: Imaginary impedance response, red sensor.

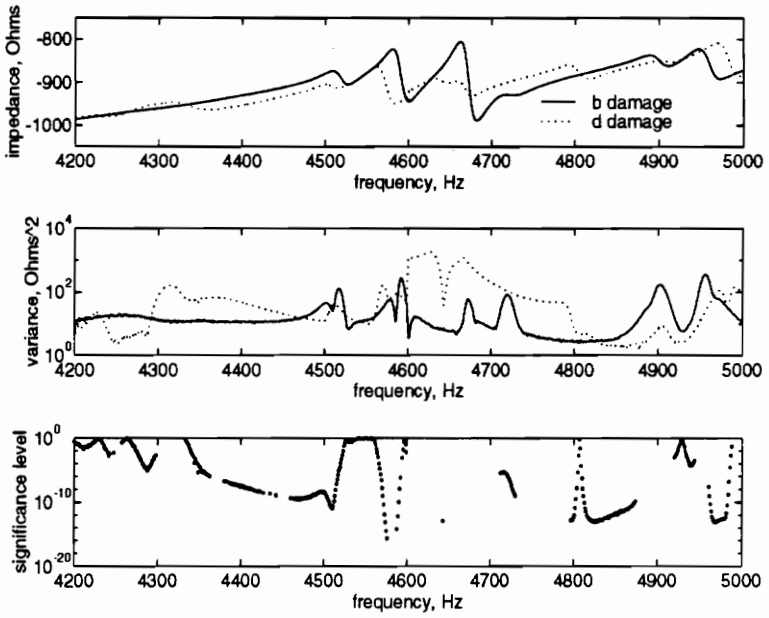


Figure 12: Imaginary impedance response, green sensor.

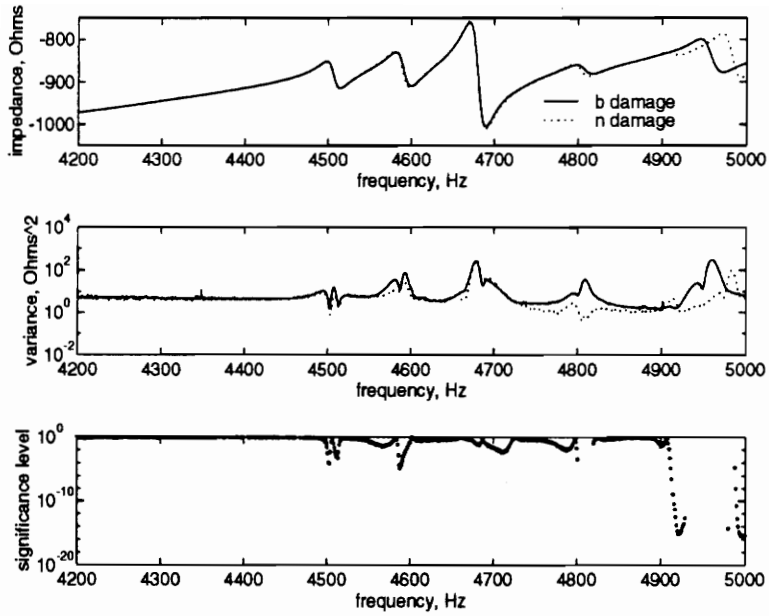


Figure 13: Imaginary impedance response, red sensor.

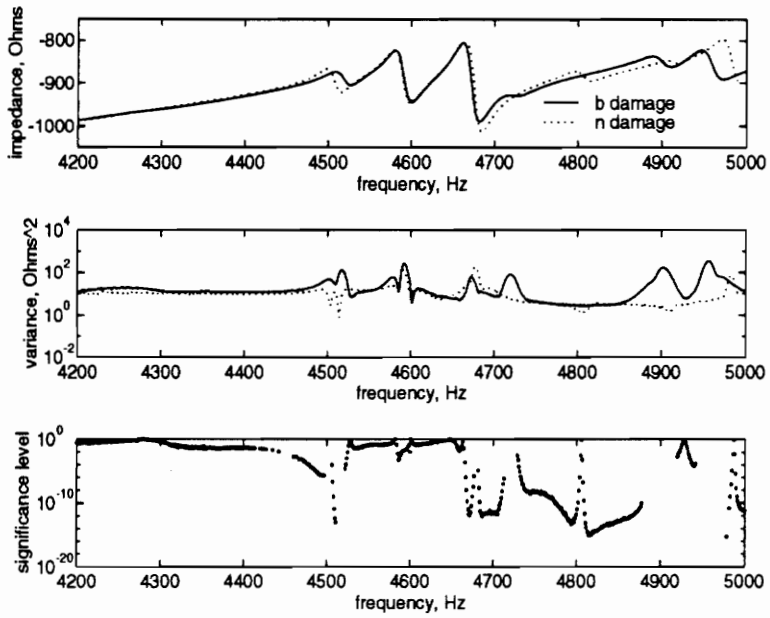


Figure 14: Imaginary impedance response, green sensor.

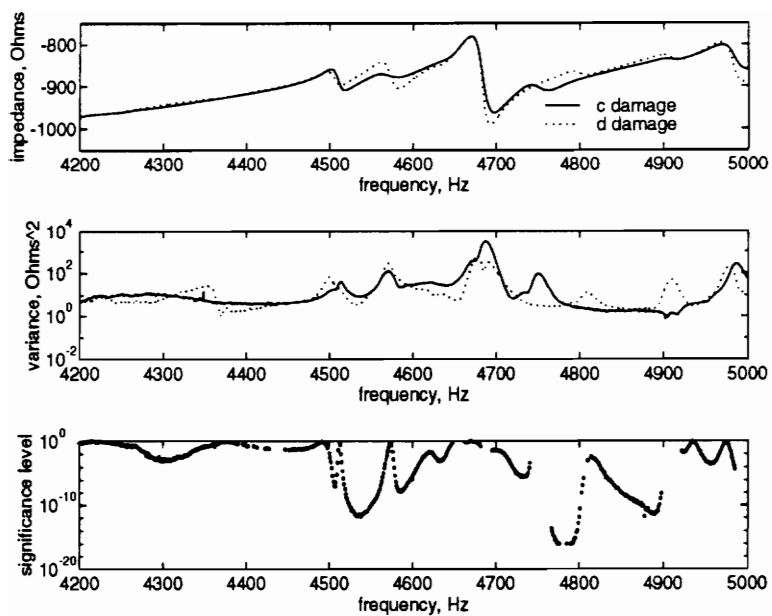


Figure 15: Imaginary impedance response, red sensor.

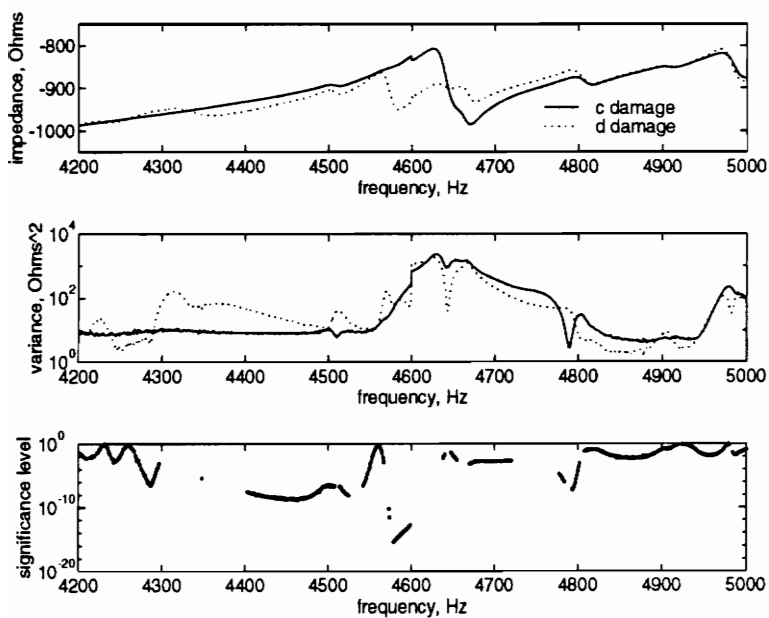


Figure 16: Imaginary impedance response, green sensor.

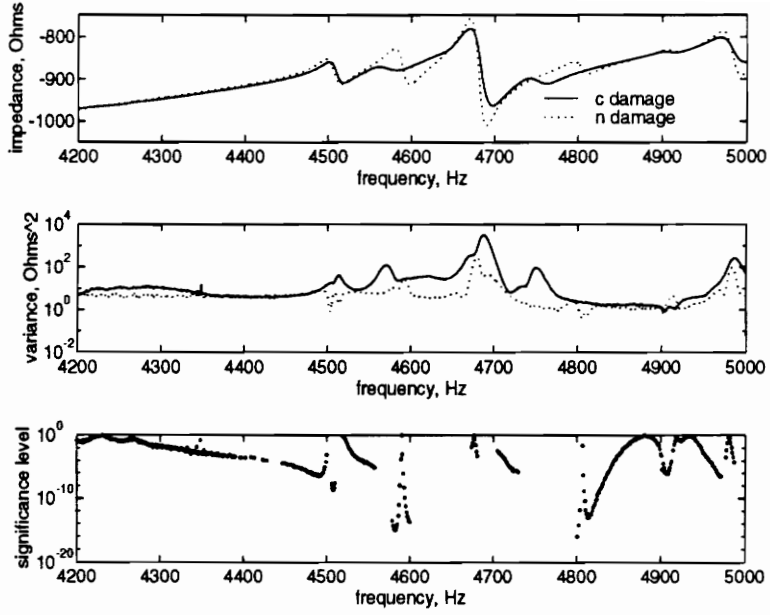


Figure 17: Imaginary impedance response, red sensor.

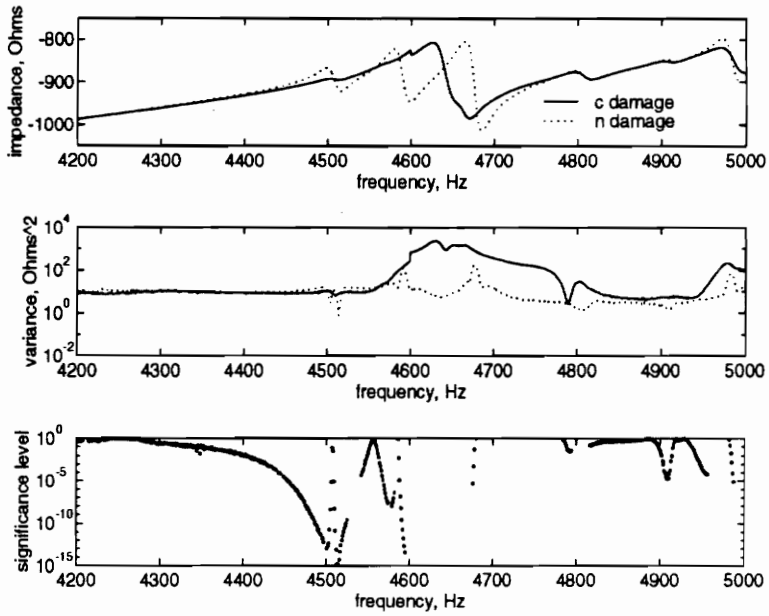


Figure 18: Imaginary impedance response, green sensor.

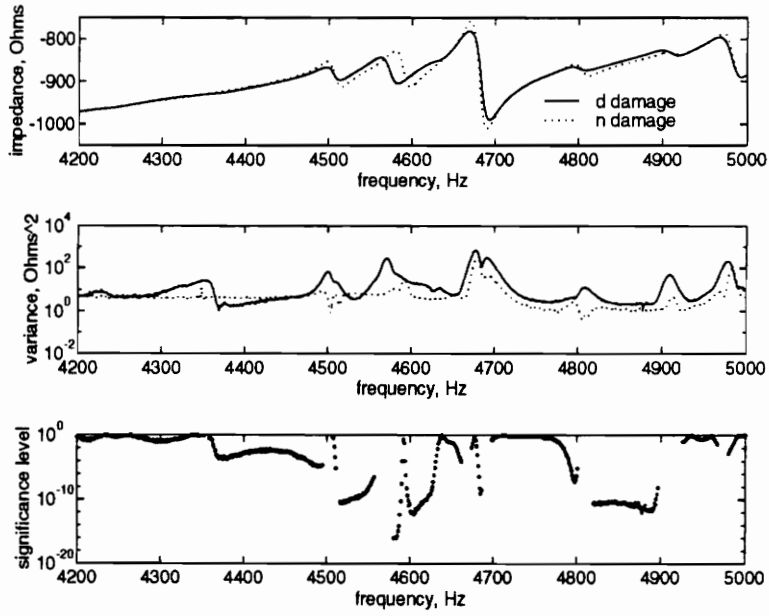


Figure 19: Imaginary impedance response, red sensor.

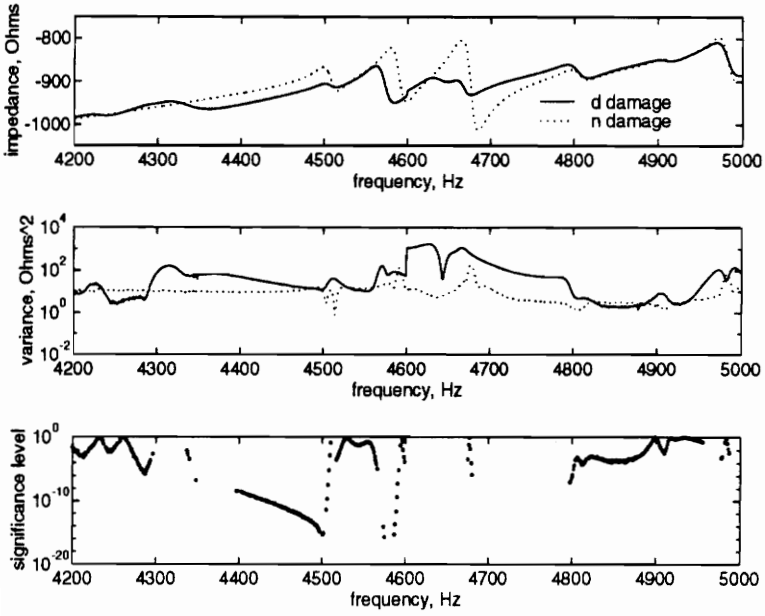


Figure 20: Imaginary impedance response, green sensor.

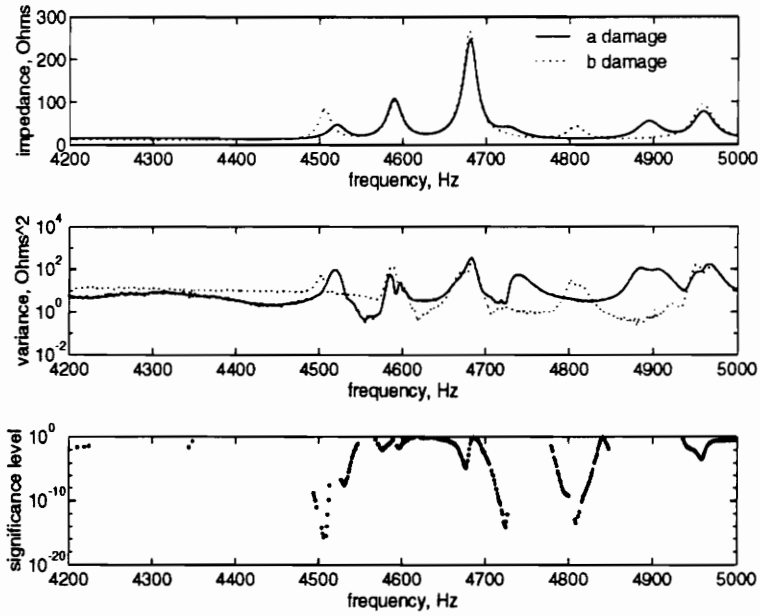


Figure 21: Real impedance response, red sensor.

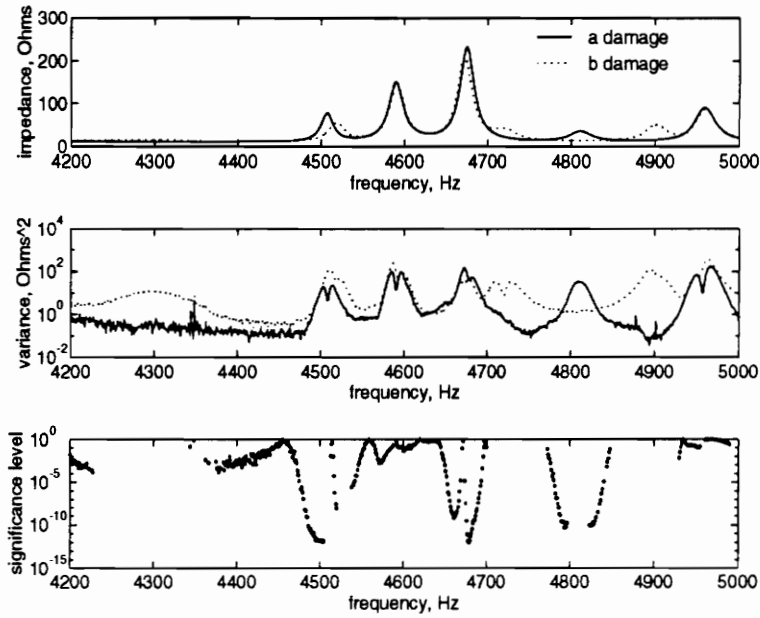


Figure 22: Real impedance response, green sensor.

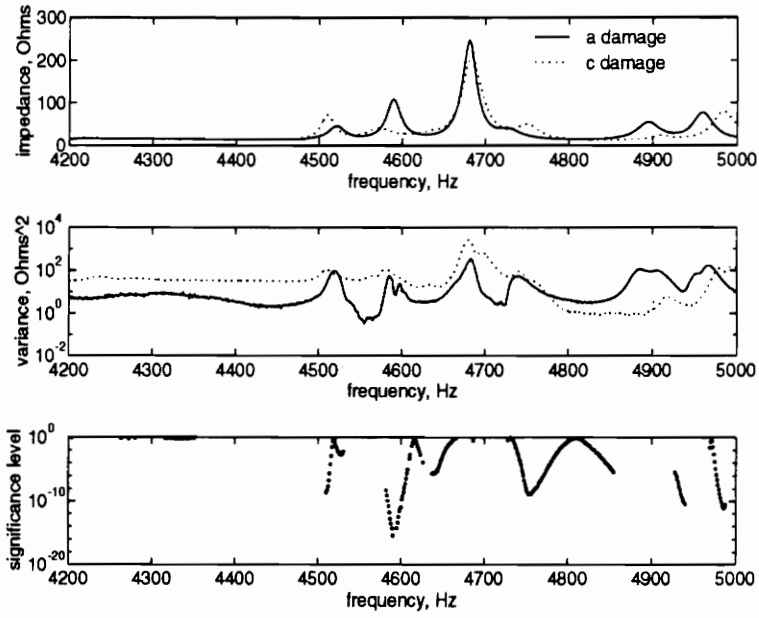


Figure 23: Real impedance response, red sensor.

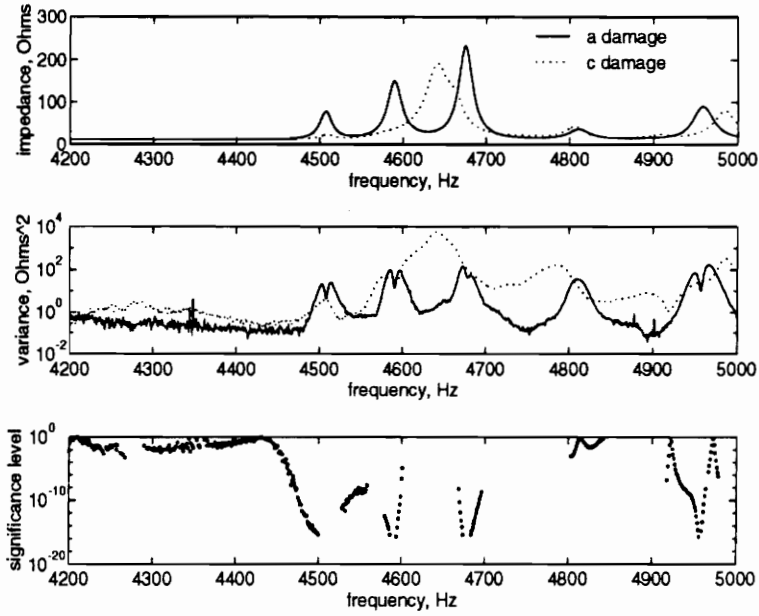


Figure 24: Real impedance response, green sensor.

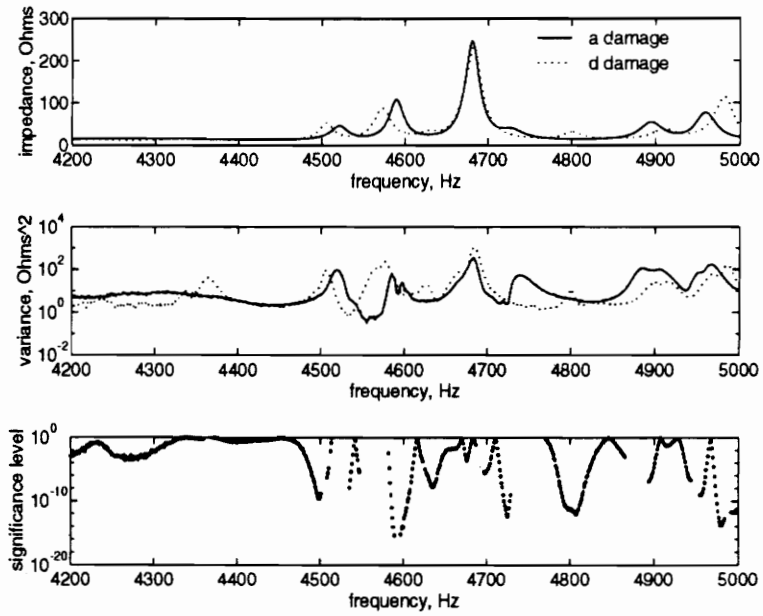


Figure 25: Real impedance response, red sensor.

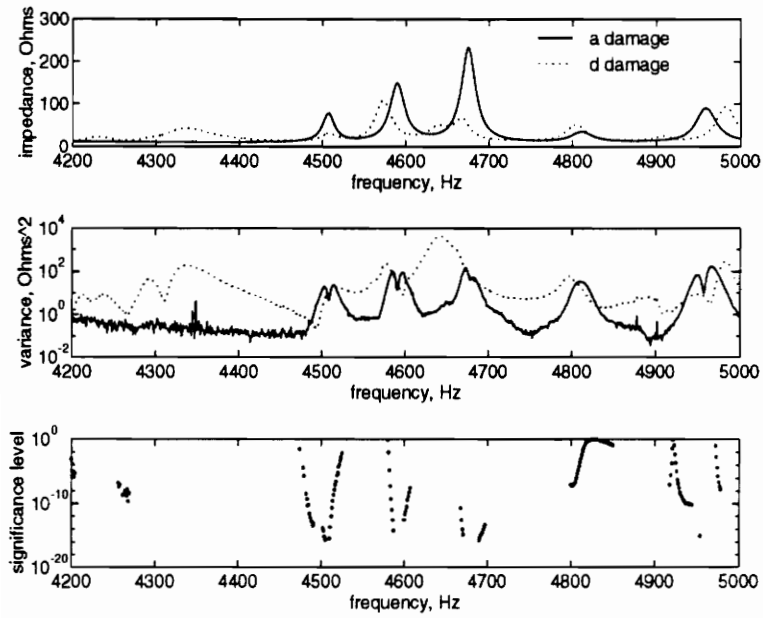


Figure 26: Real impedance response, green sensor.

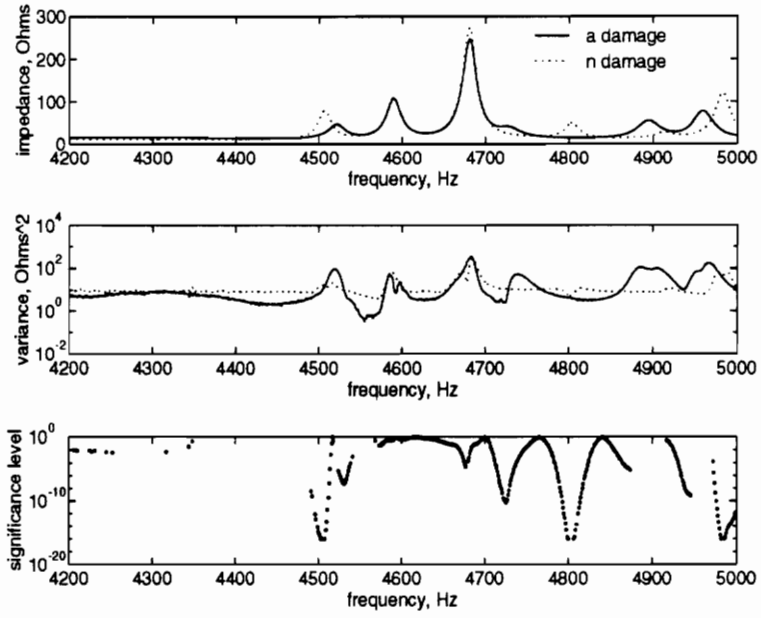


Figure 27: Real impedance response, red sensor.

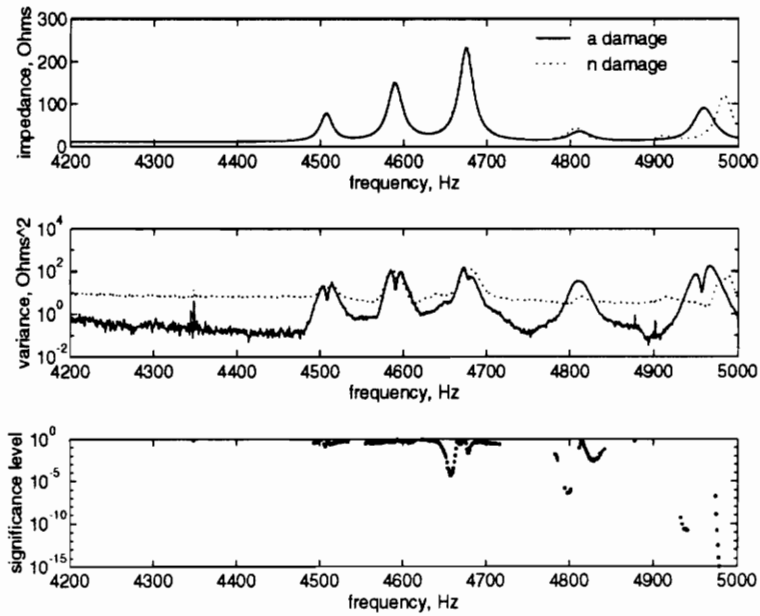


Figure 28: Real impedance response, green sensor.

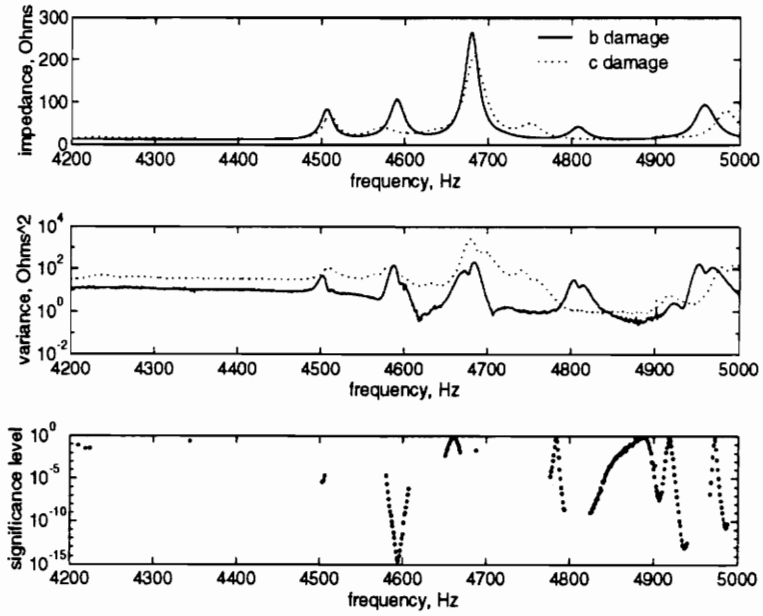


Figure 29: Real impedance response, red sensor.

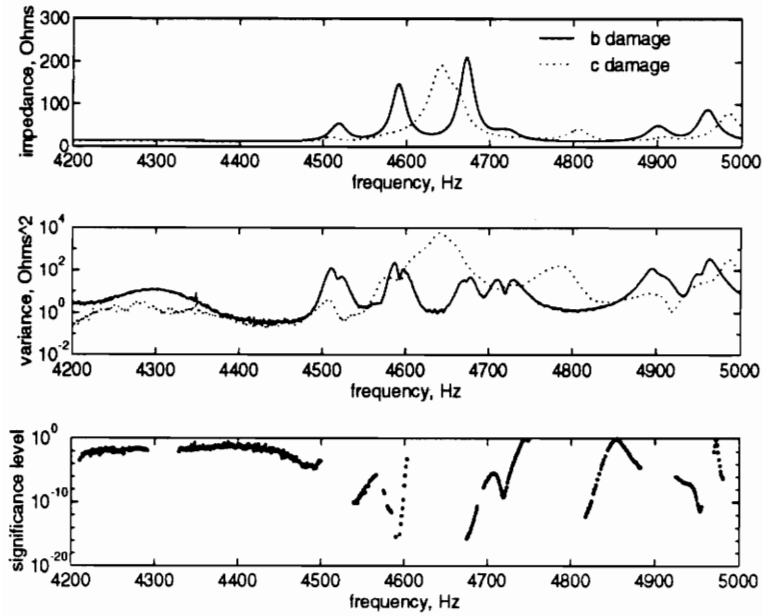


Figure 30: Real impedance response, green sensor.

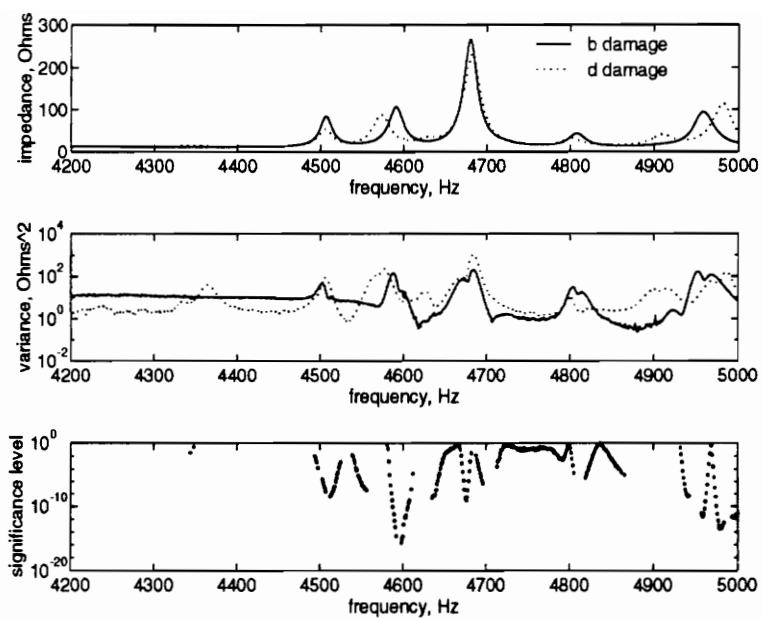


Figure 31: Real impedance response, red sensor.

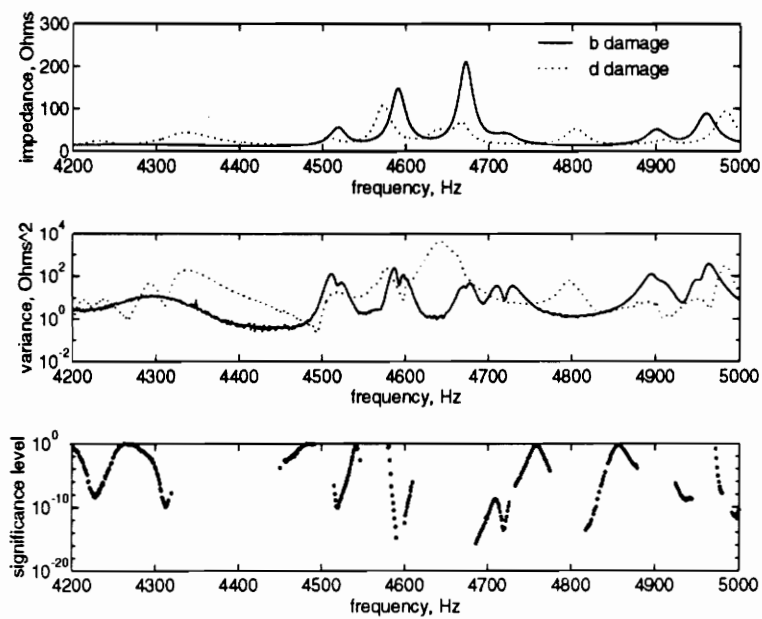


Figure 32: Real impedance response, green sensor.

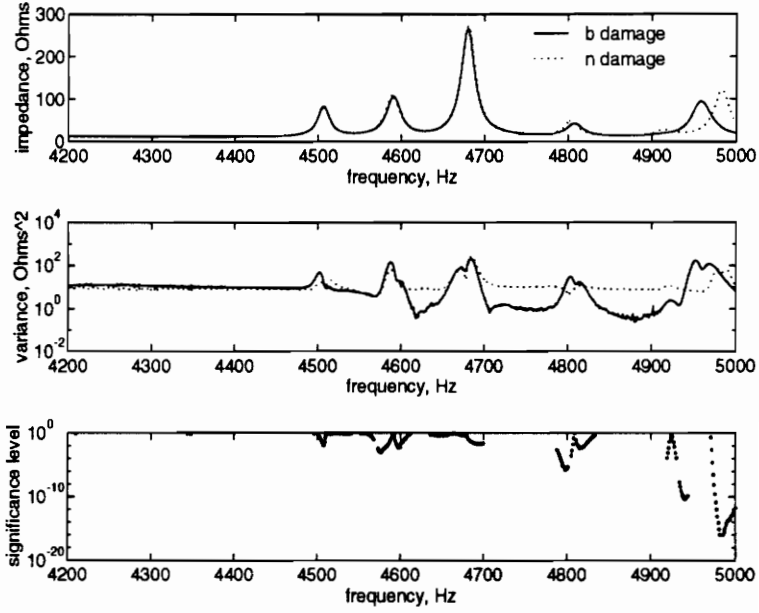


Figure 33: Real impedance response, red sensor.

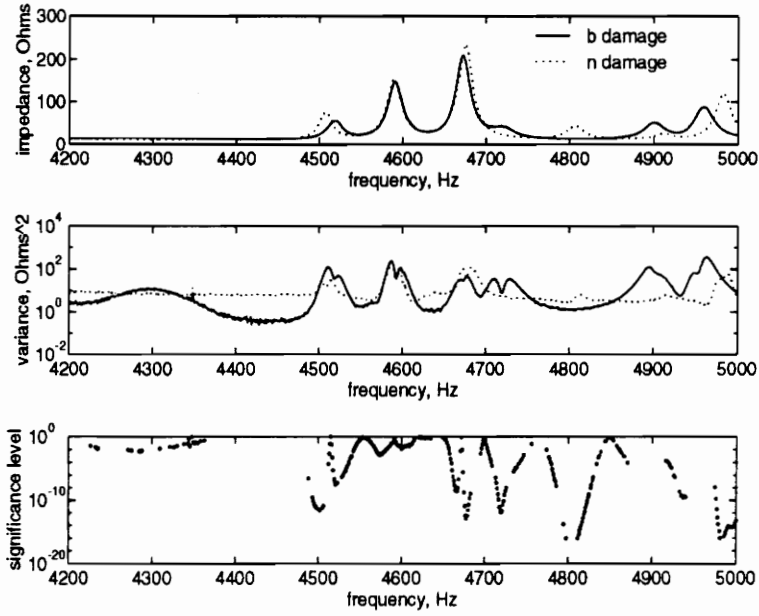


Figure 34: Real impedance response, green sensor.

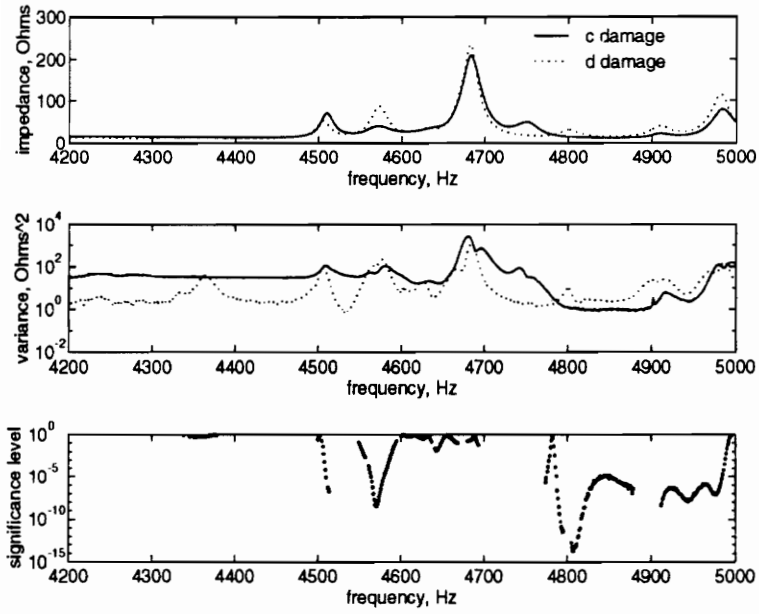


Figure 35: Real impedance response, red sensor.

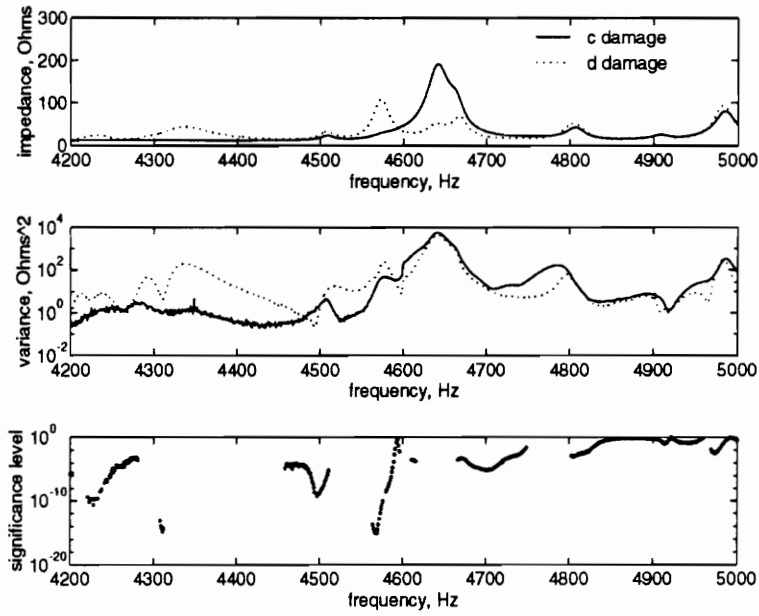


Figure 36: Real impedance response, green sensor.

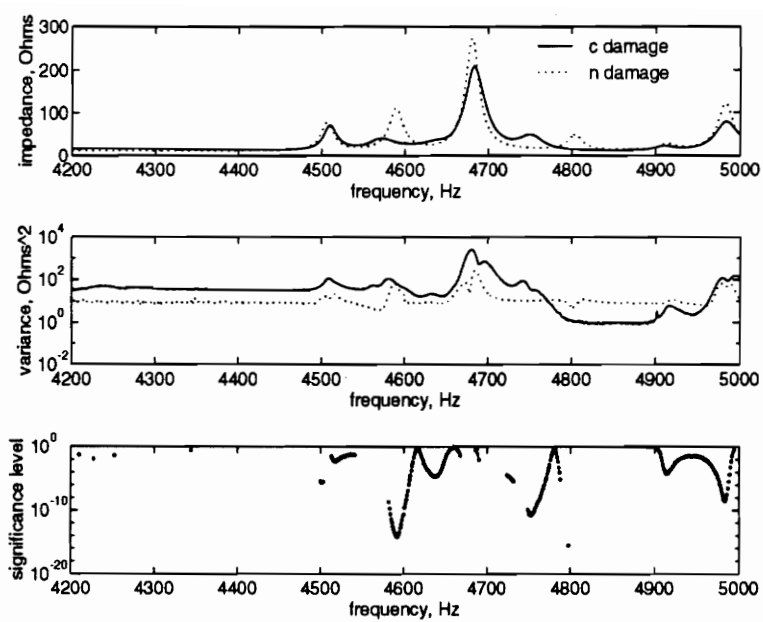


Figure 37: Real impedance response, red sensor.

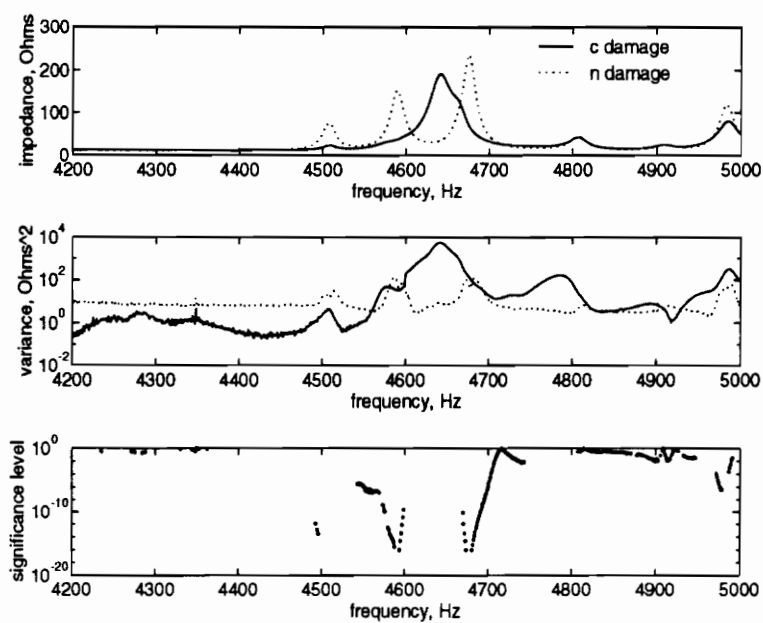


Figure 38: Real impedance response, green sensor.

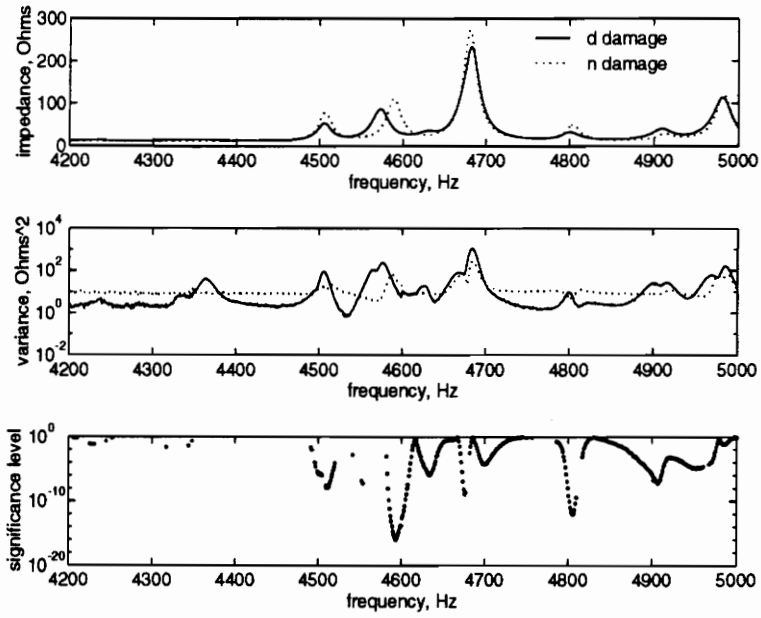


Figure 39: Real impedance response, red sensor.

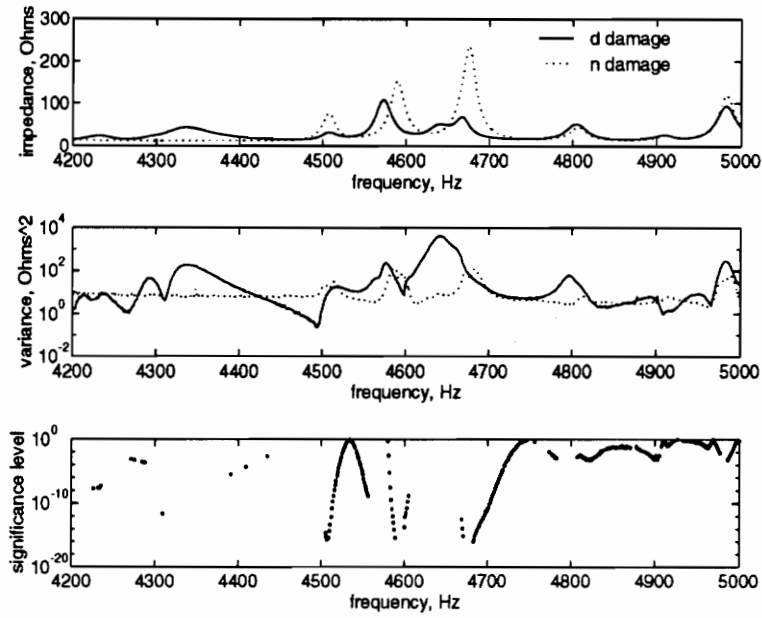
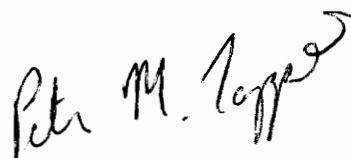


Figure 40: Real impedance response, green sensor.

Vita

Peter Matthew Tappert was born in Towson, Maryland on March 29, 1971. He spent his youth in suburban Largo, Maryland. In 1989, he began to pursue a degree in Engineering at Virginia Tech. He mosied through three and a half years of the Mechanical Engineering curriculum just to find out that he had no clue as to what he wanted to do. At that point he began to do undergraduate research... and he liked it! In May 1993, he received a Bachelor's degree in Mechanical Engineering from Virginia Tech and decided to pass up an opportunity to teach in the Peace Corps in order to continue researching in graduate school. Throughout his year and a half in graduate school, his future plans have changed weekly. The only definite future plan he has is to change his plans for the future.

A handwritten signature in black ink that reads "Peter M. Tappert". The signature is written in a cursive style with a large, sweeping flourish at the end of the name.