NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

CLASSIFICATION ANALYSIS OF VIBRATION DATA FROM SH-60B HELICOPTER TRANSMISSION TEST FACILITY

by
Gregory L. Anderson

September 1997

Thesis Advisor: Robert R. Read

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE September 1997	3. REPORT TYPE AND DATES COVERED Master's Thesis

4. TITLE AND SUBTITLE Classification Analysis of Vibration Data From SH-60B Helicopter Transmission Test Facility	5. FUNDING NUMBERS
6. AUTHOR(S) Anderson, Gregory L.	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
Approved for public release; distribution unlimited.	

ABSTRACT *(maximum 200 words)*

The U.S. Navy is currently evaluating an integrated diagnostic system for its rotary wing aircraft. The system is referred to as the Health Usage and Monitoring Systems (HUMS). The program's objective is to develop an automated diagnostic system that can identify mechanical faults within the power train of helicopters using vibration analysis. This thesis uses data provided by the Helicopter Transmission Test Facility at the Naval Air Warfare Center, Trenton, New Jersey. The goal of this thesis is to conduct data analysis to identify a fault within the helicopter test transmission using a tree-structured model. Prior to conducting tree analysis, an attempt is made to reduce the amount of data by principal component analysis. All statistical analysis was completed with S-Plus Software (MathSoft Inc., 1995).

14. SUBJECT TERMS HUMS, Helicopter Maintenance, Vibration Analysis, Classification Analysis, tree-structured classification	15. NUMBER OF PAGES 67
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFI-CATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 39-18

Approved for public release; distribution is unlimited.

CLASSIFICATION ANALYSIS OF VIBRATION DATA FROM SH-60B HELICOPTER TRANSMISSION TEST FACILITY

Gregory L. Anderson
Lieutenant, United States Navy
B.S., Georgia Institute of Technology, 1989

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1997

Author: Gregory L. Anderson

Approved by: Robert R. Read, Thesis Advisor

Harold J. Larson, Second Reader

for Richard Rosenthal, Chair
Department of Operations Research

ABSTRACT

The U.S. Navy is currently evaluating an integrated diagnostic system for its rotary wing aircraft. The system is referred to as the Health Usage and Monitoring Systems (HUMS). The program's objective is to develop an automated diagnostic system that can identify mechanical faults within the power train of helicopters using vibration analysis. This thesis uses data provided by the Helicopter Transmission Test Facility at the Naval Air Warfare Center, Trenton, New Jersey. The goal of this thesis is to conduct data analysis to identify a fault within the helicopter test transmission using a tree-structured model. Prior to conducting tree analysis, an attempt is made to reduce the amount of data by principal component analysis. All statistical analysis was completed with S-Plus Software (MathSoft Inc., 1995).

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The United States Navy is conducting research on technology which evaluates the mechanical health of a helicopter transmission. Known as the Health and Usage Monitoring System (HUMS), this technology originated in the United Kingdom for helicopter operations in the North Sea. The United States Navy has used both fleet and ground facilities to test its version of HUMS with commercial-off-the-shelf technology. In particular, Naval Air Warfare Center (NAWC), Trenton, New Jersey, in conjunction with Technology Integrated Incorporated (TII), conducted comprehensive studies of HUMS technology on which the thesis will focus.

The Naval Air Warfare Center's comprehensive program on SH-60 helicopters is called the Helicopter Integrated Diagnostic System. The program uses state of the art data acquisition, raw data storage, and algorithmic analysis provided by Technology Integration Incorporated (TII) to evaluate the propulsion and power drive system. Ground testing at NAWC provides fault detection validation in a full scale helicopter transmission test facility (HTTF) of the SH-60 power drive system. Twenty-nine sensors (accelerometers) are located throughout the transmission. These specialized sensors measure the vibration generated by gears, bearings and shafts. Each sensor measures the vibration, rpm counts and other signals of components located near it. The raw data is composed of the signatures collected from all dynamic components of the system.

Seeded faults for different components are placed in the transmission for data acquisitions. Specialized algorithms, proprietary to TII, serve as indicators of faults and their location. The indicators are the output medium by which a fault is then determined.

The data acquisitions are segregated by bearings, shafts, and gears. They are then written into Matlab matrices.

This thesis focuses on data derived from the sensors of the intermediate gear box input pinion preload bearing and the port main bevel pinion timken bearing. Its basis is to determine whether or not a faulty pinion can be distinguished by indicators calculated for a bearing using a tree classification model. The seeded faults of interest are a small integral race spall in the port main spiral bevel pinion and a edm notch in the intermediate gear box input pinion.

The pre-load bearing is located near the same sensors as the intermediate gear box input pinion. Likewise, raw data for the timken bearing originates from the same sensors as the port main spiral bevel pinion. Data acquisitions of all applicable sensor/indicators for each fault are evaluated. The data acquisitions are analyzed using principal component analysis for possible data reduction.

The tree-structured model is applicable to the HUMS research. It identifies the threshold values of indicators provides a logical decision tree for predicting the presence of a fault. Further research is necessary to totally unlock the potential of tree-classification modeling potential with the HUMS.

I. INTRODUCTION

With the increasing demand for helicopters, it is imperative that the services maintain a high operational readiness rate. The increase in usage requires reliable equipment and a structurally sound airframe. This particular platform creates a unique challenge to the services because of its dynamics and rotating machinery. [Ref. 1] Meeting this challenge would provide significant benefits in aviation safety and mission readiness.

Currently, the United States Navy is evaluating technology which could bring it closer to an ideal readiness rate. Known as the Health and Usage Monitoring System (HUMS), this technology was developed in the United Kingdom for operations in the North Sea. Transport helicopters were experiencing an unacceptable number of mechanical failures resulting in casualties. HUMS' basic concept was to use vibration analysis to detect mechanical faults in the transmission of the aircraft. It was conceived as a viable technology when the aviation press reported the first instance of a helicopter being grounded before a flight on evidence from an onboard health and usage monitoring system. In 1991, the British began to realize the huge potential that existed in terms of reduced costs upon HUMS integration in their maintenance program.[Ref. 2]

The United States Navy is investigating the benefits of HUMS. It has used both fleet and ground facilities to test its version of HUMS with commercial-off-the-shelf technology. In particular, Naval Air Warfare Center (NAWC), Trenton, New Jersey along with Technology Integrated Incorporated (TII) conducted comprehensive studies of HUMS technology.

A. BENEFITS OF HUMS

The use of aircraft-mounted sensors to monitor and record vibration, flight control positions and other parameters provide useful information to the pilots and ground crew regarding the health of the aircraft. Diagnostic data regarding the health and usage of an aircraft will provide tremendous improvements in how the United States Navy ensures safety and conducts maintenance of its helicopters. Operational readiness and maintenance savings would be increased significantly.

1. Maintenance

Currently, the United States Navy provides maintenance based upon flight hours of the aircraft. Parts are replaced in accordance with the maintenance cycle or as needed. In many instances, it is believed that parts replaced are without fault. Healthy parts which cost thousands of dollars may be prematurely extracted from helicopters. HUMS can help reduce and possibly alleviate unnecessary parts replacement by providing ongoing mechanical diagnosis of the aircraft.

The current system allows for very conservative safeguards in terms of ensuring that new parts are periodically installed. However, this process cannot control two important factors; the actual health of the new or refurbished part and human error. The probability of faulty parts in the stock system is not negligible. Even more, maintenance has its own inherent dangers because of human mistakes.

Routine squadron maintenance supplemented by HUMS analysis provides an excellent diagnostic approach. The health status of individual components in the aircraft would be available without removal or replacement. Also, HUMS would reduce functional check flights required after certain repairs. Continuous-monitoring, as opposed to time based maintenance, provides the greatest potential for parts and man-hours savings.

Additionally, HUMS provides an excellent supplement to quality assurance of the maintenance process. Procedural requirements of helicopter maintenance include inspection, paper work review by the maintenance control, a safe-for-flight authorization, and pilot approval. Critical components require three individuals to perform the maintenance. These steps are evidence that quality assurance is integrated into the maintenance process. However, human factors are a part of each process and pose the potential for error. HUMS' fault detection validation is a powerful supplement to the present quality assurance process.

2. Safety

Aircraft mishaps are evaluated for five possible causal factors: supervisory, air crew, facilities, material and maintenance. HUMS can isolate and reduce mishaps resulting from material and maintenance origins. Maintenance personnel, alerted by HUMS to problems, could immediately initiate corrective measures to prevent impending material failure problems. In addition, a continual update of exceedence parameters on all helicopter models and series can be exercised to include data garnered from mishaps due to recurring component failures. [Ref. 3]

As previously suggested, HUMS was originated for safety reasons. Its integration into Naval helicopters depends primarily upon its ability to provide accurate information about the health of a helicopter. It is at the forefront of technologies that provide the greatest potential for identifying impending mechanical failures.

B. LIMITATIONS OF HUMS

The health and usage monitoring system represents the cutting edge of today's helicopter operational safety and diagnostics technology. However, HUMS is not without its share of problems. Services across Europe have implemented their version of HUMS, but not without experiencing difficulties. The primary obstacles are data quality, false alarms and missed faults..

1. Data Quality

It is essential that the data provided for diagnosing the health of a helicopter is accurate. The usefulness of analyses of vibrations emanating from the bearings, gears, and shafts of the transmission is affected by the reliability of the accelerometers. Data quality extends into the implementation of the system. Every conceivable effort must be taken to ensure proper placement of sensors and cabling. If poor data is collected the results are worthless.

Along with the issue of data quality comes the question of data maintenance. In evaluating the health of certain components, HUMS makes a determination in one of two ways. The data for the component may exceed a defined limit called a threshold, or it might exceed a limit based on its trends. In order for the trending capability to be useful, the data for each specific component must be archived and carried along with it if it is removed and placed in another aircraft. Each critical component, as well as each aircraft, must maintain its own database for HUMS to be effective. [Ref. 1].

2. Errors

Two types of errors may occur in using HUMS, the false positive and false negative indications. The false positive alarm occurs when HUMS indicates that a healthy component has experienced some sort of fault. The false negative is the most dangerous error because HUMS fails to give warning in the case of a faulty component.

a. False Positive Alarms

Threshold values are predetermined limits set on specific components monitored by HUMS. Nearly eighty percent of the United Kingdom aircraft integrated with HUMS exceed threshold limits and do not have any faulted components. Low threshold values are the cause of these high false alarm rates. The frequency of such alarms put an organization in a situation where decisions must be made concerning the safety of their aircraft. Either excessive maintenance demands and reduced operational availability can occur or the aircraft is flown under the premise of a false alarm. Effectiveness in the system is lost in either case.

The threshold is a value set for a specific component of the aircraft that is monitored by a HUMS sensor. The HUMS sensor takes a reading from the component and compares the value of the reading to the threshold value. The challenge is to set the threshold limits to values that do not compromise safety or cause excessive false alarm rates.

b. False Negative Indication

A false negative indication occurs if no warning of a fault is given when there is a fault present. Similar to the false positive indication, the act of setting the threshold value to the appropriate level is a design challenge for the system. [Ref 4]

C. SCOPE OF THESIS

Analysis of the system will be based on the data from a developmental HUMS at NAWC, Trenton, New Jersey. Chapter II will describe the employment of HUMS in the aircraft transmission. Chapter III will describe principal components analysis as a data reduction method. Also, it will explain the non-parametric technique used to uncover structures in a data set - Classification Trees. The models and results of the analysis will be provided in Chapter IV. Finally, Chapter V will discuss their usefulness.

II. BACKGROUND

A. HELICOPTER INTEGRATED DIAGNOSTIC SYSTEM (HIDS)

The Naval Air Warfare Center Aircraft Division (NAWCAD - Trenton, NJ) is conducting a comprehensive program on SH-60 helicopters called the Helicopter Integrated Diagnostic System. The SH-60B was chosen as the ideal platform due to the large fleet throughout the Department of Defense services and its high potential for support. The program uses state of the art data acquisition, raw data storage, and algorithmic analysis provided by Technology Integration Incorporated (TII) to evaluate the propulsion and power drive system.

Ground testing at NAWCAD provides fault detection validation in a full scale helicopter transmission test facility (HTTF) of the SH-60 power drive system. The power drive consists of engines, transmission and tail drive system. As many as thirty-two accelerometers/sensors can be located throughout the power train. These specialized sensors measure the vibration generated by gears, bearings and shafts. Each sensor can be affected by the vibration of many component sources. The raw data is composed of the vibration signatures collected from all dynamic components from the loaded parts of the system.

A data acquisition consists of 30 seconds or less of recording time; typical acquistions are made over 4 to 10 seconds of operation. The records are obtained simultaneously from (up to) 32 accelerometers at 100,000 samples per second. The sampling rate of the system exceeds NAWC's requirements for a total on-board health and usage monitoring system. The records are written into Matlab matrices in the following categories: bearings, gears, and shafts.

B. FAULT EVALUATION

Vibration signatures can be acquired by injecting known faulty parts in the power drive system. Parts rejected by the fleet and turned in for overhaul have been aside for testing. These parts provided natural faults created by the SH-60 drive train. However, due to the scarcity of fleet rejected parts, good parts which have been damaged artificially are also used to imitate faults of interest. Testing has been concentrated initially on the tail drive system in order to verify the TII/BFG diagnostic system operation and performance. Subsequent testing was performed on the engines, input modules, hydraulic pumps and main gearbox. The test conditions consist of a set of sequential variations of power settings throughout the normal range of operation. Such power variation is essential to understand the sensitivity of the diagnostic algorithms as a function of changing aircraft power. Ambient temperature variation effects can also be taken into account in the analysis. The first data set of each test run is taken at low torque before the oil is warm. This provides a data base that can be compared to flat pitch maintenance ground turns for troubleshooting.

Many different computed indicators can be evaluated for each data acquisition. The indicators are the output medium and are used to identify fault thresholds. These indicators are not the same for each of the three component categories: bearing, gears, and shafts. For each data acquisition, data is received from each accelerometer that senses the component. [Ref. 5] The composition of the algorithms (indicators) are proprietary to TII/BFG and not in the scope of the thesis.

This thesis will focus on data collected from sensors located near two well separated bearings. The sensor data has been used by TII to compute measures (indicators) that are expected to detect bearing faults. There are 28 independent bearing indicators per sensor. Of the two bearings of interest, one has two sensors located near it and the other has three. The fundamental indicators are the following: bdf, Iraw_pk2, Iraw_cf, Iraw_sv, Iraw_kv, Iraw_rms, EBpk2pk, EBcf, EBsv , Ebkv, EBRms, rte, rbe, te, be, ce, bse, ie, oe, tbe, counter, EBRms, BC1, BC2, BC3, BC4, BC5, and BC6.

III. ANALYSIS

A. METHODOLOGY

1. Overview of Principal Component Analysis

For investigations involving a large number of observed variables, it is often useful to simplify the analysis by considering a smaller number of variables, such as a linear combinations of the original variables. Principal components seek a few underlying dimensions that account for patterns of variation among the observed variables. These underlying dimensions can provide ways to combine variables, simplifying subsequent analysis. For example, a few combined variables could replace many original variables in a regression. Advantages of this approach include more parsimonious models, improved measurement of indirectly observed concepts, new graphical displays, and the avoidance of multi-collinearity.

Principal components is not model based. It involves a straightforward mathematical transformation. Data on K observed variables can be re-expressed as data on K principal components. The K principal components explain all the variability of the original K variables. Data reduction is accomplished when fewer than K components account for most of the variance. If only J of the largest components (J<K) are retained, we can disregard the rest. [Ref. 6] Also of importance is the potential to discover a few of the original variables that (in effect) determine the dominant principal components.

2. Example of Principal Component Analysis

Principal components can be applied to the scholastic achievement test (SAT). The SAT typically consist of a number of examinations in different subject areas. In attempting to rate students applying for admission, college administrators frequently attempt to reduce the scores from all subject areas to a single, overall score. If the reduction can be done with minimal information loss, all the better.

An obvious choice for overall score is the mean over all subject areas. For three subject areas s_1, s_2, and s_3, the mean corresponds to the linear combination $⅓s_1+⅓s_2+⅓s_3$, or equivalently the use of weighing vector $\boldsymbol{l}$, where $\boldsymbol{l}$ is the vector of coefficients $(⅓,⅓,⅓)^T$. A linear combination with $\sum l_i^2 = 1$ is called a standardized linear combination, or SLC. By restricting attention to SLCs, one can make meaningful comparisons between various choices of linear combinations. For example, with test scores, one can seek the combination with the greatest variance as a way of both ranking the students and separating them.

Principal components analysis finds a set of SLCs, called the principal components, which form an orthogonal set of vectors and taken together explain all the variance of the original data. The principal components are defined as follows:

If x is a random vector with mean μ and covariance matrix Σ, then the principal components transformation requires us to find a matrix, Γ,

$$x \rightarrow y = \Gamma^T(x - \mu)$$

where Γ is orthogonal, $\Gamma^T\Sigma\Gamma = \Lambda$ is diagonal, and $\lambda_1 \geq \lambda_2 \ldots \geq \lambda_p \geq 0\ldots$ The ith principal component of x may be defined as the ith element of the vector y, namely as

$$y_i = \gamma^T_{(i)}(x - \mu).$$

Here $\gamma_{(i)}$ is the ith column of Γ, and is called the ith vector of principal component loadings.

The first principal component has the largest variance among all SLCs of x.

Similarly, the second principal component has the largest variance among all remaining SLCs of x and is not correlated with the first principal component, and so on.

In general, there are as many principal components as variables. However, it is usually possible to consider only a few of the principal components, which together explain most of the original variation.

Table 3-1 shows the results of qualifying examinations for 25 graduate students in mathematics at a fictional university. The students sat for examinations in each of five subject areas - differential geometry, complex analysis, algebra, real analysis, and statistics.

	diffgeom	complex	algebra	reals	statistics
1	36	58	43	36	37
2	62	54	50	46	52
3	31	42	41	40	29
4	76	78	69	66	81
5	46	56	52	56	40
6	12	42	38	38	28
7	39	46	51	54	41
8	30	51	54	52	32
9	22	32	43	28	22
10	9	40	47	30	24
11	32	49	54	37	52
12	40	62	51	40	49
13	64	75	70	66	63
14	36	38	58	62	62
15	24	46	44	55	49
16	50	50	54	52	51
17	42	42	52	38	50
18	2	35	32	22	16
19	56	53	42	40	32
20	59	72	70	66	62
21	28	50	50	42	63
22	19	46	49	40	30
23	36	56	56	54	52
24	54	57	59	62	58
25	14	35	38	29	20

Table 3-1: Examination scores for graduate students in mathematics

The differential geometry and statistics examinations were closed book and the remaining examinations were open book. The summary showing the importance of the calculated principal components are shown in Table 3-2.

	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp.5
Standard Deviation	28.4897	9.0355	6.6009	6.1336	3.7234
Proportion of Variance	0.8212	0.0826	0.0440	0.0381	0.0140
Cumulative Proportion	0.8212	0.9038	0.9471	0.0986	1.0000

Table 3-2: Summary of calculated principal components

In this example, the first component explains 82% of the total variance, and the first two principal components together explain 90% of that variance.

The principal component loadings are the coefficients of the principal components transformation. They provide a convenient summary of the influence of the original variables on the variance of the principal components, and thus a useful basis for interpretation. A large coefficient (in magnitude) corresponds to a high loading, while a coefficient near zero has a low loading.

	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp. 5
Diffgeom	0.598	-0.675	-0.185	-0.386	0
Complex	0.361	-0.245	0.249	0.829	-0.247
Algebra	0.302	0.214	0.211	0.135	0.894
Reals	0.389	0.338	0.7	-0.375	-0.321
Statistics	0.519	0.57	-0.607	0	-0.179

Table 3-3: Loading factors for test scores

The loadings for the first principal component (Table 3-3) are all the same sign and of moderate size, although, differential geometry and statistics tend to dominate. A reasonable interpretation is that this component represents a weighted average score for the five qualifying examinations. The second component contrasts the two closed book exams with the three open book exams, with the first and last exams weighted most heavily - and so forth. [Ref. 7]

B. OVERVIEW OF TREE-STRUCTURED CLASSIFICATION

Tree-based models can be used either for prediction (similar to a regression analysis) or for classification. They use a principle known as binary recursive partitioning to achieve this goal. Basically, at each step of the tree-building process, the values of the independent variables are examined for all possible binary splits of the data to find the split that most effectively separates the dependent variable into homogeneous groups. For continuous independent variables the splits are defined by a single value: an observation goes into one node if its value is less than or equal to the split value, and into the other node if its value is greater than the split value. For factors, all possible partitions of the levels into two non-overlapping groups are considered. Because of the lack of assumptions, these models perform well in cases where more parametric models might not be effective. [Ref. 8]

1. Example of Tree-Structured Classification

To identify a car owner's satisfaction with a new car, a tree-structured classification can be useful. By way of introduction to tree-structured classification, a car owner satisfaction example will be discussed.

Sixty-nine new car owners were surveyed on their overall satisfaction with their cars. Five factors were observed for each car: turning circle, weight, miles per gallon, price, and length. Each factor is a continuous variable and is also called an independent variable. Each car owner represents a case and falls into one of two classes. A satisfied owner falls in class 1 which is designated by "true." A dissatisfied owner falls into class 2 which is designated by "false." Each car owner is represented by a data point which is called a case.

A classification tree recursively splits the car owners into two classes according to the value of one of the independent variables. An ideal goal would be purity in these nodes. By definition, purity (or homogeneity) means that all the cases in a single terminal node have exactly the same dependent variable classification. In the car satisfaction example, a homogeneous node would be one in which either all new car owners in that node are satisfied in one instance or all not satisfied in another.

The root node of this binary classification tree contains all the cases in the data set. From this node, a determination is made regarding a split of the data into two separate "child" nodes. At each node the tree algorithm searches through *M* independent variables one by one, beginning with x_1 and continuing up to x_M. For our example, $M = 5$ and x_1 = "turning circle," x_2 = "weight," x_3 = "mile per gallon," x_4 = "price," and x_5 = "length." Considering each independent variable separately, it evaluates the change in homogeneity if all the cases in that node were separated by a value of that variable. That is, a split is chosen at a specific value, *j*, of a single independent variable, x_i. The right child node gets all cases for which $x_i > j$ and the left child node gets all cases for which $x_i < j$. Considering the data at the root node of our car satisfaction example, the algorithm evaluates every possible split of the cases, and picks the variable and splitting value that gives the greatest improvement in homogeneity. It first checks the turning circle variable. It evaluates the change in purity for splits made between distinct values of turning circle observed in the data set. It then does the same for the splits made between distinct values of weight, length, miles per gallon, and price, respectively. From all the possible splits, the algorithm chooses the one that gives the greatest improvement in purity. [Ref. 8]

S-Plus (Mathsoft Inc.) uses the deviance (likelihood statistic) to measure the homogeneity of the node. At each node *i* of a classification tree, there is a terminal vector of the probabilities over the k classes. Each case in node *i* is assumed to be drawn from a multinomial distribution. At node *i*, n_{ik} cases are observed in class k, where $\sum_k n_{ik} = n_i$.

The deviance at a node is defined as the negative of twice the log-likelihood,

$$D_i = -2\sum_k n_{ik} \log p_{ik} .$$

Since we do not know the probabilities, we must estimate them for node i. We now determine if node i should be split into two child nodes *l* and *r*. The split is made to maximally decrease the deviance of the node (i.e., maximize

$$\Delta D_i = D_i - D_l - D_r,$$

which measures the decrease in deviance or increase the homogeneity). [Ref. 9]

Using the data from our example, the deviance of the root node is computed for illustration. The two classes of new car owners are "TRUE," and "FALSE." Thus, each case in the root node is assumed to be drawn from a multinomial distribution with k = 2. If $\mu_1 = (p_{11,}p_{12})$, then p_{11} = prob("true") and p_{12} = prob("false"). At the root node, there are a total of $n_1 = 69$ cases, $n_{11} = 29$ with level "true" and $n_{12} = 40$ with level "false," giving $p_{11} = 29/69$ and $p_{12} = 40/69$, and the deviance at the root node is equal to

$$-2[29\ln(29/69) + 40\ln(40/69)] = 93.8932.$$

The first split of the cases in the example is made on turning radius. The split is made such that all the cases with a turning radius < 39.5 ft. be allocated to the left child node and all the cases with a turning radius > 39.5 ft. be allocated to the right child node. The split results in $n_2 = 28$ cases in the left node and $n_3 = 41$ cases in the right node. Of the 28 cases in the left node, $n_{21} = 22$ have the level "true" and $n_{22} = 6$ have the level "false." Of the 41 cases in the right node, $n_{31} = 34$ have the level "false" and $n_{32} = 7$ have the level "true." The resultant deviance is the sum of the deviances of the two child nodes,

$$-2[6\ln(6/28) + 22\ln(22/28)] - 2[34\ln(34/41) + 7\ln(7/41)] = 66.5741$$

which is the smallest possible deviance among all possible splits for all five independent variables.

Each split of a node results in a tree which has nodes that are more pure in the dependent variable. The purity of the tree is defined by the sum of deviances,

$$D = \Sigma_j D_j,$$

where j is the set of all nodes on which splits have not yet been made. This set of nodes is called the "leaf nodes." A "terminal node" is a leaf node on which no further splits are made. In growing a tree, the binary partitioning algorithm recursively splits the data in each node until either the node is homogeneous or contains too few observations.[Ref 6]

According to Figure 3-1, there are a total of 29 out of 69 car owners not satisfied with their cars. The tree splits according to whether the turning circle is less than 39.5 or not. Twenty-eight buyers were initially classed as TRUE and forty-one buyers were classed as FALSE. Of the 28 owners classified as satisfied (TRUE). Six are misclassified (false positive errors). Among the 41 owners classified as not satisfied (FALSE) by the tree model, seven were misclassified (false negative errors). These nodes continue to split along optimal threshold levels to minimize the deviance and increase homogeneity.

However, if a tree is allowed to grow until each terminal node contains one case, the tree may be compromised in its ability to predict new data.

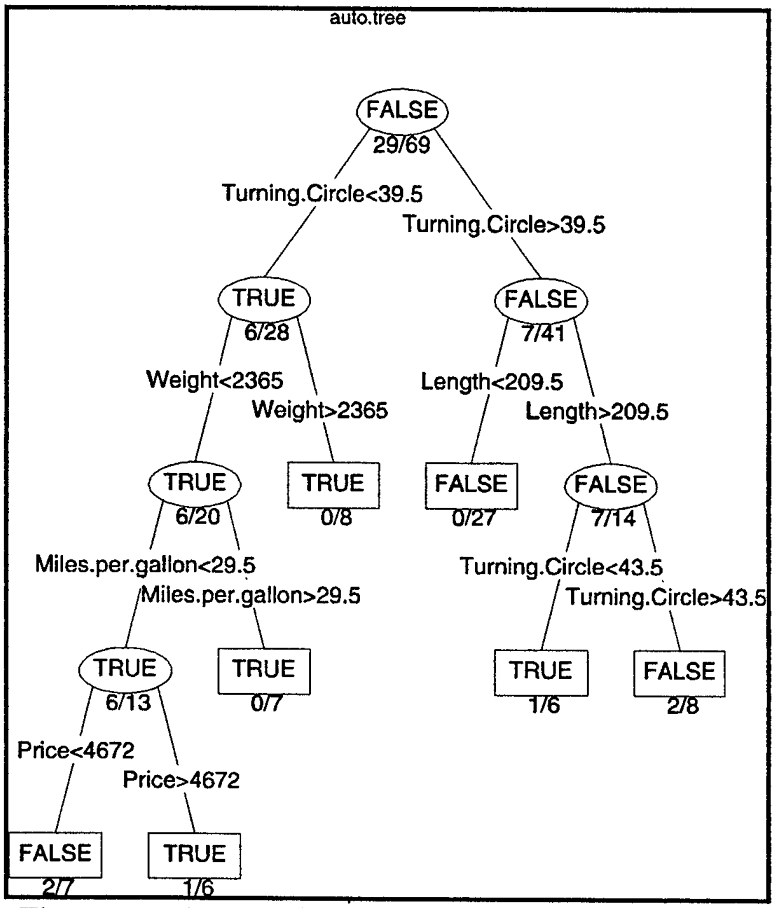


Figure 3-1. An Oversized Tree of Car Owner Satisfaction

The classification tree is described by its tree-object. Figure 3-2 is the tree-object for the owner satisfaction graph. Each node is labeled with a threshold value of the dependent variable which is displayed. The node marks (TRUE, FALSE) characterize the cases. For instance, node 2 which is formed by splitting on the condition Turning.Circle < 39.5, contains 28 cases. The deviance is 29.10. The node has a mark of "TRUE." Twenty-two cases have the value of "TRUE" (28 x 0.7857) and the remaining 6 are "FALSE" (28 x 0.2143).

```
* denotes terminal node
node), split, n, deviance, yval, (yprob)

 1) root 69 93.890 FALSE ( 0.5797 0.4203 )
   2) Turning.Circle<39.5 28 29.100 TRUE ( 0.2143 0.7857 )
     4) Weight<2365 20 24.430 TRUE ( 0.3000 0.7000 )
       8) Miles.per.gallon<29.5 13 17.940 TRUE ( 0.4615 0.5385 )
        16) Price<4672 7  8.376 FALSE ( 0.7143 0.2857 ) *
        17) Price>4672 6  5.407 TRUE ( 0.1667 0.8333 ) *
       9) Miles.per.gallon>29.5 7  0.000 TRUE ( 0.0000 1.0000 ) *
     5) Weight>2365 8  0.000 TRUE ( 0.0000 1.0000 ) *
   3) Turning.Circle>39.5 41 37.480 FALSE ( 0.8293 0.1707 )
     6) Length<209.5 27  0.000 FALSE ( 1.0000 0.0000 ) *
     7) Length>209.5 14 19.410 FALSE ( 0.5000 0.5000 )
      14) Turning.Circle<43.5 6  5.407 TRUE ( 0.1667 0.8333 ) *
      15) Turning.Circle>43.5 8  8.997 FALSE ( 0.7500 0.2500 ) *
```

Figure 3-2. Tree-object of Owner Satisfaction

Since tree size is not limited in the growing process, a tree may be more complex than necessary to describe the data. [Ref. 7] Pruning the tree reduces the original tree structure by removing nodes, at a cost of increasing deviance. Pruning will produce either a single pruned tree if the cost-complexity parameter is given, or a series of pruned trees based on a sequence of cost-complexity parameters.

The pruning method determines the homogeneity (or deviance) of the trees ranging in size from the over-sized tree, to the tree consisting of only the root node. It is intuitive that as the size of the tree increases, the deviance will decrease. Figure 3-3 shows the results from pruning the full tree in the car satisfaction example.

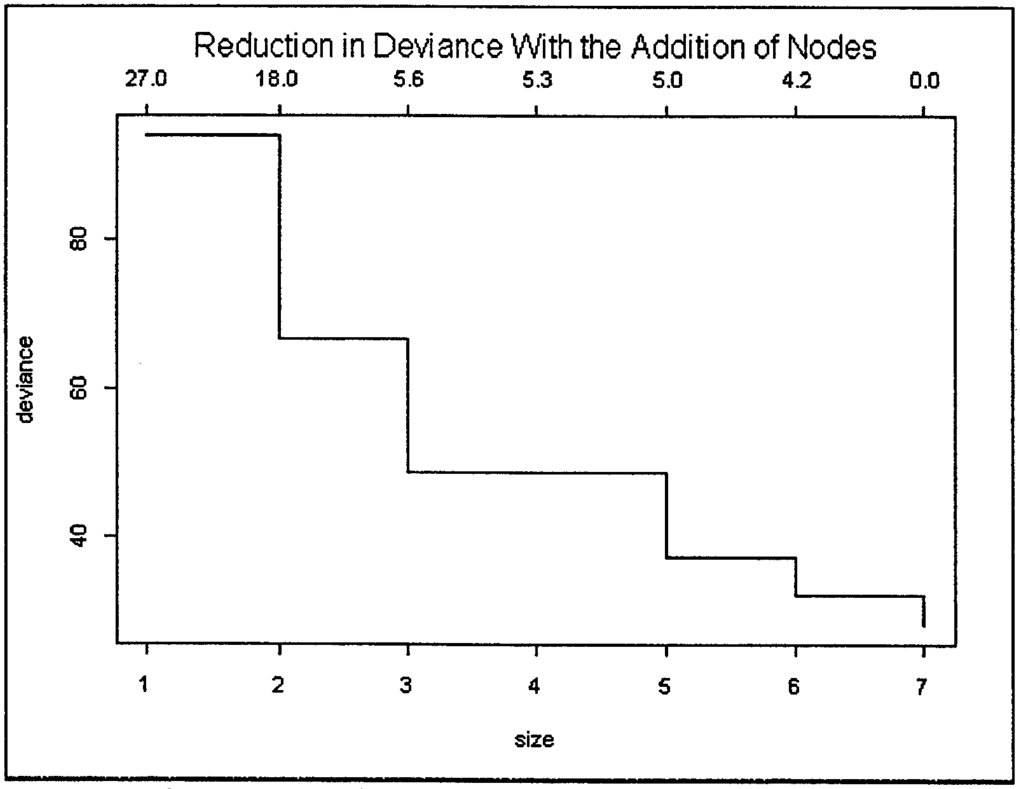


Figure 3-3. Pruning Sequence for Car Satisfaction Example

However, pruning a tree to eliminate complexity is not the only concern. After we have established a tree model, we must ensure that the tree is "right sized." Cross validation is the procedure that is implemented for right sizing.

Achieving total homogeneity is not always reached without cost. The tree may be compromised with its inability to accurately predict responses not used in the tree's construction. Cross-validation is a way of determining the size of tree that optimizes both the purity of the tree and its ability to predict from new data. If the data set is sufficiently large, nine-tenths of it can be used to grow the tree and the remaining data used to check for the tree's ability to accurately classify it.

The process involves the use of nine-tenths of the data to grow an over-sized tree. The tenth of the data removed prior to growing the tree is applied to the sequence of pruned trees to test its predictive accuracy. The deviance from the cases applied to each of the pruned trees in the sequence is recorded.

The procedure is performed nine more times for each of the unique partitions of the data set. When this is finished, there are ten deviances recorded for each size in the sequence of pruned tree. Cross-validation plots the sum of the deviances from all ten trees at each size in the sequence. In general, an increase in tree size will decrease the deviance until the size of the tree is so large that it loses its predictive ability. The minimum point of deviance is the determination of the "right-sized" tree. The series of pruned trees is what the cross validation method uses. [Ref. 10]

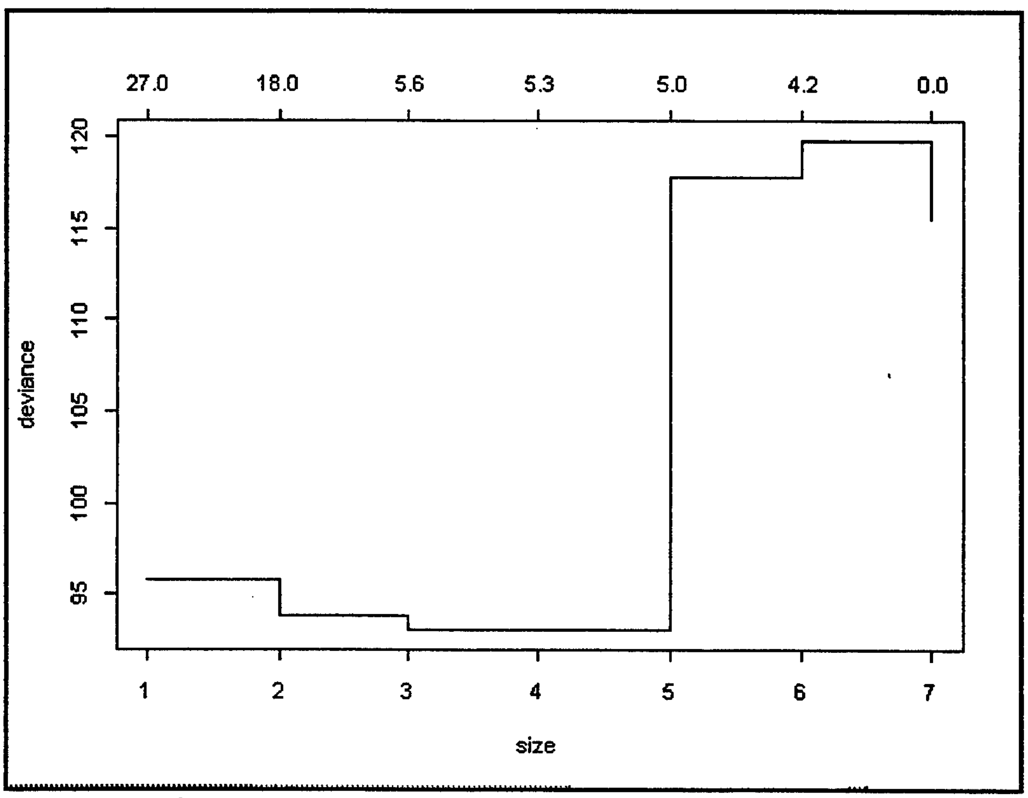


Figure 3-4. Cross-Validation of Car Owner Satisfaction

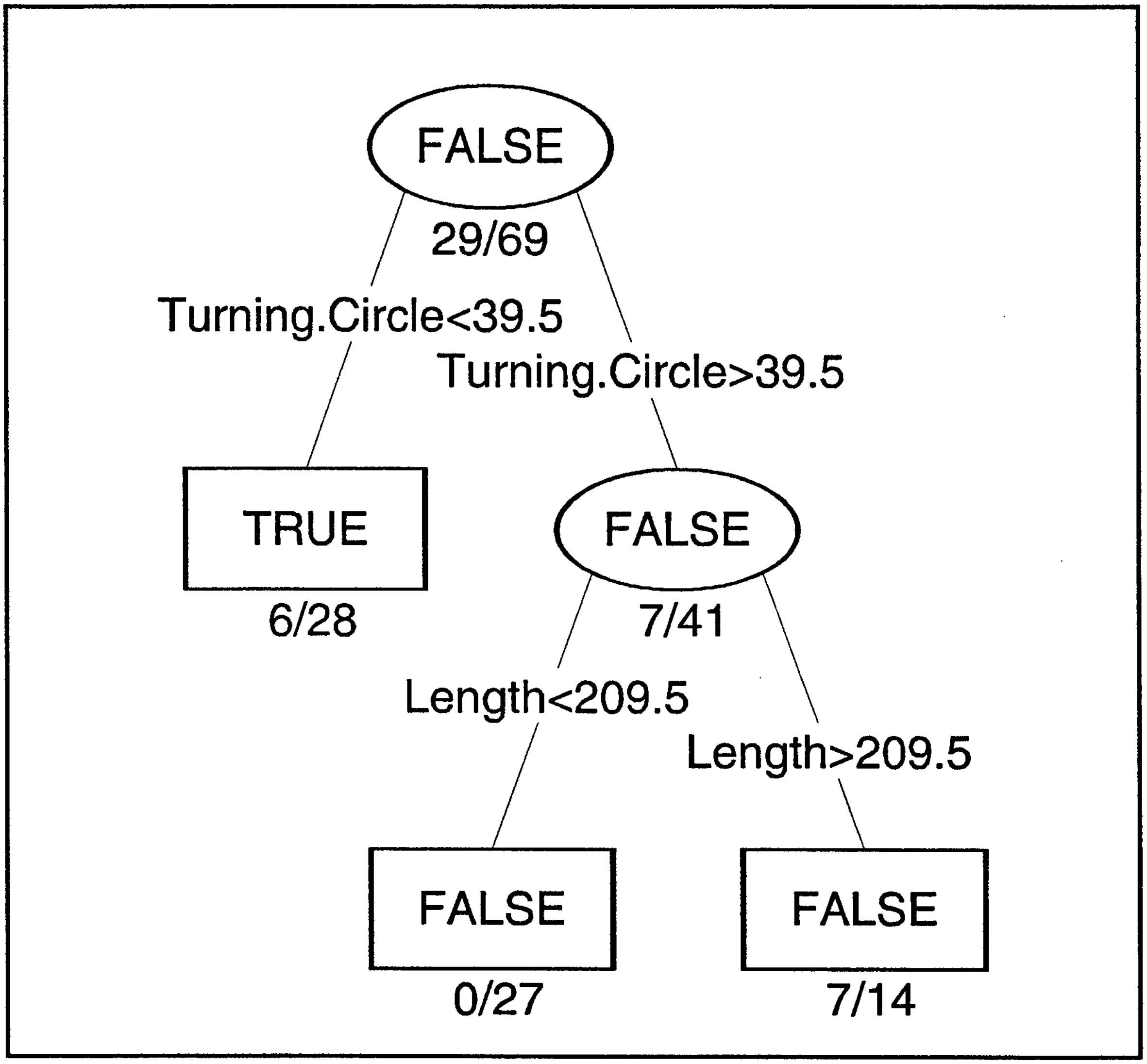


Figure 3-5. Cross-Validated Car Owner Satisfaction Tree Model

Figure 3-4 is a cross-validation example of the car owner satisfaction tree model. The original tree model has seven terminal nodes. However, the cross-validation plot reveals that there are only three terminal nodes to optimize this model. As one would probably infer from Figure 3-3, the misclassification rate increases with the reduction in terminal nodes. Figure 3-5 is the tree model chosen by the cross-validation process. The original model terminal nodes indicates six misclassifications. The cross-validated tree model misclassification increases to 13.

The analysis of the thesis focuses on tree models' predictive ability. Therefore all models will be cross-validated. Refer to APPENDIX D for cross-validation plots.

B. HTTF DATA

All data used in this thesis originated from the Naval Air Warfare Center (NAWC), Trenton, NJ Helicopter Transmission Testing Center (HTTF). A total of 618 acquisitions are used that were taken from 1 December 1994 to 3 January 1997. Of these acquisitions, a wide range of seeded faults were analyzed.

This thesis focuses on three of the numerous faults evaluated at HTTF. A small integral race spall in the port main spiral bevel pinion, one edm notch and three edm notches in the intermediate gear box input pinion were selected. The data acquisitions from the sensors close to the support bearings from which the pinions emanate were used. The indicators employed were designed to isolate bearing faults. The small integral race spall may be detected by these bearing indicators. The race spall in the pinion is a common dynamic cause for gearbox removal in the SH-60. The edm notch faults are evaluated by the data acquisitions from the sensors near the intermediate gear box. The indicators used were again designed to isolate bearing faults. The edm notch fault is a machined slit made in the tooth of the pinion. It was designed to propagate a crack in the gear from the weakness in that tooth.

Of all the acquisitions used, the initial 32 recordings were honest baseline acquisitions. That is, these recordings were taken with no known faults in the transmission system. The remaining 586 had one or more known faulty parts in the transmission.

However, these faults are not expected to affect sensors that are not located near the faulty component. Data from three sensors located near the port main bevel pinion timken bearing were utilized for the small integral race spall - 71 fault responses and 547 no-fault responses. Data from two sensors located near the intermediate gear box input pinion preload bearing were used for the one and three edm notch faults - 186 and 36 acquisitions of faulted data, respectively. There were 396 non-faulted data acquisitions to be used as well.

Sensors 19 and 20 are close to the intermediate gear box input pinion. Sensors 1, 2, and 3 are near the port main bevel pinion. The total number of indicators available is the product of the number of sensors located near the bearings. Since there are 28 indicators calculated for each sensor, the edm faults have 56 indicators (28 x 2) that are used for analysis. The small race spall fault has 84 indicators.

The indicators used in the analysis are suffixed with the sensor number to prevent redundancy (i.e., the pre-load bearing indicators are bdf.19, Iraw_pk2.19, Iraw_cf.19, …, BC6.19, bdf.20, Iraw_pk2.20, Iraw_cf.20,…, BC6.20). As mentioned earlier, all algorithms are proprietary and are not in the scope of this thesis. APPENDIX A provides a sample of data used.

Two different approaches were used for the edm faults. First, all data was grouped together to determine if a tree model could distinguish the difference in non-faulted, single edm notch, and 3 edm notch readings. Next both edm notch faults data acquisitions were grouped together and classified as one fault. The small race spall fault was studied as an isolated case, independent of the edm faults.

IV. RESULTS

The objective of the NAWC, Trenton, New Jersey helicopter transmission test facility is to accurately and efficiently determine the presence of a mechanical fault in a helicopter transmission system. NAWC's methodology is based upon mechanical signature recognition. Large amounts of raw data are stored and processed during each data acquisition at the HTTF. Actual flights would create a significantly large amount of data to be processed if one attempted to continuously monitor all sensor output. Identifying faults within the helicopter transmission is the ultimate goal. In particular, identifying faults within the system with only essential data would be best. Memory and cpu time can be reduced proportionally.

A. PRINCIPAL COMPONENT ANALYSIS

The processing time based upon the twenty-eight independently calculated algorithms(indicators) per sensor used to identify faults, can be quite significant. If faults can be categorized and tailored to each component, the health of the entire transmission system can be systematically evaluated for a spectrum of possible problems. It is plausible that 2-3 indicators can serve as the primary constituents necessary in identifying a fault. These indicators would be unique to the particular fault and component. However, safeguards must be taken to ensure misclassification errors are not increased when the data reduction is used.

By conducting principal component analysis on each fault, the loading factors can be studied to determine if original variables can be isolated as to their importance. If so, correlation between the original variables and classification tree splits would be an indication that the number of indicators used to predict a mechanical fault could be reduced. Additionally, two of the faults chosen are on one component. There is the possibility that common indicators for a particular component could identify several faults on or near the component.

Looking at the data sets, there are only two of concern. The combination of data sets for the edm faults are the same. They only differ in the factor response.

The screeplots, Figures 4-1 and 4-2, reveal the cumulative variances of the first ten principal components. Over 80% of the variability are taken into account in each data set. From these ten principal components, we can determine if specific original variables can be isolated for identifying a fault using a classification tree model. By the cumulative variances displayed in the screeplots, very little optimism can be gained regarding data reduction. The largest principal component encompasses less than one-third of the variability. At best, one of the original variables may be isolated through the

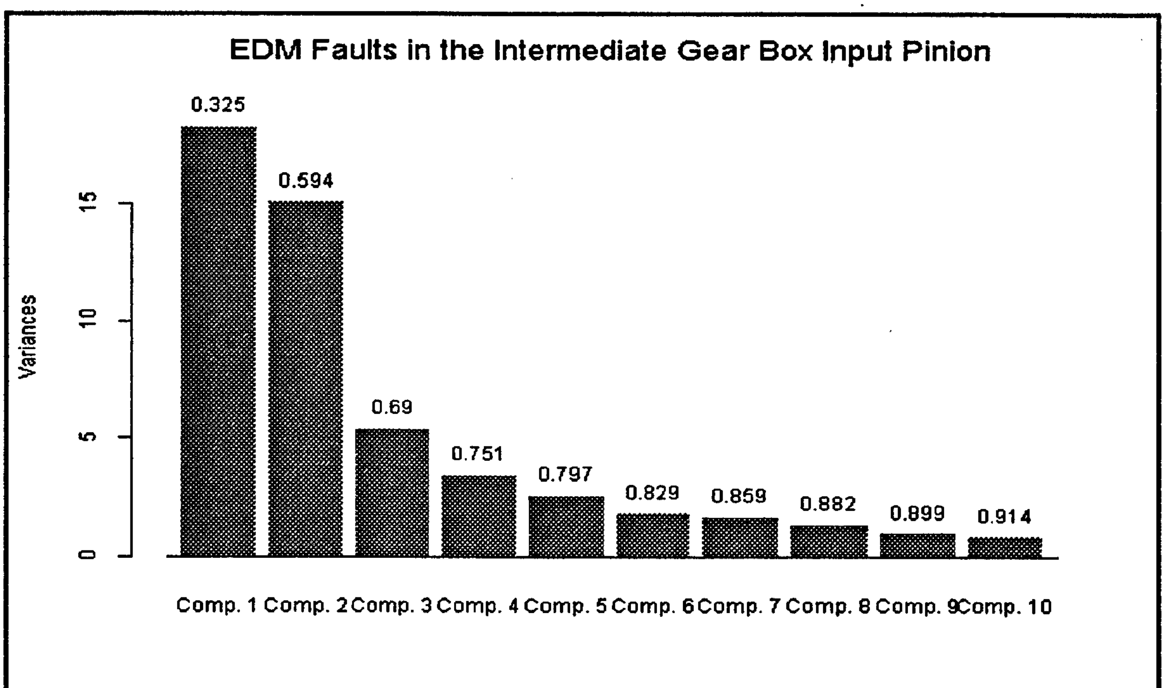


Figure 4-1

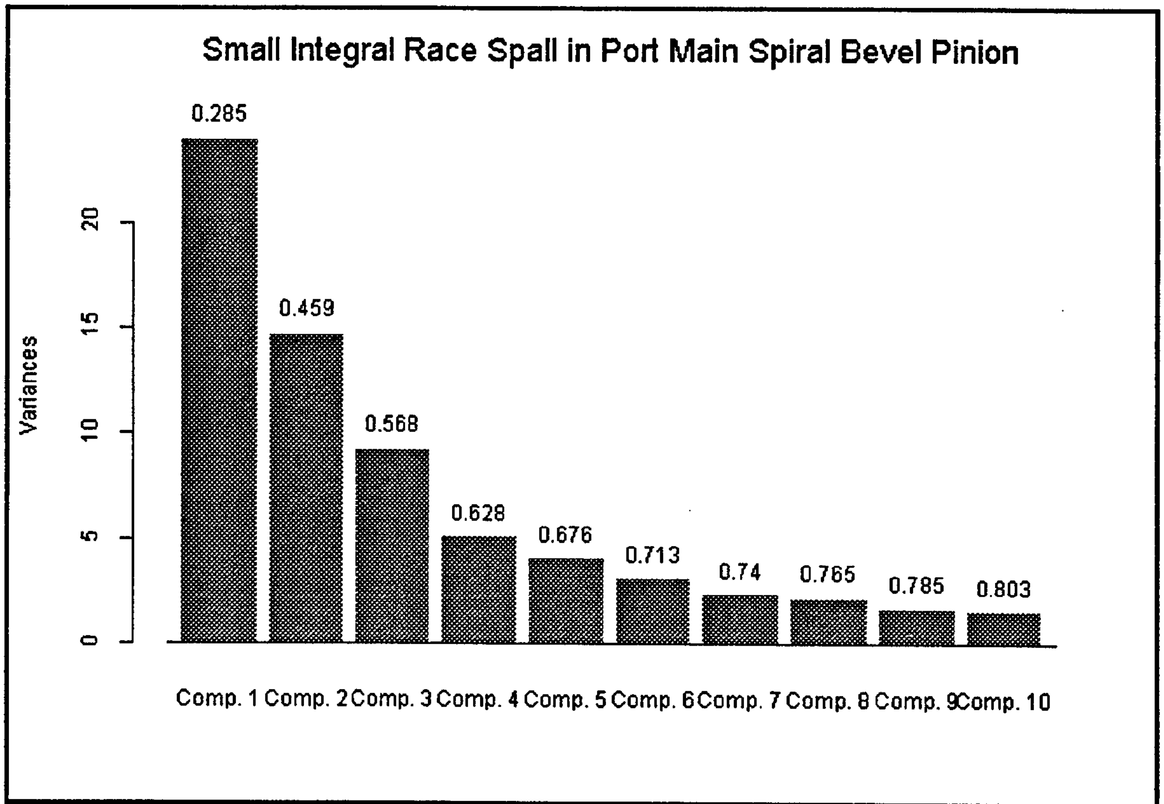


Figure 4-2

principal component method.

The factor loadings, provided in Appendix B, validates the theory derived from viewing the screenplots. They do not support this data reduction. The matrices reveal small magnitudes of the factor loadings. Variables (indicators) with loading factors close to 1.0 (i.e. 0.8 - 1.0) in magnitude would explain most of the variance among all variables observed. Most loading factors had values between 0.01 and 0.2. Based upon these results, it was decided to evaluate all of the original indicators in the pertinent data.

B. CLASSIFICATION TREES

The data structure presented to the tree-model is applicable to the objective of the model. The main goal for HUMS is to indicate whether or not a fault is present in a helicopter's transmission. The analysis conducted not only focuses on fault finding, it evaluates fault classifications as well. Data sets from two different faults are grouped together to determine if the tree-model can distinguish the presence and type of fault.

Figure 4-3 is a tree-model that classifies whether single edm notch, 3 edm notches or no-fault is present in the input pinion. The tree model is able distinguish the two faults and non-fault readings with 19 misclassifications, a 3% error rate. When both edm notch faults are grouped together as a single fault, the error rate increased to 5%. Figure 4-4 shows that there are thirty-two misclassifications. Figure 4-5 shows that the tree model of the timken bearing indicators are able to classify the small race spall with 19 misclassifications. All tree-models are cross validated for prediction accuracy. Cross validation plots are given in APPENDIX B.

The tree models' prediction capabilities cannot sell CART as a "breakthrough" methodology, particularly for integrated health detection on a helicopter. However, for such small degradations in the pinions and the fact that a bearing component indicators were used in the model with a 5% or less error rate, the tree-model worked well. Not shown in the analysis, the model worked exceptionally well in distinguishing the 3 edm notches in the pinion when evaluated alone. The addition of the single edm fault data set created problems for the tree model, resulting in an increased misclassification rate. The difference in the severity of damage to the pinion may be attributed to the model's increased error rate.

Appendix B provides more details in the summaries of each tree object.

Tree Plot of One and Three EDM Notch Faults

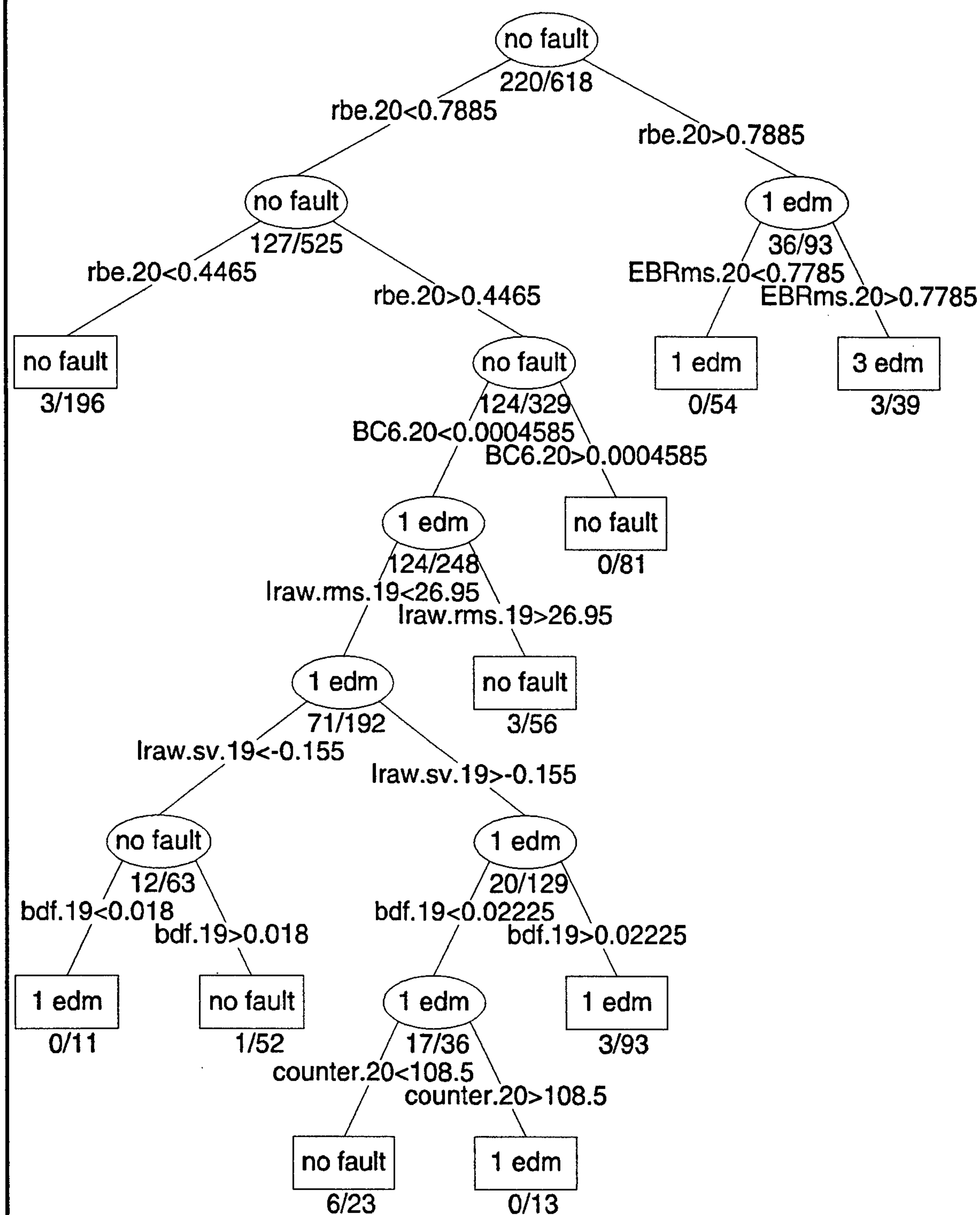


Figure 4-3

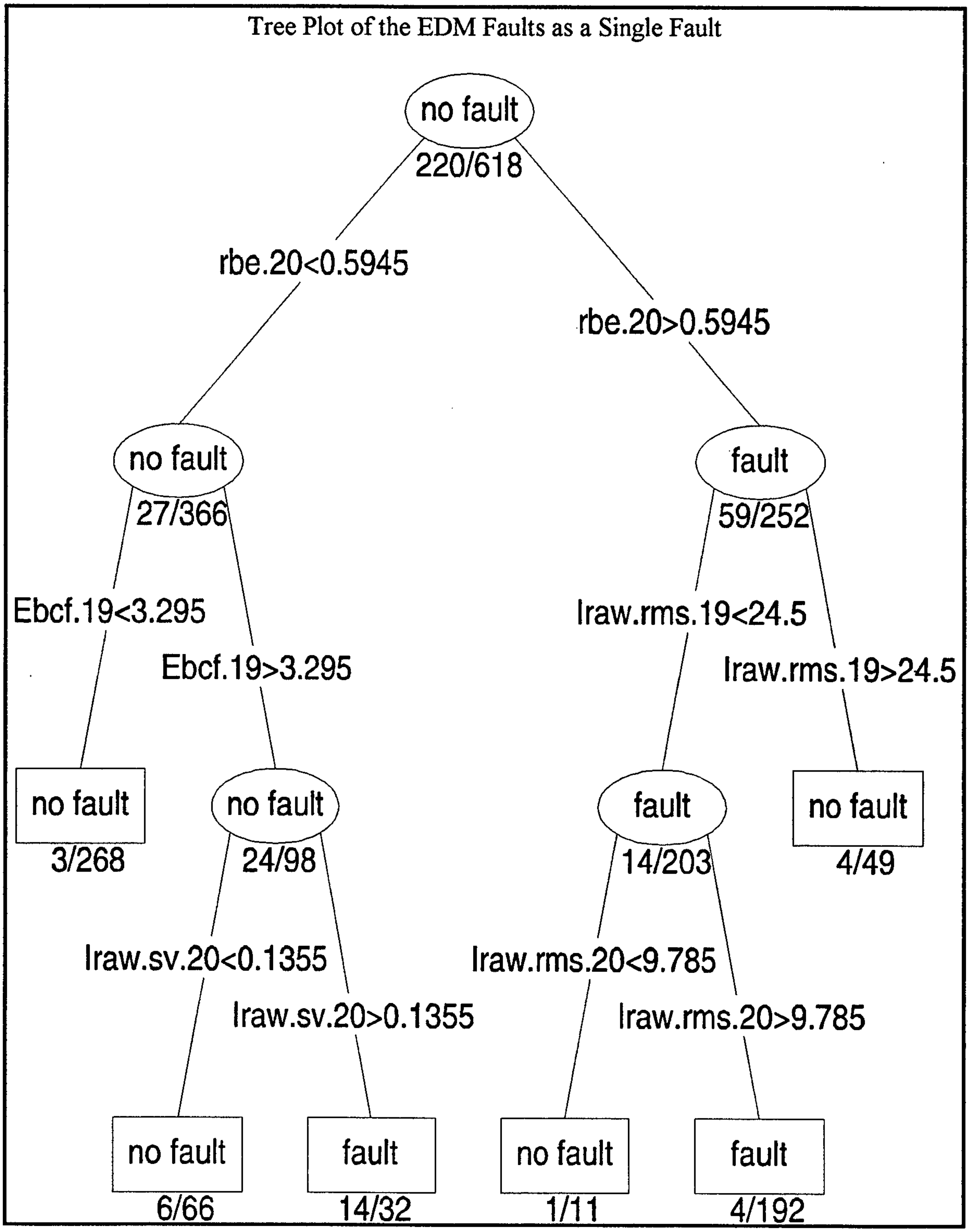

Tree Plot of the EDM Faults as a Single Fault
no fault
220/618
rbe.20<0.5945
rbe.20>0.5945
no fault
27/366
fault
59/252
Ebcf.19<3.295
Ebcf.19>3.295
Iraw.rms.19<24.5
Iraw.rms.19>24.5
no fault
3/268
no fault
24/98
fault
14/203
no fault
4/49
Iraw.sv.20<0.1355
Iraw.sv.20>0.1355
Iraw.rms.20<9.785
Iraw.rms.20>9.785
no fault
6/66
fault
14/32
no fault
1/11
fault
4/192

Figure 4-4

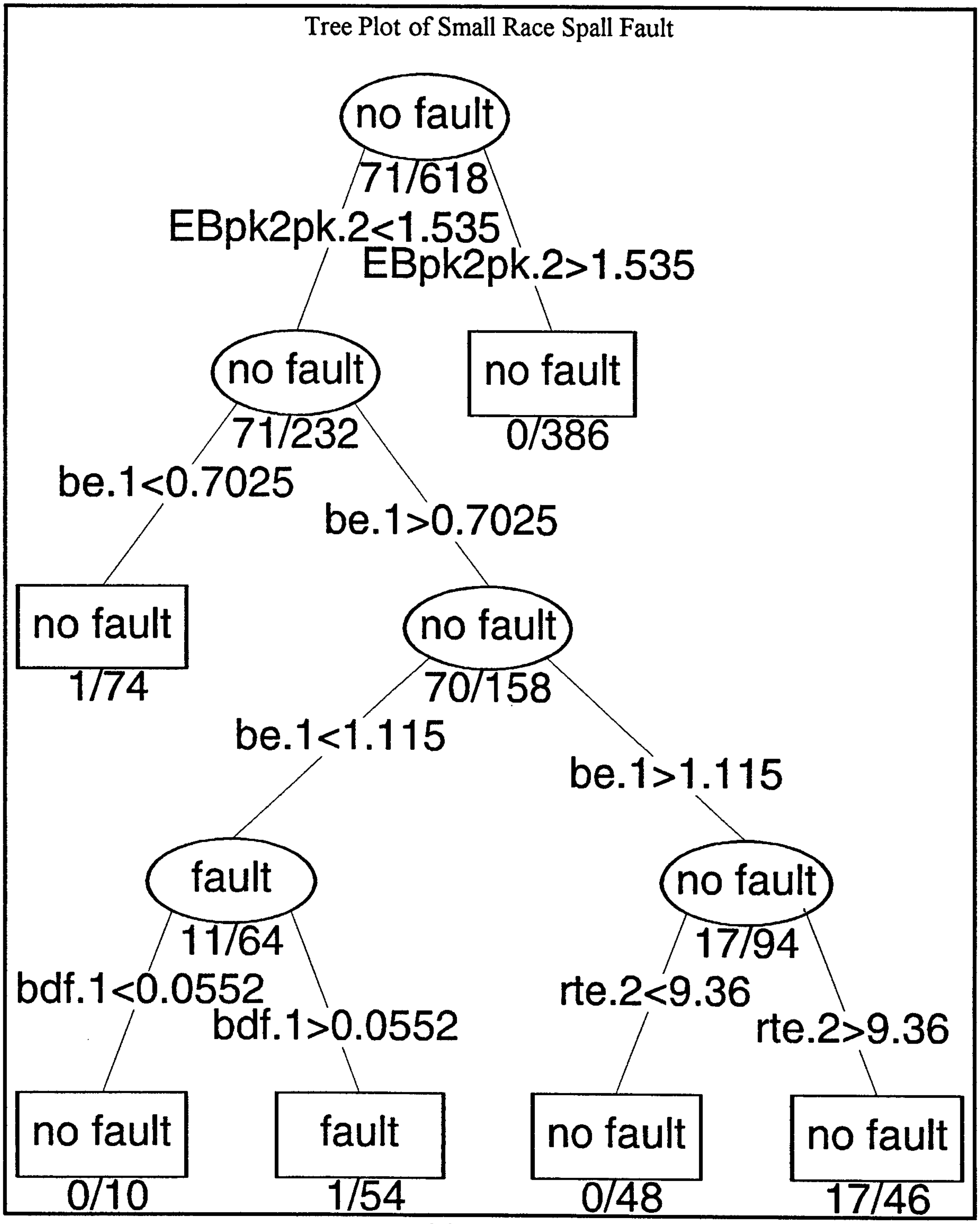

Tree Plot of Small Race Spall Fault
no fault
71/618
EBpk2pk.2<1.535
EBpk2pk.2>1.535
no fault
71/232
no fault
0/386
be.1<0.7025
be.1>0.7025
no fault
1/74
no fault
70/158
be.1<1.115
be.1>1.115
fault
11/64
no fault
17/94
bdf.1<0.0552
bdf.1>0.0552
rte.2<9.36
rte.2>9.36
no fault
0/10
fault
1/54
no fault
0/48
no fault
17/46

Figure 4-5

The errors pertaining to the tree classification models fall in one of two categories, false-negative or false-positive. Tables 4-1 through 4-3 summarizes how well the models classified the actual data. Table 4-1 describes the error rate among the three possible classifications. As mentioned earlier, there are 398 non-fault, 184 single edm notch fault, and 36 three edm notch fault acquisitions were used in the tree model. Reading down the table, one can see 3 single edm notch faults were classified as 3 edm notch faults. Three non-faults were classified as a single edm notch fault, false-positive errors. Also, the model predicted 13 single edm notch faults to be non-faults, a false-negative errors. The sums of the numbers in the table across are the actual number of data acquisitions used by the model.

Tables 4-2 and 4-3 compare actual data outcome against the tree models' predictions, as well. These tables identify only two possible classification errors, false-positive and false-negative. For example, Table 4-2 indicates that 14 out of 220 faulted acquisitions were classified incorrectly as a false-negative error with 18 false-positive misclassification among the 398 non-fault acquisitions.

Classification of Tree Model versus the Actual Outcome	**Tree Model** 3 EDM Notch	1 EDM notch	No-fault
Actual 3 EDM Notch	36	0	0
1 EDM Notch	3	168	13
No -Fault	0	3	395

Table 4-1: Actual versus observed classification of EDM Notch Faults

Classification of Tree Model versus the Actual Outcome		**Tree Model** Fault	No Fault
Actual	Fault	206	14
	No Fault	18	380

Table 4-2: Actual versus tree model classifications of the edm notch faults

Classification of Tree Model versus Actual Outcome		**Tree Model** Fault	No Fault
Actual	Fault	54	17
	No Fault	1	546

Table 4-3: Actual versus observed classification of small race spall faults/no-faults

As mentioned earlier, the edm notch fault is a small machined slit made in the pinion. This fault is difficult to detect. Similarly, the small integral race spall is a less severe fault that is difficult to detect. However, using bearing indicators as the foundation of the tree model, its ability to predict the presence of a fault indicates that tree models could be a viable supplement to the HUMS program. In addition, it reveals that faults in a component can be detected with the indicators for a different component.

V. CONCLUSIONS AND RECOMMENDATIONS

The scope of this thesis is to explore a statistical methodology in conjunction with HUMS data techniques to identify mechanical faults in a helicopter transmission system. Using a tree-based model, the analysis tries to exploit the test data provided by the HTTF at the Naval Air Warfare Center(NAWC) - Trenton, New Jersey. In particular, pinion faults are evaluated with bearing indicators.

The initial stage of the analysis was to reduce the amount of data used for the tree-based model. Data acquisitions from three separate seeded faults were selected and principal component analysis conducted. The loadings were small in magnitude which indicated that variable reduction based on this technique was not effective. The tree-structured model had to consider all bearing indicators.

The tree-structured classification model was able to identify each fault with the highest misclassification error rate of 3%. More notably, the model was able to classify minor pinion faults using bearing indicators. The single edm notch and small integral spall faults are minor component degradation which are considered difficult to identify. The tree based models were able classify the single edm notch fault among more severe faults. Despite the success of the tree model to predict the single edm and race spall faults, each had 18 false-negatives which may be cause one to scrutinize the model a little more. Still, these are the results of a tree model using bearing indicators to predict minor degradation in a pinion.

Increasing the library of true baseline observations would best serve as a foundation for determining faults. Using data with seeded faults located away from the sensors of interest as baseline data can possibly alter the values of the indicators. Eliminating any possible contamination keeps the decision making process more effective.

The HUMS technology is fairly new, yet it is a promising fault detection system. It should be installed on more operational aircraft. Each individual platform should have its own library of data. Close scrutiny of the indicators along with regular maintenance, including open and inspect, could assist in building effective libraries. If possible, during the major overhaul phase, maintenance facilities could install defected components that are common among fleet helicopters to update their libraries with threshold values (ground turns only).

Classification tree models are an excellent supplement to the HUMS analyses. Even though it is not totally accurate in determining a fault, particularly when it is in an infancy stage, adherence to the tree decision rules could help identify potential problems. Additional data and research would be needed to fully integrate CART with the HUMS, but it is a plausible methodology.

APPENDIX A. [EXAMPLE OF MATLAB MATRIX]

The following is a sample of a re-formatted Matlab matrix; 15 data acquisitions. The status column represents the y-value and the remaining columns are independent variables.

status	bdf.19	Iraw_pk2.19	Iraw_cf.19	Iraw_sv.19	Iraw_kv.19	Iraw_rms.19	EBpk2pk.19	Ebcf.19
no fault	2.95E-02	6.95E+01	3.38E+00	-4.91E-01	2.31E+00	1.34E+01	1.31E+00	3.31E+00
no fault	3.43E-02	6.98E+01	3.13E+00	-3.74E-01	2.22E+00	1.21E+01	1.26E+00	3.03E+00
no fault	3.42E-02	7.08E+01	3.48E+00	-3.32E-01	2.14E+00	1.20E+01	1.23E+00	3.03E+00
no fault	3.45E-02	7.01E+01	3.36E+00	-3.01E-01	2.12E+00	1.21E+01	1.33E+00	3.18E+00
no fault	3.40E-02	7.23E+01	3.51E+00	-2.62E-01	2.14E+00	1.24E+01	1.39E+00	3.31E+00
no fault	3.61E-02	7.19E+01	3.32E+00	-2.17E-01	2.19E+00	1.21E+01	1.44E+00	3.30E+00
no fault	3.75E-02	8.07E+01	3.48E+00	-2.28E-01	2.26E+00	1.20E+01	1.54E+00	3.44E+00
no fault	3.04E-02	9.40E+01	2.89E+00	-3.22E-01	2.53E+00	1.75E+01	1.63E+00	3.06E+00
no fault	4.20E-02	9.27E+01	3.06E+00	3.12E-02	2.68E+00	1.55E+01	1.94E+00	2.99E+00
fault	4.27E-02	1.06E+02	2.43E+00	-1.85E-01	1.93E+00	2.19E+01	3.04E+00	3.25E+00
fault	4.26E-02	1.10E+02	2.37E+00	-1.82E-01	1.86E+00	2.34E+01	2.83E+00	2.85E+00
fault	4.26E-02	1.18E+02	2.54E+00	-1.88E-01	1.84E+00	2.39E+01	2.96E+00	2.91E+00
fault	4.33E-02	1.17E+02	2.47E+00	-1.68E-01	1.84E+00	2.42E+01	3.17E+00	3.03E+00
fault	4.67E-02	1.05E+02	2.75E+00	-1.78E-01	1.90E+00	2.13E+01	2.87E+00	2.90E+00
fault	4.79E-02	1.10E+02	2.77E+00	-1.73E-01	2.07E+00	2.06E+01	2.85E+00	2.89E+00

Ebsv.19	Ebkv.19	EBRms.19	rte.19	rbe.19	te.19	be.19	ce.19	bse.19
6.05E-01	3.18E+00	3.94E-01	1.28E+01	5.62E-01	4.26E-02	3.52E-01	3.20E-03	1.27E-03
5.93E-01	3.13E+00	4.15E-01	1.16E+01	5.45E-01	5.45E-02	3.61E-01	3.82E-03	1.37E-03
6.27E-01	3.21E+00	4.09E-01	1.14E+01	5.61E-01	6.10E-02	3.48E-01	3.51E-03	1.18E-03
5.72E-01	3.07E+00	4.18E-01	1.16E+01	5.64E-01	6.25E-02	3.56E-01	2.88E-03	9.38E-04
5.94E-01	3.14E+00	4.22E-01	1.19E+01	5.38E-01	6.62E-02	3.56E-01	3.52E-03	1.12E-03
6.01E-01	3.15E+00	4.38E-01	1.15E+01	5.73E-01	7.86E-02	3.59E-01	3.92E-03	1.58E-03
6.00E-01	3.16E+00	4.49E-01	1.14E+01	5.79E-01	8.12E-02	3.68E-01	3.03E-03	1.76E-03
5.48E-01	3.02E+00	5.33E-01	1.69E+01	6.24E-01	1.19E-01	4.14E-01	4.04E-03	1.42E-03
5.71E-01	3.08E+00	6.51E-01	1.49E+01	6.33E-01	2.42E-01	4.09E-01	3.85E-03	1.36E-03
4.38E-01	2.91E+00	9.35E-01	2.08E+01	1.07E+00	2.04E-01	7.31E-01	1.64E-02	2.13E-03
3.81E-01	2.88E+00	9.95E-01	2.23E+01	1.05E+00	1.92E-01	8.03E-01	1.39E-02	2.22E-03
4.37E-01	2.97E+00	1.02E+00	2.29E+01	1.05E+00	2.19E-01	8.01E-01	1.37E-02	2.06E-03
3.87E-01	2.89E+00	1.05E+00	2.31E+01	1.04E+00	2.18E-01	8.29E-01	1.75E-02	1.78E-03
3.88E-01	2.84E+00	9.92E-01	2.02E+01	1.09E+00	2.08E-01	7.84E-01	1.26E-02	2.22E-03
4.54E-01	2.92E+00	1.00E+00	1.89E+01	1.14E+00	2.18E-01	7.83E-01	2.66E-02	2.50E-03

ie.19	oe.19	tbe.19	counter. 19	EBRms. 19	BC1.19	BC2.19	BC3.19	BC4.19
1.75E-03	5.07E-03	1.74E-02	4.40E+01	2.01E-02	1.41E-04	1.67E-04	3.19E-04	2.62E-04
1.18E-03	2.58E-03	9.85E-03	9.40E+01	2.02E-02	1.98E-04	2.10E-04	3.83E-04	3.25E-04
2.60E-03	1.15E-03	8.14E-03	2.40E+01	1.17E-02	3.05E-04	1.67E-04	4.13E-04	1.48E-04
2.39E-03	1.68E-03	7.84E-03	2.60E+01	1.13E-02	2.44E-04	1.36E-04	5.51E-04	1.66E-04
9.26E-04	1.30E-03	5.59E-03	7.80E+01	1.83E-02	7.08E-05	1.39E-04	1.12E-04	1.26E-04
1.48E-03	8.11E-04	5.73E-03	7.00E+01	1.95E-02	8.27E-05	1.07E-04	7.34E-05	1.16E-04
6.41E-04	7.25E-04	3.82E-03	9.50E+01	1.55E-02	1.05E-04	1.03E-04	8.60E-05	8.76E-05
1.98E-03	3.08E-03	1.23E-02	2.80E+01	1.45E-02	1.79E-04	1.33E-04	3.78E-04	1.99E-04
2.38E-03	1.21E-03	8.81E-03	4.80E+01	1.36E-02	1.94E-04	2.25E-04	4.30E-04	1.55E-04
2.04E-03	2.06E-03	7.84E-03	4.40E+01	1.15E-02	2.32E-04	2.10E-04	7.44E-04	2.02E-04
1.09E-03	1.21E-03	6.86E-03	4.40E+01	1.79E-02	1.47E-04	1.69E-04	1.96E-04	1.86E-04
6.01E-04	7.73E-04	4.50E-03	7.40E+01	1.46E-02	9.51E-05	1.04E-04	1.12E-04	9.48E-05
8.42E-04	1.09E-03	4.89E-03	4.20E+01	1.71E-02	7.59E-05	6.75E-05	9.45E-05	8.30E-05
1.34E-03	1.77E-03	9.92E-03	2.50E+01	1.73E-02	1.67E-04	1.85E-04	4.00E-04	3.88E-04
1.13E-03	1.35E-03	8.00E-03	3.80E+01	1.56E-02	1.62E-04	2.23E-04	3.38E-04	2.88E-04

BC5.19	BC6.19	bdf.20	Iraw_pk2. 20	Iraw_cf. 20	Iraw_sv. 20	Iraw_kv. 20	Iraw_rms. 20	EBpk2pk. 20
2.46E-04	1.18E-03	3.58E-02	8.37E+01	3.30E+00	-7.78E-01	2.62E+00	1.53E+01	1.46E+00
1.84E-04	3.09E-04	4.21E-02	6.16E+01	3.87E+00	-5.14E-01	2.63E+00	9.27E+00	1.19E+00
1.21E-04	2.82E-04	6.65E-02	5.84E+01	3.65E+00	-2.62E-01	2.67E+00	8.05E+00	1.36E+00
1.40E-04	2.88E-04	6.75E-02	5.68E+01	3.77E+00	-2.67E-01	2.66E+00	8.15E+00	1.35E+00
1.23E-04	2.49E-04	1.36E-02	1.27E+02	3.31E+00	1.90E-01	2.57E+00	1.98E+01	8.50E-01
1.18E-04	2.24E-04	1.30E-02	1.27E+02	3.31E+00	1.58E-01	2.53E+00	2.00E+01	8.39E-01
7.60E-05	6.61E-05	1.48E-02	9.30E+01	3.54E+00	1.67E-01	2.78E+00	1.33E+01	6.39E-01
2.03E-04	8.66E-04	3.99E-02	8.99E+01	3.39E+00	-7.02E-01	2.72E+00	1.50E+01	1.41E+00
1.31E-04	5.14E-04	6.22E-02	5.67E+01	4.41E+00	-3.23E-01	2.96E+00	7.08E+00	1.15E+00
1.85E-04	3.23E-04	5.94E-02	6.14E+01	3.69E+00	-1.73E-01	2.75E+00	8.46E+00	1.40E+00
2.10E-04	2.35E-04	1.91E-02	8.95E+01	3.43E+00	-2.43E-01	2.48E+00	1.41E+01	9.59E-01
1.07E-04	1.88E-04	1.15E-02	1.30E+02	3.07E+00	9.69E-02	2.37E+00	2.13E+01	7.84E-01
7.55E-05	1.31E-04	1.86E-02	9.89E+01	3.99E+00	2.02E-01	2.91E+00	1.26E+01	8.07E-01
3.42E-04	7.57E-04	2.61E-02	8.58E+01	3.09E+00	-7.78E-01	2.46E+00	1.65E+01	1.33E+00
2.55E-04	4.05E-04	3.67E-02	6.21E+01	3.65E+00	-4.58E-01	2.53E+00	9.39E+00	1.12E+00

Ebcf.20	Ebsv.20	Ebkv.20	EBRms. 20	rte.20	rbe.20	te.20	be.20	ce.20
2.69E+00	2.23E-01	2.70E+00	5.47E-01	1.49E+01	3.51E-01	1.88E-01	3.59E-01	5.96E-03
3.06E+00	5.72E-01	3.11E+00	3.91E-01	8.97E+00	2.97E-01	9.27E-02	2.98E-01	4.49E-03
2.55E+00	2.03E-01	2.67E+00	5.35E-01	7.74E+00	3.04E-01	2.25E-01	3.10E-01	4.19E-03
2.45E+00	1.87E-01	2.66E+00	5.51E-01	7.86E+00	2.91E-01	2.33E-01	3.17E-01	4.88E-03
3.17E+00	5.24E-01	3.01E+00	2.68E-01	1.94E+01	3.52E-01	6.14E-02	2.07E-01	2.13E-03
3.24E+00	6.13E-01	3.22E+00	2.60E-01	1.96E+01	3.42E-01	4.85E-02	2.11E-01	2.88E-03
3.25E+00	5.18E-01	2.97E+00	1.97E-01	1.30E+01	3.09E-01	5.06E-02	1.46E-01	2.46E-03
2.37E+00	1.61E-01	2.64E+00	5.99E-01	1.46E+01	3.96E-01	2.43E-01	3.56E-01	1.16E-02
2.61E+00	3.37E-01	2.89E+00	4.41E-01	6.79E+00	2.94E-01	1.54E-01	2.87E-01	6.38E-03
2.80E+00	3.06E-01	2.76E+00	5.02E-01	8.19E+00	2.72E-01	2.12E-01	2.90E-01	6.70E-03
3.56E+00	6.37E-01	3.25E+00	2.70E-01	1.38E+01	3.00E-01	6.82E-02	2.02E-01	4.20E-03
3.20E+00	5.93E-01	3.12E+00	2.46E-01	2.10E+01	3.40E-01	6.13E-02	1.84E-01	2.92E-03
3.44E+00	5.35E-01	3.09E+00	2.35E-01	1.23E+01	3.27E-01	7.37E-02	1.61E-01	6.03E-03
3.09E+00	4.64E-01	2.91E+00	4.30E-01	1.61E+01	4.23E-01	1.07E-01	3.24E-01	5.69E-03
3.25E+00	5.17E-01	2.98E+00	3.44E-01	9.05E+00	3.41E-01	7.00E-02	2.74E-01	3.65E-03

bse.20	ie.20	oe.20	tbe.20	counter. 20	EBRms. 20	BC1.20	BC2.20	BC3.20
7.41E-04	3.74E-04	4.79E-03	1.19E-02	4.30E+01	2.17E-02	8.66E-05	1.17E-04	6.88E-05
7.75E-04	5.56E-04	1.74E-03	7.56E-03	5.40E+01	1.93E-02	9.93E-05	1.09E-04	6.70E-05
4.58E-04	3.58E-04	7.20E-04	5.73E-03	1.30E+01	1.07E-02	1.27E-04	7.98E-05	6.15E-05
6.04E-04	3.63E-04	9.80E-04	6.83E-03	1.60E+01	1.24E-02	8.34E-05	9.75E-05	6.26E-05
6.90E-04	4.40E-04	1.22E-03	4.48E-03	4.00E+01	1.67E-02	9.75E-05	9.93E-05	9.04E-05
4.75E-04	1.03E-03	6.32E-04	5.02E-03	2.90E+01	1.93E-02	1.12E-04	9.85E-05	7.92E-05
3.73E-04	3.78E-04	4.57E-04	3.67E-03	9.10E+01	1.87E-02	9.67E-05	8.98E-05	6.74E-05
6.37E-04	3.93E-04	4.04E-03	1.66E-02	3.20E+01	2.78E-02	8.66E-05	9.40E-05	7.98E-05
6.11E-04	2.84E-04	1.07E-03	8.35E-03	4.20E+01	1.89E-02	8.19E-05	1.08E-04	5.11E-05
3.90E-04	2.88E-04	1.93E-03	9.31E-03	2.80E+01	1.85E-02	6.54E-05	9.56E-05	4.11E-05
7.54E-04	3.61E-04	8.97E-04	6.21E-03	7.50E+01	2.30E-02	9.31E-05	1.40E-04	9.37E-05
4.73E-04	2.71E-04	6.86E-04	4.35E-03	3.80E+01	1.77E-02	8.03E-05	9.26E-05	5.45E-05
4.98E-04	3.06E-04	6.35E-04	7.47E-03	1.46E+02	3.18E-02	1.20E-04	1.16E-04	8.67E-05
7.68E-04	3.35E-04	1.55E-03	8.34E-03	4.40E+01	1.94E-02	1.14E-04	1.59E-04	8.63E-05
8.98E-04	4.76E-04	8.53E-04	5.87E-03	5.80E+01	1.71E-02	1.24E-04	1.20E-04	9.84E-05

BC4.20	BC5.20	BC6.20
1.56E-04	9.47E-05	3.68E-04
9.77E-05	9.96E-05	5.58E-04
1.42E-04	5.24E-05	1.18E-03
7.29E-05	4.55E-05	1.34E-03
1.18E-04	7.99E-05	4.85E-04
9.52E-05	1.08E-04	4.96E-04
5.27E-05	7.63E-05	1.87E-04
1.29E-04	7.49E-05	1.48E-03
1.29E-04	6.95E-05	1.25E-03
1.40E-04	4.27E-05	1.52E-03
1.09E-04	1.40E-04	4.29E-04
6.40E-05	8.90E-05	4.94E-04
9.86E-05	8.79E-05	2.01E-04
1.14E-04	1.41E-04	4.09E-04
1.58E-04	1.08E-04	6.22E-04

APPENDIX B. [LOADING FACTORS]

The following are the values of the first 10 principal components' factor loadings.

Principal Components of EDM Notch Fault

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
bdf.19	0.020420981	-0.098533402	0.167486864	0.194398838	0.067837269
Iraw.pk2.	0.1535681	-0.154744381	-0.078292093	-0.108968632	-0.155552142
Iraw.cf.1	0.072080752	-0.11943282	-0.258799241	0.110927985	-0.115306391
Iraw.sv.1	-0.058976287	0.146970692	0.272139067	-0.135077022	0.107374729
Iraw.kv.1	0.045345469	-0.115607432	-0.296992401	0.194543965	-0.215721549
Iraw.rms.	0.134191088	-0.106734372	0.083527144	-0.24780874	-0.213531995
EBpk2pk	0.087435356	-0.216194632	-0.124235947	0.077197372	0.040255113
Ebcf.19	0.058752611	-0.140205696	-0.317126496	0.104741238	-0.094036507
Ebsv.19	0.056730036	-0.136579016	-0.314025172	0.131659817	-0.138076859
Ebkv.19	0.042456326	-0.109215551	-0.298397972	0.1905901	-0.225442492
EBRms.1	0.09137042	-0.206222187	0.164267157	0.036642184	0.039330451
rte.19	0.133149576	-0.099890586	0.0799396	-0.253341859	-0.222423125
rbe.19	0.11449678	-0.197606953	0.121754696	-0.076911968	0.005820926
te.19	0.082702993	-0.187722779	0.164074077	0.134949673	-0.001267109
be.19	0.091601884	-0.206342651	0.157906323	-0.005726958	0.054892703
ce.19	0.088464188	-0.207440331	0.029228025	-0.010597812	0.101400486
bse.19	0.086114999	-0.192635282	0.125920024	-0.02976791	0.02130233
ie.19	0.075969887	-0.164588795	0.142009842	-0.048750623	0.006767029
oe.19	0.074952735	-0.163280583	0.144063654	-0.070270023	-0.031293715
tbe.19	0.093439641	-0.215773722	0.062374224	-0.022891807	0.080686952
counter.	0.04737473	-0.001002442	-0.158707169	-0.188896036	0.012759631
EBRms.1	0.067046347	-0.063458816	-0.058201925	-0.27916889	-0.131974798
BC1.19	0.081213812	-0.203170077	-0.050285628	-0.025549553	0.191456
BC2.19	0.084019049	-0.208398486	-0.035639523	-0.035070418	0.185873855
BC3.19	0.084656492	-0.20887488	-0.035925439	-0.025214148	0.183071353
BC4.19	0.091759457	-0.21857458	-0.009490157	-0.002118745	0.156390311
BC5.19	0.085683163	-0.211810059	-0.044038369	-0.017737729	0.173875033
BC6.19	0.077691881	-0.182635404	0.112170197	0.061206823	0.080926758

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
bdf.20	0.123759	0.007127	0.161152	0.270411	0.025104
Iraw.pk2.2	0.179025	0.043015	0.040301	-0.20634	-0.18359
Iraw.cf.20	0.03376	-0.0197	-0.04138	-0.17956	0.084555
Iraw.sv.20	0.024596	0.017335	-0.07252	-0.07459	0.04111
Iraw.kv.20	0.036346	0.040841	-0.08867	-0.03022	0.195783
Iraw.rms.2	0.186201	0.052604	0.050552	-0.16067	-0.20269
EBpk2pk.2	0.196449	0.102371	-0.01036	0.028132	0.068465
Ebcf.20	0.042058	0.054897	-0.19367	-0.21239	0.289056
Ebsv.20	0.021678	0.057226	-0.22798	-0.21208	0.306349
Ebkv.20	0.017093	0.061175	-0.20264	-0.13214	0.338794
EBRms.20	0.212333	0.097974	0.030563	0.079619	0.013146
rte.20	0.178953	0.045747	0.055306	-0.17575	-0.222
rbe.20	0.212273	0.104455	-0.00422	0.014532	0.020747
te.20	0.206163	0.081157	0.075909	0.127779	-0.01429
be.20	0.211516	0.101609	0.017038	0.064473	0.021061
ce.20	0.204902	0.112946	-0.0104	0.053554	0.026854
bse.20	0.202136	0.102099	-0.00023	0.050877	0.031992
ie.20	0.196933	0.102357	0.002248	0.0582	0.02029
oe.20	0.206782	0.097862	0.010097	0.040881	0.005994
tbe.20	0.208319	0.111257	-0.00575	0.053641	0.025197
counter.20	0.033472	0.00901	-0.1377	-0.18122	0.09043
EBRms.20	0.060745	0.016311	-0.08449	-0.34502	-0.14833
BC1.20	0.202134	0.117801	-0.03043	0.045447	0.054658
BC2.20	0.200851	0.118425	-0.03708	0.028567	0.073779
BC3.20	0.199802	0.117723	-0.02366	0.063948	0.043961
BC4.20	0.202392	0.112014	-0.00846	0.065249	0.038847
BC5.20	0.202762	0.11694	-0.02619	0.053852	0.048685
BC6.20	0.201447	0.097846	0.020604	0.072952	0.027813

	Comp.6	Comp.7	Comp.8	Comp.9	Comp.10
bdf.20	0.418625	-0.16932	-0.26243	0.060519	0.123621
Iraw.pk2.2	-0.08737	0.01493	0.033769	-0.05173	-0.18887
Iraw.cf.20	0.219643	-0.0627	-0.1301	0.122802	-0.16243
Iraw.sv.20	-0.00989	-0.02539	-0.17887	-0.00447	-0.0238
Iraw.kv.20	0.050842	-0.06551	-0.01624	-0.07961	0.043208
Iraw.rms.2	-0.13031	0.01154	0.106733	-0.14781	-0.19545
EBpk2pk.2	-0.11577	0.022616	-0.0104	0.05835	-0.00329
Ebcf.20	-0.01634	-0.06565	-0.02995	-0.0487	0.055792
Ebsv.20	-0.0029	-0.06101	-0.02292	-0.04612	0.028451
Ebkv.20	0.060113	-0.08069	-0.0117	-0.08676	0.04535
EBRms.20	0.081652	-0.07058	-0.05729	-0.11597	-0.00777
rte.20	-0.13516	0.014346	0.116951	-0.14931	-0.19993
rbe.20	-0.00633	-0.03916	-0.10404	-0.07651	-0.05715
te.20	0.032262	-0.12166	-0.07377	-0.10787	-0.12388
be.20	0.098859	-0.04655	-0.04801	-0.11467	0.041203
ce.20	0.002176	0.053827	0.145153	-0.08251	0.111099
bse.20	0.181204	-0.05316	-0.13138	-0.14195	0.077376
ie.20	0.09613	-0.02873	-0.13308	-0.03025	0.15163
oe.20	0.144822	-0.08	0.030745	0.090694	0.252054
tbe.20	0.04147	0.027618	0.096896	-0.07175	0.136537
counter.20	0.298422	0.299073	-0.00732	0.007082	-0.08481
EBRms.20	0.248229	0.09053	0.317971	0.025425	0.413328
BC1.20	-0.19156	0.104193	-0.03092	0.180189	-0.02788
BC2.20	-0.17601	0.096559	-0.02752	0.172599	-0.01863
BC3.20	-0.17198	0.091052	-0.01502	0.170701	-0.00915
BC4.20	-0.09245	0.057681	0.012065	0.015169	-0.01619
BC5.20	-0.16831	0.088532	-0.00325	0.14945	-0.01688
BC6.20	-0.02406	-0.03999	0.122756	0.098156	-0.10058

	Comp.6	Comp.7	Comp.8	Comp.9	Comp.10
bdf.20	0.268757	0.049483	0.210074	-0.04771	-0.06783
lraw.pk2.2	-0.03151	-0.20565	-0.06921	-0.03634	-0.0645
lraw.cf.20	0.433615	-0.0125	0.312813	0.079804	-0.53359
lraw.sv.20	0.005238	0.352922	-0.47723	-0.49835	-0.01449
lraw.kv.20	-0.02085	0.165999	0.354082	-0.51335	0.181732
lraw.rms.2	-0.06859	-0.17109	-0.12755	0.013715	-0.01519
EBpk2pk.2	-0.01816	-0.12874	0.018326	-0.02528	0.006966
Ebcf.20	0.047892	-0.31535	-0.06493	-0.1433	0.097509
Ebsv.20	0.051249	-0.24876	-0.02616	-0.02845	-0.0169
Ebkv.20	0.025977	-0.38115	-0.0329	-0.07882	-0.06373
EBRms.20	0.006733	-0.00311	-0.00958	0.010787	-0.00519
rte.20	-0.07304	-0.18389	-0.13343	0.017328	-0.01768
rbe.20	-0.01395	-0.01373	-0.04558	-0.01876	0.009277
te.20	0.019477	0.014971	0.009156	-0.07087	-0.03337
be.20	0.003027	-0.00922	-0.01473	0.034155	0.003091
ce.20	-0.03564	0.051946	-0.01358	-0.00133	0.022654
bse.20	0.00633	0.009091	-0.02763	0.104813	0.017543
ie.20	0.00955	0.069157	-0.01312	0.122764	0.02738
oe.20	0.031042	0.018094	0.055127	0.092776	0.083995
tbe.20	-0.01909	0.046258	-0.00864	0.035825	0.029332
counter.20	0.240108	0.383973	-0.30336	0.140691	-0.14389
EBRms.20	0.075646	0.076346	-0.06252	0.279153	0.358869
BC1.20	-0.01978	0.068071	-0.00841	0.016979	0.009048
BC2.20	-0.02335	-0.01088	0.000825	0.002169	0.017942
BC3.20	-0.00928	0.087687	0.038046	0.053339	0.013622
BC4.20	-0.00532	0.071011	0.032725	0.014389	0.015024
BC5.20	-0.02267	0.060505	0.019674	0.014332	0.008902
BC6.20	-0.00346	0.050656	0.040135	-0.09281	-0.00564

Principal Components of Small Race Spall

	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp. 5
bdf.1	-0.038298068	-0.08836711	-0.167262219	0.254264651	-0.003300468
Iraw.pk2.1	0.186992557	0.034180028	0.024698764	-0.106972429	-0.013783478
Iraw.cf.1	-0.138689725	-0.085160064	-0.044729826	0.076268291	0.063599929
Iraw.sv.1	0.012386616	0.066020023	-0.037964903	-0.013851486	-0.066492341
Iraw.kv.1	-0.141985616	-0.083977102	-0.044598705	0.050536558	0.073840286
Iraw.rms.1	0.177704392	0.035537731	0.040339245	-0.097742413	-0.024184782
EBpk2pk.1	0.167429947	0.005618486	-0.16380259	0.062014592	-0.081745761
Ebcf.1	-0.015860625	-0.089838409	-0.033949253	-0.084858643	-0.091408636
Ebsv.1	-0.012121198	-0.129863851	-0.080900028	-0.140642294	-0.149814132
Ebkv.1	-0.013550034	-0.126781252	-0.077919938	-0.140423768	-0.151993708
EBRms.1	0.170527426	0.018614076	-0.159160363	0.073354145	-0.06766281
rte.1	0.175782874	0.035189707	0.043849344	-0.100661135	-0.023611894
rbe.1	0.158504748	0.030003907	-0.104119898	0.056576626	-0.032916716
te.1	0.150211544	0.031699894	-0.16716469	0.015663408	-0.126027737
be.1	0.171873483	0.010604984	-0.146199216	0.099560176	-0.033321188
ce.1	0.154134763	0.063601042	-0.155705198	0.052422498	-0.072263491
bse.1	0.148947551	-0.017370015	-0.112449539	0.139209646	0.040289455
ie.1	0.154600875	0.013956998	-0.107640815	0.102558295	-0.007656424
oe.1	0.148657716	0.01528467	-0.125924109	0.114654105	0.017969066
tbe.1	0.163953917	0.050758869	-0.155584107	0.075499332	-0.051301899
counter.1	0.020152288	0.168403119	0.015648539	-0.013814886	-0.129831208
EBRms.11	0.052656754	0.102660933	-0.063943709	0.097477388	0.023639051
BC1.1	0.161555227	0.052401165	-0.153012956	0.053642294	-0.061901926
BC2.1	0.15955436	0.044978168	-0.154518041	0.040223533	-0.073025499
BC3.1	0.147772495	0.022345704	-0.158482034	0.054618651	-0.055181373
BC4.1	0.155618333	0.044717415	-0.155633037	0.056596897	-0.076274234
BC5.1	0.158025897	0.045486582	-0.155572956	0.071690006	-0.048683167
BC6.1	0.162545671	0.045734669	-0.150145194	0.059249524	-0.062550368

	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp. 5
bdf.2	-0.0533	-0.12422	0.018736	0.260581	0.018484
Iraw.pk2.2	0.143616	-0.12903	0.074439	-0.04402	0.02594
Iraw.cf.2	-0.07694	0.030845	-0.03829	0.27505	0.066646
Iraw.sv.2	-0.02784	0.051802	0.018931	-0.0776	-0.10842
Iraw.kv.2	-0.05299	0.0587	-0.01055	0.288431	0.074282
Iraw.rms.2	0.153364	-0.10233	0.069	-0.15999	0.009812
EBpk2pk.2	0.076852	-0.21608	0.057841	0.091662	0.067816
Ebcf.2	-0.00127	0.05605	-0.0679	0.065268	0.27741
Ebsv.2	0.016013	0.016249	-0.06054	0.10576	0.374405
Ebkv.2	0.020042	0.029032	-0.06938	0.094664	0.36665
EBRms.2	0.072458	-0.22699	0.077739	0.058756	-0.02938
rte.2	0.152876	-0.10263	0.068544	-0.15984	0.009224
rbe.2	0.159913	-0.08997	0.079096	-0.15462	0.019885
te.2	0.051441	-0.18378	0.083119	0.082647	0.029752
be.2	0.074399	-0.22697	0.072054	0.049103	-0.04415
ce.2	0.02869	-0.20454	0.082217	0.028321	-0.19195
bse.2	0.074434	-0.19494	0.06402	0.030644	-0.03244
ie.2	0.09381	-0.15356	0.055388	0.044395	0.049567
oe.2	0.076212	-0.17959	0.087986	0.083192	-0.00796
tbe.2	0.045917	-0.21727	0.086623	0.037517	-0.15712
counter.2	-0.00482	0.072952	0.050619	-0.05194	-0.20622
EBRms.22	-0.02359	-0.09778	0.053851	-0.02077	-0.28203
BC1.2	0.054466	-0.19367	0.083508	0.1484	0.029447
BC2.2	0.060123	-0.18857	0.096598	0.088728	0.028656
BC3.2	0.072017	-0.15281	0.076437	0.100197	0.118999
BC3.23	0.072759	-0.18817	0.067487	0.071415	0.03459
BC5.2	0.060603	-0.1978	0.077469	0.101421	-0.02585
BC6.2	0.035619	-0.21379	0.032775	0.035242	-0.06993

	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp. 5
bdf.3	-0.11203	0.110932	0.04827	0.161064	-0.17147
Iraw.pk2.3	0.170414	0.027192	0.029229	-0.15588	0.094439
Iraw.cf.3	-0.12251	0.042293	-0.04487	-0.03123	-0.16397
Iraw.sv.3	0.041424	-0.04461	-0.09315	-0.11662	-0.0089
Iraw.kv.3	-0.08294	0.037082	-0.06288	-0.09237	-0.1611
Iraw.rms.3	0.174937	0.019922	0.02769	-0.14199	0.10386
EBpk2pk.3	0.094077	0.127232	0.204377	0.069775	-0.01859
Ebcf.3	-0.04888	0.017031	0.023016	-0.0242	0.008092
Ebsv.3	-0.08533	0.025765	-0.00655	-0.08224	0.014102
Ebkv.3	-0.07155	0.027552	-0.00029	-0.09217	0.015788
EBRms.3	0.113554	0.12386	0.200674	0.076441	-0.02064
rte.3	0.174652	0.019502	0.027138	-0.14339	0.103965
rbe.3	0.170498	0.037997	0.038114	-0.06971	0.087096
te.3	0.09287	0.088063	0.11888	0.148247	-0.1362
be.3	0.105511	0.120419	0.204668	0.036052	0.028765
ce.3	0.015425	0.109847	0.145322	0.191667	-0.17484
bse.3	0.054021	0.075739	0.168446	0.064291	0.059869
ie.3	0.045336	0.095603	0.168876	0.034332	0.041366
oe.3	0.080721	0.104779	0.124074	0.026021	-0.00858
tbe.3	0.0381	0.13003	0.184832	0.179137	-0.13522
counter.3	-0.0808	-0.03429	-0.0466	0.159484	-0.14947
EBRms.34	-0.07692	0.049861	0.03388	0.161337	-0.17385
BC1.3	0.117238	0.068402	0.169063	-0.04332	0.018006
BC2.3	0.102969	0.100677	0.16433	-0.00895	0.031576
BC3.3	0.094208	0.112063	0.16802	0.043627	0.047458
BC4.3	0.098432	0.105055	0.152161	0.044864	-0.00375
BC5.3	0.04959	0.112863	0.198325	0.089424	-0.06013
BC6.3	0.04311	0.113078	0.192229	0.099781	-0.04352

	Comp. 6	Comp. 7	Comp. 8	Comp. 9	Comp. 10
bdf.1	0.080764	0.03078	-0.11871	0.060362	-0.07549
Iraw.pk2.1	-0.03458	-0.01151	0.030376	-0.09732	0.102096
Iraw.cf.1	0.095263	0.052634	-0.03061	0.278386	-0.11594
Iraw.sv.1	-0.05905	0.028915	-0.1027	0.146007	-0.00745
Iraw.kv.1	0.082596	0.054964	-0.02926	0.312438	-0.10493
Iraw.rms.1	-0.04117	-0.05582	0.036847	-0.19261	0.140549
EBpk2pk.1	0.047131	0.071725	-0.00961	-0.00393	0.027345
Ebcf.1	0.030204	0.325038	0.009884	-0.0534	0.126275
Ebsv.1	0.04717	0.401424	-0.02733	-0.08447	0.104713
Ebkv.1	0.049278	0.404307	-0.02555	-0.09456	0.114471
EBRms.1	0.042063	0.024485	-0.00812	0.00373	0.010578
rte.1	-0.0423	-0.06068	0.03922	-0.20063	0.144418
rbe.1	0.017045	0.147706	-0.06435	0.195914	-0.0653
te.1	0.06796	0.017447	0.036619	-0.06966	0.030033
be.1	0.026185	0.026921	-0.03112	0.041843	-0.00028
ce.1	0.046983	-0.07664	0.051447	-0.01608	0.001623
bse.1	0.026893	-0.01619	-0.07182	0.073264	-0.12327
ie.1	0.028566	0.006262	-0.03508	0.076058	-0.04597
oe.1	0.045983	-0.0265	-0.0656	0.07875	-0.04948
tbe.1	0.046731	-0.06457	0.023615	0.008784	-0.02071
counter.1	0.051778	-0.25032	0.121397	-0.17018	-0.11809
EBRms.11	0.036617	-0.28124	0.075549	0.047767	-0.17696
BC1.1	0.046296	-0.02396	0.034971	-0.02831	0.018812
BC2.1	0.058553	-0.00991	0.017113	-0.04235	0.001915
BC3.1	0.060982	0.010107	0.012863	-0.02637	-0.01057
BC4.1	0.040955	0.025836	0.016693	-0.04117	0.005137
BC5.1	0.042832	-0.03441	-0.00561	-0.00823	0.002548
BC6.1	0.050856	-0.00946	0.03325	-0.00257	0.002506

	Comp. 6	Comp. 7	Comp. 8	Comp. 9	Comp. 10
bdf.2	0.135418	-0.05531	-0.16407	0.06143	0.244812
lraw.pk2.2	-0.07477	0.046024	0.184571	-0.01907	0.020037
lraw.cf.2	0.003096	0.051659	0.095996	-0.06926	0.400211
lraw.sv.2	0.257796	-0.0099	-0.10469	-0.11878	-0.27334
lraw.kv.2	0.006423	0.027977	0.10547	-0.04827	0.400522
lraw.rms.2	-0.0555	0.012853	0.104479	0.043848	-0.10009
EBpk2pk.2	0.03124	0.04442	0.069017	-0.01811	-0.04512
Ebcf.2	-0.04368	0.190643	0.237403	-0.17151	-0.19538
Ebsv.2	-0.00352	0.187163	0.182868	-0.15414	-0.17599
Ebkv.2	-0.0188	0.201058	0.204804	-0.16051	-0.14882
EBRms.2	0.042603	-0.01931	-0.00402	0.017109	0.010164
rte.2	-0.05655	0.012879	0.106097	0.044337	-0.10086
rbe.2	-0.02518	0.007655	0.05968	0.036761	-0.07735
te.2	0.012234	-0.00645	0.053782	0.027259	0.077362
be.2	0.048698	-0.02174	-0.01978	0.01347	-0.00894
ce.2	0.005318	-0.06821	0.097086	0.002712	-0.05366
bse.2	0.050449	-0.07328	-0.02871	-0.03086	0.01652
ie.2	0.053813	-0.01936	-0.01276	-0.08505	-0.00163
oe.2	0.041215	0.018837	-0.02385	-0.02177	0.048439
tbe.2	0.018024	-0.06415	0.071683	-0.00935	-0.0366
counter.2	-0.13034	-0.03291	0.0863	0.029758	0.133708
EBRms.22	-0.05269	-0.11889	0.181911	-0.04912	-0.09446
BC1.2	0.045828	-0.0151	0.029379	0.015517	-0.00984
BC2.2	0.054618	-0.02588	-0.03817	-0.04665	-0.00759
BC3.2	0.057873	-0.06472	-0.03565	-0.13552	-0.06764
BC3.23	0.063393	-0.02111	0.024835	-0.06163	-0.04166
BC5.2	0.040531	-0.01872	0.037036	-0.04668	-0.07788
BC6.2	0.064092	0.01979	-0.00071	-0.02487	-0.11507

	Comp. 6	Comp. 7	Comp. 8	Comp. 9	Comp. 10
bdf.3	0.093313	0.064664	-0.07202	-0.12289	-0.15267
lraw.pk2.3	-0.0167	0.037368	0.065911	0.16361	0.04897
lraw.cf.3	0.245894	0.113822	0.057224	-0.16243	-0.08453
lraw.sv.3	0.005587	0.166233	-0.21175	0.289615	0.005817
lraw.kv.3	0.242557	0.200396	0.060403	-0.11495	-0.16574
lraw.rms.3	-0.06325	-7.7E-05	0.056495	0.160488	0.06482
EBpk2pk.3	0.124241	0.068541	0.054651	0.093315	0.016467
Ebcf.3	0.278418	-0.04607	0.258212	0.213934	0.060218
Ebsv.3	0.372371	-0.05738	0.288261	0.157253	0.10041
Ebkv.3	0.37836	-0.06223	0.307918	0.154684	0.130915
EBRms.3	0.038226	0.076939	-0.01942	0.02529	0.005934
rte.3	-0.06424	-0.00242	0.058031	0.15804	0.066834
rbe.3	-0.01527	0.094047	-0.00682	0.2457	-0.02823
te.3	-0.17485	0.124154	0.162965	0.108151	0.001897
be.3	0.119125	0.046291	-0.0899	-0.01207	0.006761
ce.3	-0.13023	0.139295	0.231595	0.11259	-0.07384
bse.3	0.078825	0.025123	-0.13336	0.015665	-0.11974
ie.3	0.114796	0.061278	-0.21138	-0.00769	-0.07358
oe.3	0.108883	0.014933	-0.15907	0.01425	-0.0369
tbe.3	-0.06939	0.130192	0.128989	0.098739	-0.09404
counter.3	-0.31353	0.075939	-0.02245	0.05986	-0.10118
EBRms.34	-0.18809	0.068761	0.255063	0.094626	-0.12271
BC1.3	0.055023	0.021787	-0.02144	-0.02497	0.024787
BC2.3	0.075841	0.049707	-0.08685	-0.04224	0.012148
BC3.3	0.090133	0.055982	-0.12181	-0.06802	0.042063
BC4.3	0.105136	0.053292	-0.07653	-0.09102	0.033817
BC5.3	0.036898	0.096034	-0.01966	0.040732	-0.00443
BC6.3	0.052207	0.093604	-0.03106	0.011433	0.024887

APPENDIX C. [TREE CLASSIFICATION SUMMARIES]

This appendix contains the S-Plus output for each tree model constructed. It contains the details of the tree. Each line of the tree has the node, the numeric split that separated the cases, the deviance at that node, the y-value of the node, and a vector with the probabilities of each case in the node. An asterisk denotes a terminal node. Each tree object corresponds to a figure in the text of the thesis - ordered in the same sequence that tree graphs appear.

One and Three EDM Notch Faults - Tree Object

```
node), split, n, deviance, yval, (yprob)
      * denotes terminal node

  1) root 618 1001.000 no fault ( 0.29770 0.05825 0.64400 )
    2) rbe.20<0.7885 525  580.900 no fault ( 0.24190 0.00000 0.75810 )
       4) rbe.20<0.4465 196  31.030 no fault (0.01531 0.00000 0.98470) *
      5) rbe.20>0.4465 329  435.900 no fault ( 0.37690 0.00000 0.62310)
       10) BC6.20<0.0004585 248  343.800 1 edm ( 0.50000 0.00000 0.5000)
         20) Iraw.rms.19<26.95 192 253.000 1 edm ( 0.63020 0.00000 0.36980 )
             40) Iraw.sv.19<-0.155 63   61.350 no fault (0.19050 0.00000 0.81)
             80) bdf.19<0.018 11    0.000 1 edm ( 1.00000 0.00000 0.00000 ) *
             81) bdf.19>0.018 52    9.883 no fault ( 0.01923 0.00000 0.98080)*
           41) Iraw.sv.19>-0.155 129  111.300 1 edm ( 0.84500 0.00000 0.15500)
             82) bdf.19<0.02225 36   49.800 1 edm ( 0.52780 0.00000 0.47220 )
              164) counter.20<108.5 23 26.400 no fault(0.26090 0.00000 73910)*
              165) counter.20>108.5 13    0.000 1 edm (1.00000 0.00000 0.0000)*
             83) bdf.19>0.02225 93   26.510 1 edm ( 0.96770 0.00000 0.03226) *
         21) Iraw.rms.19>26.95 56  23.400 no fault ( 0.05357 0.00000 0.94640)*
       11) BC6.20>0.0004585 81    0.000 no fault ( 0.00000 0.00000 1.00000 ) *
    3) rbe.20>0.7885 93  124.100 1 edm ( 0.61290 0.38710 0.00000 )
      6) EBRms.20<0.7785 54    0.000 1 edm ( 1.00000 0.00000 0.00000 )*
      7) EBRms.20>0.7785 39   21.150 3 edm ( 0.07692 0.92310 0.00000 )*
```

Summary of One and Three EDM Notch Faults Tree Object

```
Classification tree:
snip.tree(tree = edm13.tree, nodes = c(164, 83, 7, 4, 21))
Variables actually used in tree construction:
[1] "rbe.20"       "BC6.20"       "Iraw.rms.19" "Iraw.sv.19"
[5] "bdf.19"       "counter.20"   "EBRms.20"
Number of terminal nodes:  10
Residual mean deviance:  0.2276 = 138.4 / 608
Misclassification error rate: 0.03074 = 19 / 618
```

EDM Notch Faults - Tree Object

```
node), split, n, deviance, yval, (yprob)
      * denotes terminal node

 1) root 618 804.700 no fault ( 0.35600 0.64400 )
   2) rbe.20<0.5945 366 192.700 no fault ( 0.07377 0.92620 )
     4) Ebcf.19<3.295 268  32.920 no fault ( 0.01119 0.98880 ) *
     5) Ebcf.19>3.295 98 109.100 no fault ( 0.24490 0.75510 )
      10) Iraw.sv.20<0.1355 66  40.210 no fault ( 0.09091 0.90910 ) *
      11) Iraw.sv.20>0.1355 32  43.860 fault ( 0.56250 0.43750 ) *
   3) rbe.20>0.5945 252 274.300 fault ( 0.76590 0.23410 )
     6) Iraw.rms.19<24.5 203 101.900 fault ( 0.93100 0.06897 )
      12) Iraw.rms.20<9.785 11   6.702 no fault ( 0.09091 0.90910 ) *
      13) Iraw.rms.20>9.785 192  38.890 fault ( 0.97920 0.02083 ) *
     7) Iraw.rms.19>24.5 49  27.710 no fault ( 0.08163 0.91840 ) *
```

Summary of EDM Notch Faults Tree Object

```
Classification tree:
snip.tree(tree = edm2.tree, nodes = c(4, 10, 13, 11, 7))
Variables actually used in tree construction:
[1] "rbe.20"      "Ebcf.19"     "Iraw.sv.20"  "Iraw.rms.19"
"Iraw.rms.20"
Number of terminal nodes:  6
Residual mean deviance:  0.3109 = 190.3 / 612
Misclassification error rate: 0.05178 = 32 / 618
```

Small Integral Race Spall - Tree Object

```
node), split, n, deviance, yval, (yprob)
      * denotes terminal node

 1) root 618 440.800 no fault ( 0.11490 0.88510 )
   2) EBpk2pk.2<1.535 232 285.800 no fault ( 0.30600 0.69400 )
     4) be.1<0.7025 74  10.590 no fault ( 0.01351 0.98650 ) *
     5) be.1>0.7025 158 217.000 no fault ( 0.44300 0.55700 )
      10) be.1<1.115 64  58.730 fault ( 0.82810 0.17190 )
        20) bdf.1<0.0552 10   0.000 no fault ( 0.00000 1.00000 ) *
        21) bdf.1>0.0552 54   9.959 fault ( 0.98150 0.01852 ) *
      11) be.1>1.115 94  88.860 no fault ( 0.18090 0.81910 )
        22) rte.2<9.36 48   0.000 no fault ( 0.00000 1.00000 ) *
        23) rte.2>9.36 46  60.600 no fault ( 0.36960 0.63040 ) *
   3) EBpk2pk.2>1.535 386   0.000 no fault ( 0.00000 1.00000 ) *
```

Summary of Small Integral Race Spall Tree Object

```
Classification tree:
snip.tree(tree = port.tree, nodes = c(21, 4, 23))
Variables actually used in tree construction:
[1] "EBpk2pk.2" "be.1"      "bdf.1"      "rte.2"
Number of terminal nodes:  6
Residual mean deviance:  0.1326 = 81.16 / 612
Misclassification error rate: 0.03074 = 19 / 618
```

APPENDIX D. [CROSS-VALIDATION PLOTS]

The following plots were used to determine the right size for each tree model derived. The plots are in the same order as the tree models in chapter 4.

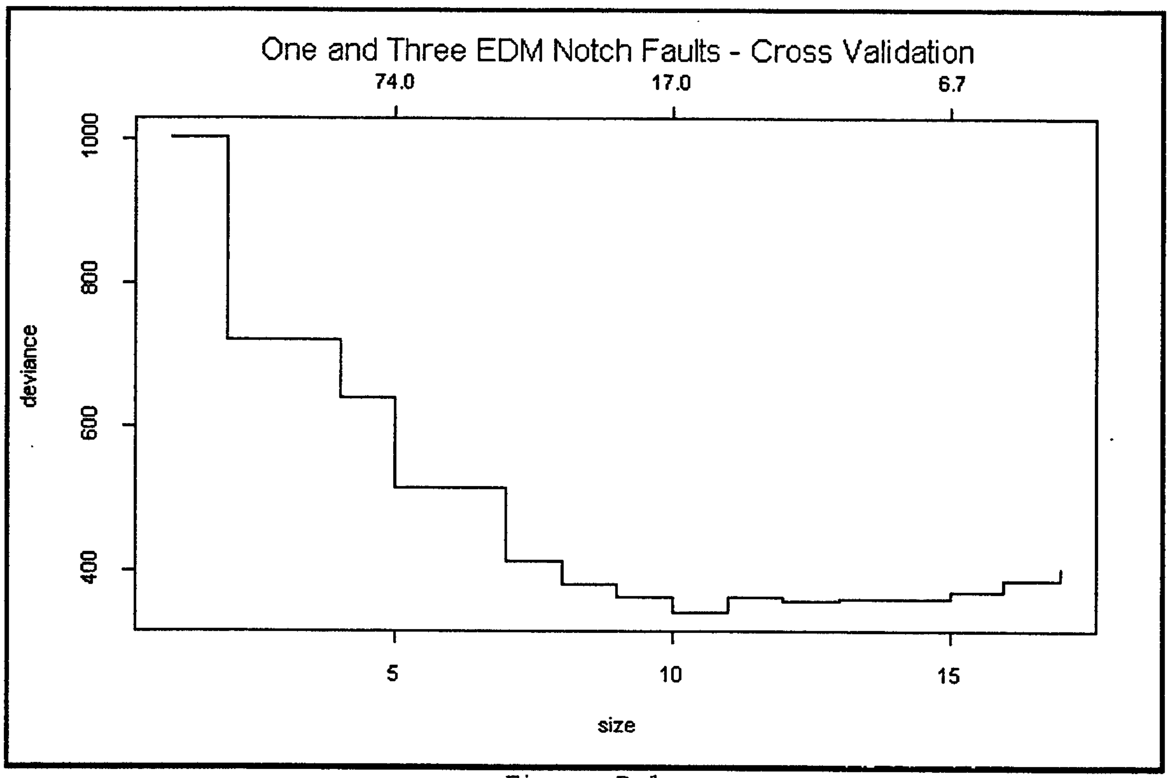
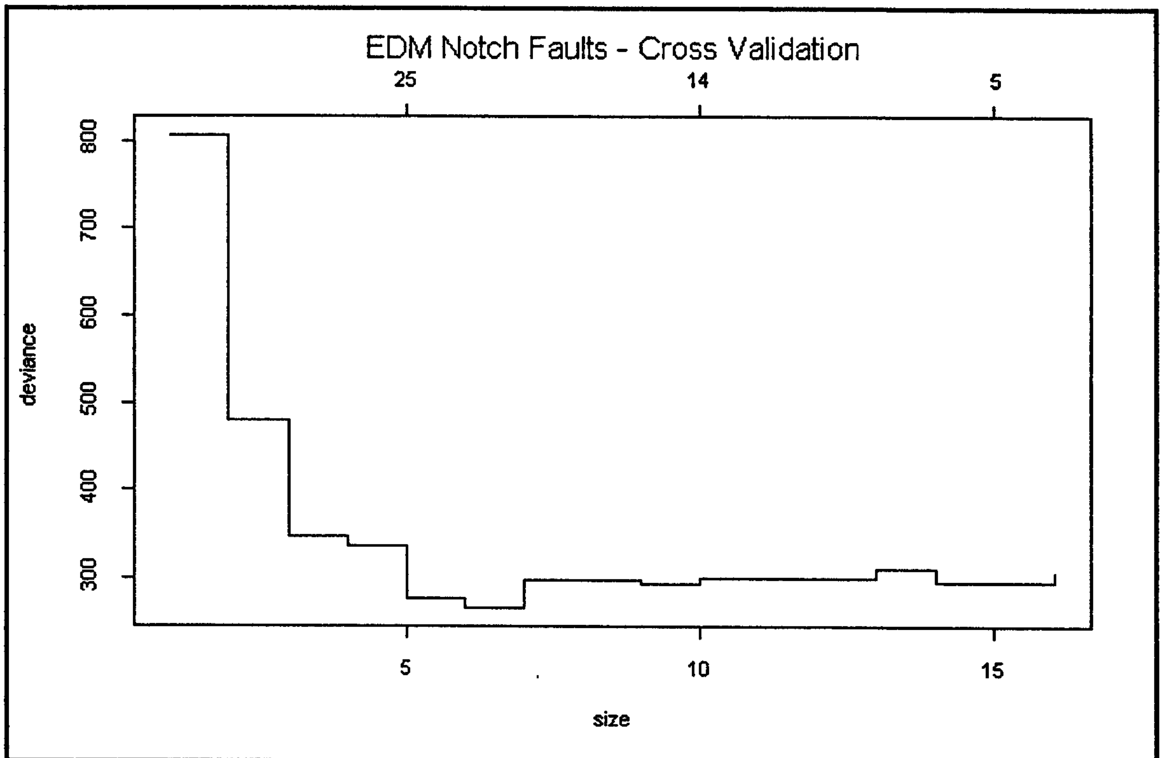
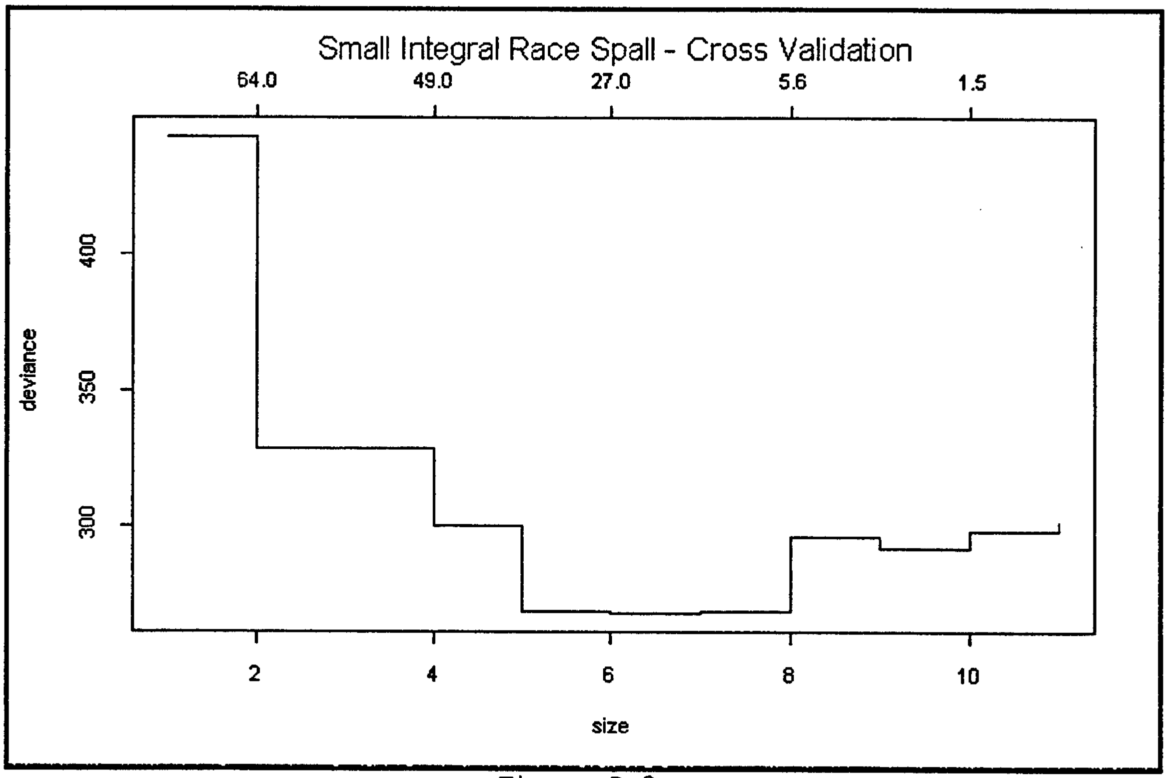

One and Three EDM Notch Faults - Cross Validation
74.0
17.0
6.7
1000
800
600
400
deviance
5
10
15
size

Figure D-1



EDM Notch Faults - Cross Validation
25
14
5
800
700
600
500
400
300
deviance
5
10
15
size

Figure D-2


Small Integral Race Spall - Cross Validation
64.0
49.0
27.0
5.6
1.5
deviance
400
350
300
2
4
6
8
10
size
Figure D-3

List of References

1. Cleveland, G.P.; Trammel, C., An Integrated Health and Usage Monitoring System for the SH-60B Helicopter, p. 1 - 3, American Helicopter Society, Inc., Washington, D.C., Jun. 1996.

2. Parry, D., Evaluating HUMS, Avionics Magazine, p. 28 - 32, February 1996.

3. Loeslein, G. F.; Kinker, L. E.; Vetere, P., Cost-Benefit Analysis for U.S. Navy and Marine Corps - Helicopter Safety-of-Flight Systems, p. 1 - 11, Flight Dynamics and Safety Inc, 1995.

4. Rovenstine, M. J., Classification Analysis of Vibration Data from SH-60b Helicopter Transmission Test Facility (HTTF), p. 2 - 30, Naval Postgraduate School, 1997.

5. Emmerling, W. C.; Hess, A. J.; Hayden, R. E., Helicopter Integrated Diagnostic System (HIDS), p. 1-10, American Helicopter Society, Inc., 1996.

6. Hamilton, L.C., Regression with Graphics - A Second Course in Applied Statistics, 249 - 257, Wadsworth, Inc., Belmont, California, 1992.

7. Statistical Sciences, S-PLUS Guide To Statistical & Mathematical Analysis, Version 3.3, Seattle: StatSci, a division of MathSoft, Inc., 1995.

8. Spector, P., An Introduction to S and S-Plus, p.244 - 246,, Wadsworth, Inc., Belmont, California, 1994.

9. Chambers, J. M.; Hastie, T. J., Statistical Models in S, p.412 - 417, Wadsworth Inc., 1992.

10. Breiman, L.; Friedman, J. H.; Olshen, R. A.; Stone, C. J., Classification and Regression Trees, p. 174 - 189, Wadsworth Inc., 1984.

INITIAL DISTRIBUTION LIST

No. of copies

1. Defense Technical Information Center.. 2
   8725 John J. Kingman Rd., STE 0944
   Ft. Belvoir, VA 22060-6218

2. Dudley Knox Library 2
   Naval Postgraduate School
   411 Dyer Rd.
   Monterey, CA 93943-5101

3. Prof. Robert Read, Code OR/De... 1

4. Prof. Richard Larson, Code OR/De... 1

5. LT. Gregory Anderson.. 1
   3959 Ester Dr.
   Atlanta, GA 30331