Final Report For:

DEMONSTRATION OF THE USE OF ADAPT TO DERIVE PREDICTIVE MAINTENANCE ALGORITHMS FOR THE KSC CENTRAL HEAT PLANT

AVSD-0084-73-RR

NOVEMBER 1972

By

Herbert E. Hunter

Prepared For:

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
JOHN F. KENNEDY SPACE CENTER
KENNEDY SPACE CENTER, FLORIDA 32899

CONTRACT NO. NAS10-7926

Prepared By:

AVCO GOVERNMENT PRODUCTS GROUP
AVCO SYSTEMS DIVISION
WILMINGTON, MASSACHUSETTS 01887

ACKNOWLEDGEMENT

This work was supported by NASA Kennedy Space Center under Contract NAS10-7926. The NASA technical representative was Mr. N. P. Salvail. In addition, the author wishes to thank Messrs. Nething and Guggenheim of Kennedy Space Center for their many helpful discussions concerning the heating plant and their efforts in gathering and supplying the data used for this study. Thanks are also due to Mr. Richard Amato of Avco for his assistance in the analysis of this data, in particular Mr. Amato selected the variables to be used in the data vector, supervised the transformation of the data from the log sheets to the punch cards and performed much of the ADAPT analysis. Thanks are also due to Mr. Jake Apsel and Mrs. Patricia Gaudet for their assistance in processing this data through the ADAPT programs.

TABLE OF CONTENTS

ABSTRACT

The Avco Data Analysis and Prediction Techniques (ADAPT), (a series of empirical data analysis programs based on the concept that pattern recognition and regression should be preceded by a reduction of dimensionality based on the Karhunen-Loeve Expansion), were applied to two years of historical data recorded on the Kennedy Space Flight Center Central Heating Plant. Detection laws capable of detecting failures in the heat plant up to three days in advance of the occurrence of the failure were successfully derived and demonstrated. The projected performance of these algorithms yielded a detection probability of 90% with false alarm rates of the order of 1 per year for a sample rate of 1 per day with each detection followed by 3 hourly samplings. This performance was verified on 173 independent test cases. The program also demonstrated diagnostic algorithms and the ability to predict the time to failure to approximately plus or minus 8 hours up to three days in advance of the failure.

The ADAPT programs produce simple algorithms which have a unique possibility of a relatively low cost updating procedure. The algorithms have been implemented on general purpose computers at Kennedy Space Flight Center and will be tested against current data.

The study concludes that the successful demonstration of the detection and classification algorithms demonstrates the feasibility of a new maintenance concept based on the demand rather than a preset schedule. This approach will save cost and avoid the possibility of introducing failures as a part of the inspection procedure. This maintenance concept should have applicability to a large variety of industrial and government facilities as well as the maintenance of complex systems such as spacecraft and other large complex systems.

List of Figures

1.0 INTRODUCTION

This report presents the results of a study program which demonstrated the feasibility of a demand preventive maintenance (DPM) approach to maintenance of the KSC central heat plant. This feasibility was demonstrated by using the Avco Data Analysis and Prediction Techniques (ADAPT) to derive simple algorithms for 1) detecting incipient failures of the central heat plant, 2) diagnosing expected cause of this incipient failure, and 3) determining the time remaining to the occurrence of this failure. Demonstration of the feasibility of providing these algorithms leads directly to the feasibility of utilizing a demand preventive maintenance scheme as a replacement or adjunct to the present schedule preventive maintenance (PM) scheme.

The objective of the conventional scheduled preventive maintenance scheme is to avoid failures by scheduling the maintenance of various elements of a complex system in such a way that each element is inspected and/or repaired prior to the occurrence of a failure. The DPM approach replaces this concept or at least complements it with the idea that diagnostic measurements will be taken on the system and used to predict an incipient failure before it occurs. When this incipient failure has been detected. the corrective action and maintenance required to prevent this failure from occurring will be performed. Thus the availability of ADAPT detection algorithms allows preventive maintenance to be performed on demand rather than on a scheduled basis.

This report presents the derivation of, performance projections for, and test verification that simple detection algorithms can be derived which would detect approximately 90% of the failures occurring in the KSC heat plant with a false alarm rate of approximately one per year for sample rate of one per day with each detection followed by three hourly samplings. The potential to derive detection algorithms with even greater performance is demonstrated; however, the requirements for maintenance on the KSC central heat plant would not justify the additional effort required to derive, verify, and implement the more complex sequence of algorithms required to achieve this gain in performance.

The demonstration algorithm for detecting incipient failures was developed and its expected performance projected from the ADAPT analysis of the learning data. This performance was then verified by testing 173 independent test cases. Algorithms were also developed and their performance projected to demonstrate the diagnosis of failures in the atomizing steam boiler and in boiler No. 1. An algorithm was developed and its performance projected for predicting the number of hours remaining until failure of the atomizing steam boiler.

The application of these algorithms to the KSC central heat plant has been illustrated by two scenarios. The first scenario illustrates how one would apply the ADAPT maintenance algorithms in a manual mode. The maintenance algorithms derived can be utilized in this manual mode without any modifications to the present KSC heat plant and its instrumentation by using existing computers at KSC. The second scenario shows how these same algorithms can be used in conjunction with an automated data monitoring system to completely automate the entire diagnosis and analysis of the KSC central heat plant. For this case the same algorithms can be incorporated in the computer dedicated to the monitoring system and the entire DPM program implemented in a completed automated fashion.

The programs required to implement ADAPT algorithms on existing KSC computers have been developed and implemented on these computers by KSC personnel. It is also possible to implement the programs required to make use of the optimum representation to update the algorithms on existing KSC computers. This would allow KSC to update the algorithms to account for minor changes in the system.

The next section of this report will summarize the results and recommendations resulting from this study. This will be followed by a description of the KSC central heat plant and the scenarios illustrating the application of these algorithms to the DPM of the KSC central heat plant. Section 4 reviews the ADAPT programs and approach to empirical data analysis. The derivation and evaluation of the detection, diagnostic and time to failure algorithms are presented in Sections 5, 6 and 7 respectively.

2.0 RESULTS AND RECOMMENDATIONS

The major result of this study was the demonstration of the feasibility of developing a predictive maintenance scheme based on the use of ADAPT derived algorithms for the Kennedy Space Flight Center central heat plant. This system can be implemented with the central heat plant in its present configuration using a manual mode of data collection, transportation and submission to existing general purpose computers, or the algorithms may be incorporated into a completely automated data retrieval and logging system. In this latter system the entire predictive preventive maintenance scheme can be incorporated in a maintenance computer which would automate all of the functions required to furnish the maintenance instructions. The requirements to implement either system, assuming that in the latter case one is already procuring the computer and data collection system for the automated data gathering and recording system, is the development of the complete set of diagnostic algorithms and a small effort to develop the software and logic required to implement and interpret the detection in diagnostic algorithms.

The feasibility of using a demand preventive maintenance system rests primarily on the ability to detect incipient failure such than maintenance may be performed on demand, that is when the system is just about to fail rather than on a scheduled basis. The implementation of this demand preventive maintenance system is considerably simplified if one can also diagnose which component is about to fail such that the maintenance instructions can be specific. Thus the primary requirement to establishing feasibility was the demonstration of the feasibility of using the ADAPT programs to derive algorithms to detect incipient failures of the central heat plant. A secondary requirement was to show the feasibility of deriving diagnostic algorithms and a tertiary objective was to show the feasibility of estimating the time of failure once the failure mode had been diagnosed. Since the detection algorithm is most critical to the feasibility of implementing the predictive preventive maintenance system, the major effort was to demonstrate the feasibilibility of the detection algorithm. The first step was to investigate three different types of detection algorithms. These three types were universal detection algorithms, algorithms based on subdivision by types of failures and algorithms based on subdivisions by natural ADAPT grouping. Exploratory studies were carried out with an initial data set ranging from 30 to 100 cases. The application of the ADAPT programs to these data sets resulted in the detection algorithms whose performance are summarized in Figure 2.1.

Figure 2.1 presents the detection probability versus the false alarm rate for each of the six detection algorithms studied. These performances are based on projections of the learning data. There are many advantages to multiple applications of the algorithm prior to initiating corrective maintenance action. Some of these advantages include less severe requirements on the performance of the algorithm and significantly smaller amounts of test data required to proof test the algorithm. These multiple applications allow one to obtain false alarm rates of the order of one per year and detection probabilities greater than 85% for any algorithm whose single application performance exceeds a detection probability of approximately 9 for a false alarm rate of .3. Since this performance is adequate for maintenance of the KSC heat plant, it can be used to separate acceptable from unacceptable detection algorithms.

Applying the above criteria to the results shown in Figure 2.1, we see that all of the detection algorithms are acceptable from a performance standpoint. The universal boiler #1 detection algorithm has significant advantages over all of the other algorithms in terms of ease of application, cost of both the development and use of the predictive preventive maintenance system, and the break-in time required to debug this system. For these reasons the universal boiler #1 detection algorithm was selected as the best algorithm for application to the KSC central heat plant. It must be emphasized that another algorithm might be required or more desirable for other applications.

Since the detection algorithm represents the most critical element in the feasibility of the demand preventive maintenance system, further development of this algorithm was carried out to provide independent test results, verify the projected performance and demonstrate the ability of the ADAPT programs to project learning data performance to test cases. The results of these studies are summarized in Figure 2.2 as plots of the detection probability versus false alarm rate for the 20 dimensional boiler #1 detection algorithm. Again this algorithm is not the best performing of the universal detection algorithms; however, taking all factors into consideration it is the recommended algorithm for the KSC central heat plant. The solid symbols show the results of applying this algorithm to 173 independent test cases which were not used in the original data set. These test cases included considerable additional variation over time of day as well as day of year relative to the original learning data. In addition to the testing of these 173 cases, testing was also performed on 15 cases where boiler #2 was substituted for boiler #1. This algorithm proved to be effective in diagnosing failures of the boiler #2 configuration. Tests were also performed on 19 cases which were taken prior to the major changes which were made in the distribution system early in 1970. These test cases showed that this algorithm could not account for these major changes in the distribution system. The details of these tests are presented in section 5.4. In summary, the testing demonstrated that the ADAPT projections of performance were valid and therefore the feasibility of deriving an algorithm for detecting incipient failure of the KSC central heat plant was verified.

The same exploratory analysis used to project the performance of the detection algorithm was applied to two different diagnostics algorithms. The first was on algorithms to diagnose boiler #1 failure versus all other failues and the second was algorithms to diagnose the atomizing boiler versus all other failures. The projected performance for these two algorithms is summarized in Figure 2.3 as plots of detection probability versus false alarm rate. Examination of this figure shows that the performance of these algorithms should be superior to the performance of any of the detection algorithms. Single applications of these algorithms will yield detection probabilities in excess of .96 with failure rates of one per year. This type of algorithm can only be constructed for failure modes for which failures have previously occurred. Thus there is always the possibility that a new type of failure will occur and this specific diagnostics will not be possible. For this situation it is recommended that the ADAPT programs capability to provide nearest neighbor analysis and relative importance information be utilized to provide additional information to assist in diagnosing the cause of the new type of failure. Algorithms should also be derived for isolating failures to certain sub-systems of the KSC central heat plant. In fact, the atomizing steam boiler diagnostic algorithms is actually such an algorithm since it was developed using several different types of failures occurring in the atomizing steam boiler subsystem. It is likely that any failure in the atomizing steam boiler subsystem would be diagnosed even if it were not identical to the specific failure which was used in the learning data.

For those cases where experience with a specific type of failure is sufficient to provide a reasonable number of cases, one might expect to be able to use the ADAPT parameter estimation capability to estimate the time remaining until the failure will occur. In order to demonstrate this capability, the data on the atomizing steam boiler failures were used in the ADAPT program to derive a time-to-failure algorithm. Figure 2.4 is a plot of the time-to-failure as estimated by the ADAPT algorithms versus the actual time-to-failure. Examination of this figure shows that the ADAPT algorithm is able to predict to within approximately six hours the time-to-failure up to three days in advance for approximately 70% of the cases.

Tables 2.1 thru 2.4 present the 20-dimensional universal detection algorithm, the two diagnostic algorithms and the time-to-failure algorithms which were derived as a result of this study. Each detection or classification algorithm consists of two steps: Step 1 is an equation (i.e. dot product) to compute a number and Step 2 is the rule for using the number. For a prediction algorithm Step 1 provides the number to be predicted. Examination of these tables shows that the implementation of these algorithms is a simple procedure which if necessary could be implemented by hand, although it is far more convenient and reliable to implement these algorithms on a computer. They have already been implemented on the general purpose computers at KSC. Table 2.5 lists the measurements which are associated with each of the index values for the algorithm presented in Table 2.1 and Table 2.6 lists each of the measurements which would be associated with the indices for Tables 2.2 through 2.4.

The availability of these three types of algorithms allows us to implement a demand preventive maintenance system. The recommended procedure for accomplishing this is illustrated in Figure 2.5. The heat plant measurements would be taken and processed through the incipient failure detection algorithm which was presented in Table 2.1. If this algorithm produced a value greater than zero, the system is operating normally and no action is required. If this algorthm produces a value less than zero, an incipient failure is indicated. This would initiate further analysis actions. The data would be recorded each hour for the next three hours and the algorithm repeated. If three confirming detections were achieved, then the decision would be made that an incipient failure was to be expected. The data would then be processed through the diagnostic and time-to-failure algorithms such as those presented in Tables 2.3 through 2.5. These algorithms would provide the basis upon which maintenance instructions would be prepared by the maintenance decision logic. This process could be carried out exactly as described with the present KSC central heat plan instrumentation by using the current measurements as recorded in the heating plant log, punching these measurements onto punch cards and feeding them to an existing general purpose computer at KSC. Alternatively, if the new automated data collection and recording system is obtained, the entire procedure from the recording of the measurements through the application of the algorithms and the performing of the maintenance decision logic can be carried out within the maintenance computer required to control the data collection.

The tests performed on variations from the learning system including the substitution of boiler #2 for boiler #1 and the use of test cases obtained before the major changes of early 1970 were incorporated into the distribution system have shown that the algorithm presented in Table 2.1 is insensitive to relatively minor changes such as the substitution for boiler #2 for boiler #1, but the major changes associated with the major modifications of the distribution system seriously degraded the performance of this algorithm. This indicates that it will be desirable to have a capability to update the ADAPT algorithms from time to time. This updating capability also allows one to incorporate new failures into the learning base as they occur. This can be accomplished as is outlined in Section 3. In order to do this it is necessary to store certain portions of the data obtained during the normal processing. A random sampling of the passing cases is required to keep the good class up to date. It is also desirable to keep each of the failed cases as a future learning case. Thus, the flow diagram of Figure 2.5 shows a random sampling of the passing cases and a complete sampling of the failure cases. Again, this can be done manually or with the automated system.. The key result of this review of the application of the procedure is that the ADAPT algorithms can be incorporated into a demand preventive maintance scheme at KSC without any additional hardware procurement in either its present configuration or in the planned automated data recording configuration.

The successful achievement of detection and diagnostic algorithms required to implement a demand preventive maintenance scheme on the KSC central heat plant implies that other base facilities at KSC, other NASA centers and in industry in general which are made up of a large number of interrelated subsystems may be maintained by a demand preventive maintenance technique such as described here. The success of this technique on this complex but relatively unsophisticated system also indicates a good prognosis for the application of this approach to detecting incipient failures in more sophisticated systems such as space shuttle and other spacecraft checkout and post-flight maintenance.

Careful examination of the ADAPT produced relative importance vector, provided information useful to improving the scheduled preventive maintenance approach and the design of the system. For example, the times to preventive maintenance which have positive values in the relative importance vector are an indication of preventive maintenance which may be being performed too often since the performance of the system is better when one is a long time away from the preventive maintenance. On the other hand, those preventive maintenance index values which have negative values are items which require more preventive maintenance. This phenomena is discussed in more detail in Section 5.5.

The successful demonstration of the feasibility of developing algorithms for detecting incipient failure leads to the immediate recommendation that as much experience as possible should be obtained with practical application of this algorithm. The best way to achieve this is to start an immediate monitoring of the present central heating facility by applying the algorithm presented in Table 2.1 to this facility on a regular basis. Based on the results obtained in evaluating the case where boiler #2 is substituted for boiler #1, it is also recommended that this algorithm be applied to either boiler #1 or boiler #2 operating by themselves.

It is also recommended that the effort be initiated to develop the remaining algorithms, optimize those algorithms and provide the proof testing of the algorithms required to provide a complete set of algorithms for implementing the demand preventive maintenance system on the KSC heat plant. The use of the ADAPT programs to provide a demand preventive maintenance capability for other base facilities and other spacecraft systems and spacecraft checkout problems should be implemented. The primary requirement to accomplish this is the availability of data which can be used as learning data to derive the required detection and diagnostic algorithms. The results on the KSC central heat plant provide an extremely high confidence that given a relatively complete monitoring of most any complex system, the ADAPT programs can derive algorithms capable of detecting incipient failures and diagnosing the cause of this failure.

TABLE 2.1

ALGORITHM FOR CLASSIFICATION OF NORMAL AND OUT OF TOLERANCE OPERATION OF THE KSC HEAT PLANT (20 DIMENSIONS)

STEP 1 CALCULATE W_T FROM THE VALUES OF THE MEASUREMENTS, V_i^T, AT TIME "T" USING

$$W_T = A_i^{C1} \; V_i^T \; + 19.0$$

WHERE A_i^{C1} ARE GIVEN AS A FUNCTION OF THE MEASUREMENT INDEX, i BELOW

i	A^{C1}	i	A^{C1}
1	0.3370E-01	26	0.4601E-01
2	0.1476E 00	27	-0.1478E [illegible]
3	-0.2899E-01	28	-0.2101E 00
4	0.3997E-01	29	0.1079E 00
5	[illegible].3365E-01	30	-0.6919E-01
6	-0.3485E-01	31	-0.8580E-01
7	0.9211E-01	32	-0.5522E-01
8	0.6459E-01	33	-0.3759E-01
9	-0.1112E 00	34	0.4345E-01
10	-0.6745E-01	35	0.5909E-01
11	0.1669E-02	36	0.9351E-01
12	-0.1393E 00	37	0.6946E-01
13	0.1190E 00	38	-0.5041E-01
14	0.1012E 00	39	0.6280E-01
15	-0.7079E-01	40	0.2663E-02
16	-0.8282E-01	41	-0.2277E-01
17	0.9696E-02	42	-0.8099E-01
18	0.9323E-01	43	-0.5216E-02
19	-0.2355E-01	44	-0.7022E-02
20	0.5817E-01	45	-0.1014E 00
21	-0.5092E-01	46	[illegible]
22	-0.1702E-01	47	-0.7764E-01
23	-0.3057E-01	48	-0.1764E-01
24	0.7942E-01	49	-0.5168E-01
25	0.3678E-02	50	-0.4295E-01

STEP 2:

$W_T > 0$ Heat Plant Operation Normal

$W_T < 0$ Heat Plant Operation Out of Tolerance Apply Diagnostic Algorithms

TABLE 2.2

ALGORITHM FOR DIAGNOSING FAILURE OF ATOMIZING STEAM BOILER

STEP 1 AFTER DETECTING FAILURE CALCULATE $W_T^{D1} = \sum A_i^{D1} \; V_i^T - 300$

i	A_i^{D1}	i	A_i^{D1}	i	A_i^{D1}	i	A_i^{D1}
1	-0.2132D 00	51	0.0	101	0.0	151	0.2840D 00
2	0.5564D-01	52	0.0	102	0.0	152	0.8311D-01
3	0.2536D 00	53	0.0	103	0.0	153	0.7695D-01
4	-0.5337D-02	54	0.0	104	-0.2764D 00	154	0.1895D 00
5	-0.1732D 00	55	0.0	105	0.7230D-01	155	0.7722D-01
6	-0.8104D-01	56	0.0	106	0.8657D-01	156	0.3494D-01
7	-0.3819D-02	57	0.9246D-01	107	-0.5866D-01	157	-0.9306D-01
8	-0.4914D 00	58	-0.4708D-01	108	-0.1131D 00	158	-0.4859D 00
9	-0.2695D 00	59	0.3232D 00	109	-0.6051D-01	159	-0.9306D-01
10	0.1425D 00	60	-0.1570D-01	110	0.8681D-02	160	-0.4982D 00
11	-0.1929D 00	61	-0.5748D-01	111	0.2383D-01	161	-0.1526D 00
12	0.3064D 00	62	0.4537D-01	112	-0.4014D-01	162	0.2697D-01
13	0.2269D-01	63	0.6210D-01	113	-0.1335D-01	163	0.2697D-01
14	0.2076D 00	64	0.0	114	-0.1711D 00	164	0.2697D-01
15	0.1647D 00	65	0.0	115	0.4049D-02	165	0.2697D-01
16	0.2163D 00	66	0.0	116	0.1551D-01	166	0.9058D-01
17	0.0	67	0.0	117	0.7115D-01	167	0.1819D 00
18	0.0	68	0.0	118	0.8551D-01	168	0.3601D-01
19	0.0	69	0.0	119	0.1506D 00	169	0.2425D-01
20	0.0	70	0.0	120	-0.3309D 00	170	-0.6736D-01
21	0.0	71	0.0	121	0.1839D 00	171	0.7713D-01
22	0.0	72	0.0	122	-0.5610D-01	172	0.5007D-01
23	0.0	73	0.0	123	-0.7954D-01	173	0.2786D-01
24	0.0	74	0.0	124	-0.3670D-01	174	-0.9915D-01
25	0.0	75	0.0	125	-0.4566D-01	175	-0.2427D-01
26	0.0	76	0.0	126	0.2543D 00	176	-0.9915D-01
27	0.0	77	0.0	127	-0.2372D 00	177	0.9859D-01
28	0.0	78	0.6732D-01	128	-0.3515D 00	178	-0.1009D 00
29	0.0	79	0.5723D-01	129	-0.1311D 00	179	0.5319D-01
30	0.0	80	-0.1226D 00	130	0.7801D-01	180	0.3601D-01
31	-0.2378D 00	81	-0.2157D-01	131	0.7289D-01	181	0.5319D-01
32	-0.1031D 00	82	-0.3690D-01	132	-0.1503D 00	182	0.3601D-01
33	0.2405D 00	83	0.9302D-01	133	-0.1308D-01	183	0.5319D-01
34	-0.7819D-01	84	0.7677D-01	134	-0.1547D 00	184	-0.6736D-01
35	-0.3700D 00	85	-0.3930D-01	135	-0.1152D-01	185	0.7713D-01
36	-0.1066D 00	86	0.2233D 00	136	-0.1311D 00	186	-0.6736D-01
37	0.7912D-01	87	0.2546D 00	137	0.2989D-01	187	-0.6736D-01
38	-0.2486D 00	88	-0.6907D-01	138	-0.1533D 00	188	0.1540D 00
39	0.3214D 00	89	-0.1227D 00	139	0.2272D-01	189	-0.8787D-01
40	0.1663D 00	90	-0.1048D-01	140	-0.1711D 00	190	0.3648D 00
41	-0.2316D 00	91	-0.1246D-01	141	0.2347D 00	191	0.1052D 00
42	-0.2074D 00	92	0.4553D 00	142	-0.8152D-01	192	0.8414D-01
43	-0.4736D-01	93	-0.1328D 00	143	0.2314D 00	193	0.0
44	-0.2510D 00	94	0.4126D 00	144	0.6953D-01	194	0.0
45	0.4505D 00	95	-0.2923D 00	145	-0.1117D-01	195	0.0
46	-0.6907D 00	96	-0.1853D 00	146	-0.1838D 00	196	0.0
47	0.4956D-01	97	-0.1212D 00	147	-0.6466D-01	197	0.0
48	0.1879D 00	98	0.0	148	-0.3571D 00	198	0.0
49	0.2671D 00	99	0.0	149	-0.3871D 00	199	0.0
50	-0.5663D 00	100	0.0	150	0.6559D-01	200	0.0

STEP 2 IF: $W_T^{D1} > 0$ FAILURE WILL OCCUR IN ATOMIZING STEAM BOILER

$W_T^{D1} < 0$ FAILURE WILL OCCUR ELSEWHERE

TABLE 2.3

ALGORITHM FOR DIAGNOSING FAILURE OF BOILER NO. 1

STEP 1 AFTER DETECTING FAILURE CALCULATE: $W_T^{D2} = \sum_i A_i^{D2} V_i^T - 51.2$

i	A_i^{D2}	i	A_i^{D2}	i	A_i^{D2}	i	A_i^{D2}
1	4.86510D-02	51	0.0	101	0.0	151	8.75482D-03
2	4.40946D-02	52	0.0	102	0.0	152	7.45773D-03
3	-9.16094D-03	53	0.0	103	0.0	153	1.13283D-02
4	5.23347D-04	54	0.0	104	9.39297D-03	154	2.79537D-02
5	3.47870D-04	55	0.0	105	8.48327D-02	155	-1.31791D-03
6	-1.99815D-01	56	0.0	106	-9.98555D-02	156	8.22751D-03
7	1.03595D-03	57	0.0	107	4.81049D-02	157	2.59404D-03
8	1.09414D-03	58	3.21523D-02	108	9.03274D-03	158	-3.13417D-03
9	2.38571D-02	59	-2.77907D-02	109	2.43903D-02	159	2.59404D-03
10	3.58231D-02	60	-1.02201D-03	110	6.29706D-02	160	-4.43655D-03
11	2.28144D-03	61	5.50108D-02	111	2.44739D-02	161	-7.22698D-03
12	-3.94471D-02	62	-1.50923D-03	112	3.53226D-02	162	-1.47098D-02
13	-2.60851D-02	63	3.27205D-02	113	4.66172D-02	163	-1.47098D-02
14	3.23174D-02	64	0.0	114	9.59325D-03	164	-1.47098D-02
15	9.62923D-03	65	0.0	115	-4.94980D-02	165	-1.47098D-02
16	1.08367D-02	66	0.0	116	2.93772D-02	166	-9.35165D-04
17	0.0	67	0.0	117	2.01263D-02	167	4.52433D-03
18	0.0	68	0.0	118	2.19234D-02	168	3.67917D-02
19	0.0	69	0.0	119	2.70535D-02	169	4.51974D-02
20	0.0	70	0.0	120	6.94420D-03	170	-8.50016D-03
21	0.0	71	0.0	121	5.55624D-02	171	1.18912D-02
22	0.0	72	0.0	122	5.28442D-02	172	5.45344D-02
23	0.0	73	0.0	123	4.85347D-02	173	4.25216D-02
24	0.0	74	0.0	124	2.80578D-02	174	2.44466D-03
25	0.0	75	0.0	125	2.68753D-02	175	6.25644D-02
26	0.0	76	0.0	126	4.61449D-02	176	2.44466D-03
27	0.0	77	0.0	127	4.21174D-02	177	-4.39298D-02
28	0.0	78	5.30597D-03	128	2.25324D-02	178	4.73922D-03
29	0.0	79	-1.07082D-02	129	1.83631D-02	179	2.16207D-02
30	0.0	80	-2.13240D-02	130	1.53337D-03	180	3.67917D-02
31	-2.25368D-03	81	8.01295D-02	131	-1.77270D-03	181	2.16207D-02
32	4.41380D-03	82	1.32030D-02	132	3.04714D-02	182	3.67917D-02
33	-2.26990D-02	83	-1.16000D-02	133	4.40667D-02	183	2.16207D-02
34	2.94523D-02	84	1.92300D-02	134	1.33295D-02	184	-8.50016D-03
35	4.33318D-02	85	-1.50770D-02	135	3.69757D-02	185	1.18912D-02
36	-8.69232D-03	86	2.04513D-02	136	1.83631D-02	186	-8.50016D-03
37	2.66861D-02	87	2.17179D-02	137	1.94987D-02	187	-8.50016D-03
38	-7.28495D-03	88	3.39455D-02	138	2.42417D-02	188	-1.13794D-01
39	3.34265D-02	89	3.6286D-02	139	1.32552D-02	189	2.32612D-02
40	-1.52530D-02	90	3.43495D-02	140	9.59325D-03	190	1.08842D-01
41	-4.18653D-03	91	-9.11438D-03	141	1.25251D-02	191	4.11079D-03
42	-1.36906D-04	92	-7.85253D-02	142	5.05192D-02	192	3.28856D-03
43	3.33899D-03	93	4.46470D-02	143	1.71842D-02		
44	-6.39224D-02	94	-4.44652D-02	144	2.92695D-02		
45	-1.03611D-01	95	5.21361D-02	145	4.75511D-02		
46	2.59059D-02	96	4.63533D-02	146	2.09525D-03		
47	-3.61378D-02	97	1.10596D-02	147	2.31197D-02		
48	4.17447D-02	98	0.0	148	8.31301D-03		
49	4.89406D-02	99	0.0	149	-3.49755D-02		
50	5.28301D-02	100	0.0	150	1.80771D-02		

STEP 2 IF: W_T^{D2} >0 FAILURE WILL OCCUR IN BOILER NO. 1

D2 ... OCCUR ELSEWHERE

TABLE 2.4

ALGORITHMS FOR PREDICTING TIME TO FAILURE OF ATOMIZING STEAM BOILER

$$\text{TIME TO FAILURE (HR)} = \sum A_i^P \quad V_i^T \quad + 29.3$$

i	A_i^P	i	A_i^P	i	A_i^P	i	A_i^P
1	-1.76544180D-01	49	7.69162430D-03	97	2.09834940D-02	145	1.86610330D-02
2	-1.33790860D-02	50	3.10857680D-02	98	0.0	146	6.19402910D-03
3	8.64296930D-02	51	0.0	99	0.0	147	4.10463750D-02
4	3.52513850D-06	52	0.0	100	0.0	148	2.01349670D-02
5	3.78578740D-02	53	0.0	101	0.0	149	2.71054360D-02
6	2.85498850D-03	54	0.0	102	0.0	150	1.71670980D-02
7	-9.01440790D-05	55	0.0	103	0.0	151	1.77749200D-02
8	-3.06569940D-01	56	0.0	104	-1.98051640D-02	152	1.71970980D-02
9	3.66557190D-03	57	-1.46870870D-02	105	-1.20708290D-01	153	2.44514780D-02
10	-1.69025050D-01	58	-5.31684020D-02	106	-2.57869770D-01	154	-6.57555610D-03
11	4.67359060D-02	59	-3.49608870D-02	107	6.06282820D-01	155	3.11280360D-02
12	2.26342300D-02	60	2.92445900D-02	108	2.71054360D-02	156	1.53009850D-02
13	5.46520550D-02	61	-5.60001670D-03	109	3.28956820D-02	157	2.36004720D-03
14	1.57226850D-02	62	9.39663750D-03	110	5.01400540D-03	158	2.01349670D-02
15	3.96067280D-03	63	-1.05380460D-02	111	2.68115250D-02	159	2.36004720D-03
16	-5.75452210D-05	64	0.0	112	3.28956820D-02	160	2.01349670D-02
17	0.0	65	0.0	113	1.99549440D-02	161	1.42348490D-02
18	0.0	66	0.0	114	4.07524640D-02	162	1.27609150D-02
19	0.0	67	0.0	115	7.91516060D-02	163	1.27609150D-02
20	0.0	68	0.0	116	1.74810090D-02	164	1.27609150D-02
21	0.0	69	0.0	117	1.71670960D-02	165	1.27609150D-02
22	0.0	70	0.0	118	1.08044510D-02	166	1.63009850D-02
23	0.0	71	0.0	119	-1.12869800D-02	167	-1.39409360D-02
24	0.0	72	0.0	120	2.01349670D-02	168	2.36004720D-03
25	0.0	73	0.0	121	2.55395820D-03	169	6.19402910D-03
26	0.0	74	0.0	122	2.01349670D-02	170	3.78045940D-02
27	0.0	75	0.0	123	1.89549440D-02	171	2.44514780D-02
28	0.0	76	0.0	124	3.40759060D-02	172	2.01349670D-02
29	0.0	77	0.0	125	2.59254130D-02	173	3.83398180D-03
30	0.0	78	1.37578660D-01	126	4.33490940D-02	174	1.63009850D-02
31	1.59905320D-01	79	-2.96035890D-02	127	-3.74733090D-02	175	6.19402910D-03
32	-5.67074210D-03	80	-7.05449700D-02	128	-2.35323710D-02	176	1.63009850D-02
33	-1.26404750D-01	81	4.56214060D-02	129	2.63115250D-02	177	2.71054360D-02
34	1.54629760D-02	82	0.0	130	-3.69524410D-02	178	1.63009850D-02
35	0.0	83	-1.14493690D-02	131	3.86193730D-02	179	2.01349670D-02
36	-1.93043380D-02	84	2.00624950D-03	132	2.01349670D-02	180	2.36004720D-03
37	-2.18335680D-02	85	3.34568970D-02	133	5.01400540D-03	181	2.01349670D-02
38	-1.44825160D-01	86	-3.99020910D-03	134	4.10453750D-02	182	2.36004720D-03
39	-5.75566190D-02	87	-6.64292920D-02	135	2.59254130D-02	183	2.01349670D-02
40	-3.74123910D-02	88	-5.19146570D-02	136	2.68115250D-02	184	3.78045940D-02
41	1.14658010D-02	89	-5.20065940D-02	137	2.59254130D-02	185	2.44514760D-02
42	1.62506920D-02	90	-2.82448140D-02	138	3.40759060D-02	186	3.78045940D-02
43	3.43143120D-02	91	-1.87971930D-02	139	3.40759060D-02	187	3.78045940D-02
44	3.97016180D-02	92	-1.07471970D-02	140	4.07524640D-02	188	-5.60895580D-02
45	-7.44490430D-04	93	2.29006760D-02	141	7.27292380D-02	189	7.67173590D-02
46	-2.22355920D-02	94	-4.14590720D-02	142	1.99549440D-02	190	-2.19378220D-02
47	-1.40215750D-01	95	3.65790730D-02	143	7.27292380D-02	191	-1.46955510D-03
48	3.94734380D-02	96	-1.97680600D-04	144	3.83398180D-03	192	-1.17564410D-03

TABLE 2.5

IDENTIFICATION OF 50 INDEPENDENT VARIABLES
(MEAS.) USED FOR DETECTION DEMONSTRATION ALGORITHM

NDEX	DESCRIPTION OF MEASUREMENT	INDEX	DESCRIPTION OF MEASUREMENT
1	Rainfall During Hour x 10^2	26	Supply Temp. 12 Hr Avg(Dev from Nom)
2	Temperature During Hour	27	Boiler 1 Wind Box x 10^1 12 Hr Avg
3	Av. Rainfall During Past 12 hrs x 10^2	28	Boiler 1 Furnace x 10^2 12 Hr Avg
4	Av. Temp. During Past 12 hrs	29	Avg. Fuel Temp.
5	Δ Boiler 1 H_20 Flow	30	Gal Oil Used in Atom Boiler
6	Boiler 1 Operating Pressure	31	Days Since PM On No.1 Blr
7	Oil Temperature Inlet	32	Days Since EPM On No.1 FD Fan
8	Used Gal. of Fuel Oil	33	Days Since PM On $C0_2$ Recdr 2
9	Steam Pressure A	34	Days Since PM On No.1 F.O. Pump
10	Steam Pressure B	35	Days Since EPM On No.1 F.O. Pump
11	Δ Zone Flow Zone 1	36	Days Since PM On Pump LTW 1
12	Δ Zone Flow Zone 2	37	Days Since PM On Pump LTW 2
13	Δ Zone Flow Zone 3	38	Days Since EPM On Makeup Feed Pump
14	Supply Temp. (Deviation from Nominal)	39	Days Since EPM On Chem Pump 1
15	Return Temp Zone 1	40	Days Since PM On Sump Pump 1
16	Return Temp Zone 2	41	Days Since PM On Sump Pump 2
17	Boiler 1 Outlet x 10^2	42	Days Since EPM On Sump Pump 2
18	Oil Inlet Temp. 12 Hr. Avg.	43	Days Since PM On HTHW Lines Zone 1
19	Oil Gallons Used 12 Hr. Avg.	44	Days Since PM On HTHW Lines Zone 2
20	Steam Pressure A 12 Hr. Avg.	45	Days Since PM On Cooling TWR.
21	Steam Pressure B 12 Hr. Avg.	46	Days Since EPM On Cooling TWR.
22	Δ Zone Flow Zone 1 12 Hr. Avg.	47	Days Since EPM On No. 3 Boiler
23	Δ Zone Flow Zone 2 12 Hr. Avg.	48	Days Since PM On Makeup Pump Atm Blr A
24	Δ Zone Flow Zone 3 12 Hr. Avg.	49	Days Since PM On Portable Blr
25	Supply Pressure 12 Hr. Avg.	50	Rainfall During Last 3 Days x 10^2

TABLE 2.6

IDENTIFICATION OF INDEX NOS. FOR KSC HEAT PLANT DATA VECTORS

INDEX	DESCRIPTION OF MEASUREMENT	INDEX	DESCRIPTION OF MEASUREMENT
1	Day of Year	38	Δ Zone Flow Zone 2
2	Hour of Day x 10^{-1}	39	Δ Zone Flow Zone 3
3	Rainfall During Hour x 10^2	40	Supply Pressure
4	Baro Pressure during Hour	41	Return Pressure Zone 1
5	Temperature During Hour	42	Return Pressure Zone 2
6	Av. Rainfall During Past 12 hrs x 10^2	43	Return Pressure Zone 3
7	Av. B/P During Past 12 hrs x10^2	44	Supply Temp.(Deviation from Nominal, Index 192)
8	Av. Temp. During Past 12 hrs	45	Return Temp Zone 1
9	Days Since Apr. 1 1971	46	Return Temp Zone 2
10	Δ Boiler 1 H_2O Flow	47	Return Temp Zone 3
11	Boiler 1 Operating Pressure	48	Boiler 1 Wind Box x 10^1
12	H_2O Inlet Temperature	49	Boiler 1 Furnace x 10^2
13	H_2O Outlet Temp. (Deviation from nominal, Index 191)	50	Boiler 1 Outlet x 10^2
14	Fuel Oil Pressure	51	Boiler 2 W/B x 10^1
15	Fuel Oil Temperature	52	Boiler 2 Furnace x 10^2
16	Steam Pressure	53	Boiler 2 Outlet x 10^2
17	Δ Boiler 2 H_2O Flow	54	Boiler 3 W/B x 10^1
18	Boiler 2 Operating Pressure	55	Boiler 3 Furnace x 10^2
19	H_2O Inlet Temperature	56	Boiler 3 Outlet x 10^2
20	H_2O Outlet Temp. (Deviation from nominal, Index 191)	57	Δ Boiler 1 H_2O Flow 12 Hr Avg
21	Fuel Oil Pressure	58	Operating Pressure 12 Hr Avg
22	Fuel Oil Temperature	59	H_2O Inlet Temp. 12 Hr Avg
23	Steam Pressure	60	H_2O Outlet Temp. 12 Hr Avg (Dev. from Index 191)
24	Δ Boiler 3 H_2O Flow	61	Fuel Oil Pressure 12 Hr Avg
25	Boiler 3 Operating Pressure	62	Fuel Oil Temp. 12 Hr Avg.
26	H_2O Inlet Temperature	63	Steam Pressure 12 Hr Avg.
27	H_2O Outlet " (deviation from Nominal, Index 191)	64	Δ Boiler 2 H_2O Flow 12 Hr Avg
28	Fuel Oil Pressure	65	Operating Pressure 12 Hr Avg.
29	Fuel Oil Temperature	66	H_2O Inlet Temperature 12 Hr Avg
30	Steam Pressure	67	H_2O Outlet Temp. 12 Hr. Avg.(Dev. from Index 191)
31	Oil Temperature Inlet	68	Fuel Oil Pressure 12 Hr. Avg.
32	Oil Temperature Outlet	69	Fuel Oil Temp. 12 Hr. Avg
33	Oil Discharge Pressure	70	Steam Pressure 12 Hr. Avg.
34	Used Gal. of Fuel Oil	71	Δ Boiler 3 H_2O Flow 12 Hr. Avg.
35	Steam Pressure A	72	Operating Pressure 12 Hr Avg.
36	Steam Pressure B	73	H_2O Inlet Temperature 12 Hr. Avg.
37	Δ Zone Flow Zone 1	74	H_2O Outlet Temperature 12 Hr. Avg.(Dev. from Index 191)

(cont'd Page 2)

IDENTIFICATION OF INDEX NOS. FOR KSC HEAT PLANT DATA VECTORS

INDEX	DESCRIPTION OF MEASUREMENT	
75	Fuel Oil Pressure 12 Hr. Avg.	
76	Fuel Oil Temp. 12 Hr. Avg.	
77	Steam Pressure 12 Hr. Avg.	
78	Oil Inlet Temp. 12 Hr. Avg.	
79	Oil Outlet Temp. 12 Hr. Avg.	
80	Oil Discharge Pressure 12 Hr. Avg.	
81	Oil Gallons Used 12 Hr. Avg.	
82	Steam Pressure 1 12 Hr. Avg.	
83	Steam Pressure 2 12 Hr. Avg.	
84	Δ Zone Flow Zone 1 12 Hr. Avg.	
85	Δ Zone Flow Zone 2 12 Hr. Avg.	
86	Δ Zone Flow Zone 3 12 Hr. Avg.	
87	Supply Pressure 12 Hr. Avg.	
88	Return Pressure Zone 1 12 Hr Avg	
89	Return Pressure Zone 2 12 Hr Avg	
90	Return Pressure Zone3 12 Hr Avg	
91	Supply Temperature 12 Hr Avg(Dev. from 192)	
92	Return Temp Zone 1 12 Hr Avg	
93	Return Temp Zone 2 12 Hr Avg	
94	Return Temp Zone 3 12 Hr Avg	
95	Boiler A Wind Box x 10^1 12 Hr Avg	
96	Boiler A Furnace x 10^2 12 Hr Avg	
97	Boiler A Outlet x 10^2 12 Hr Avg	
98	Boiler B Wind Box x 10^1 12 Hr Avg	
99	Boiler B Furnace x 10^2 12 Hr Avg	
100	Boiler B Outlet 12 Hr Avg	
101	Boiler C W/B 12 Hr Avg x 10^1	
102	Boiler C Furnace 12 Hr Avg x 10^2	
103	Boiler C Outlet 12 Hr Avg x 10^2	
104	BTU Output	
105	Make up Feed Meter x 10^{-1}	
106	Soft H_20 Meter x 10^{-1}	
107	Gal Oil Used in Atom Boiler	
108	Days Since PM H01-001 (AEX)	No. 1 Blr
109	Days Since PM H01-002 (AEY)	No. 1 FD Fan
110	Days Since PM H01-002 (E34)	No. 1 FD Fan
111	Days Since PM HO2-002(AEY)	No. 2 FD Fan

INDEX	DESCRIPTION OF MEASUREMENT	
112	Days Since PM HO2-002(AEY)	No. 2 FD Fan
113	Days Since PM HO2-002(E34)	No. 2 FD Fan
114	Days Since PM HO4-001(AFJ)	$C0_2$ Recdr 1
115	Days Since PM HO4-002(AFS)	$C0_2$ Recdr 2
116	Days Since PM HO5-001(AFK)	FO Storage tank
117	Days Since PM HO5-002(AK)	Fuel Oil Oper. Tank
118	Days Since PM HO5-003(AFL)	No. 1 F. O. Pump
119	Days Since PM HO5-003(E34)	No. 1 F. O. Pump
120	Days Since PM HO5-004(AFL)	No. 2 F. O. Pump
121	Days Since PM HO5-004(E34)	No. 2 F. O. Pump
122	Days Since PM HO6-001(AFN)	No. 1 Blr Circ Pump
123	Days Since PM HO6-001(E34)	No. 1 Blr Circ Pump
124	Days Since PM HO6-002(AFN)	No. 2 Blr Circ Pump
125	Days Since PM HO6-002((E34)	No. 2 Blr Circ Pump
126	Days Since PM HO6-003(AFN)	No. 3 Blr Circ Pump
127	Days Since PM HO6-003(E34)	No. 3 Blr Circ Pump
128	Days Since PM HO6-004(AFP)	Meter Flow No. 1
129	Days Since PM HO6-005(AFP)	Meter Flow No. 2
130	Days Since PM HO6-005(AGG)	Tank Exp #1 Blr
131	Days Since PM HO6-005(ANG)	Tank Exp #2 Blr
132	Days Since PM HO7-001(AFN)	Pump Dist #1
133	Days Since PM HO7-001(E34)	Pump Dist #1
134	Days Since PM HO7-002(AFN)	Pump Dist #2
135	Days Since PM HO7-002(E34)	Pump Dist #2
136	Days Since PM HO7-003(HFN)	Pump Dist #3
137	Days Since PM HO7-003(E34)	Pump Dist #3
138	Days Since PM HO7-004(AFP)	Meter Flow Zone 1
139	Days Since PM HO7-005(AFP)	Meter Flow Zone 2
140	Days Since PM HO7-005(AFP)	Meter Flow Zone 3
141	Days Since PM HO8-001(M25)	Pump LTW 1
142	Days Since PM HO8-001(E34)	Pump LTW 1
143	Days Since PM HO8-002(M25)	Pump LTW 2
144	Days Since PM HO8-002(E34)	Pump LTW 2
145	Days Since PM HO9-001(AHD)	Makeup Feed Pump
146	Days Since PM HO9-001(E43)	Makeup Feed Pump
147	Days Since PM HO9-002(AHD)	Makeup Feed Pump
148	Days Since PM HO9-002(E34)	Makeup Feed Pump

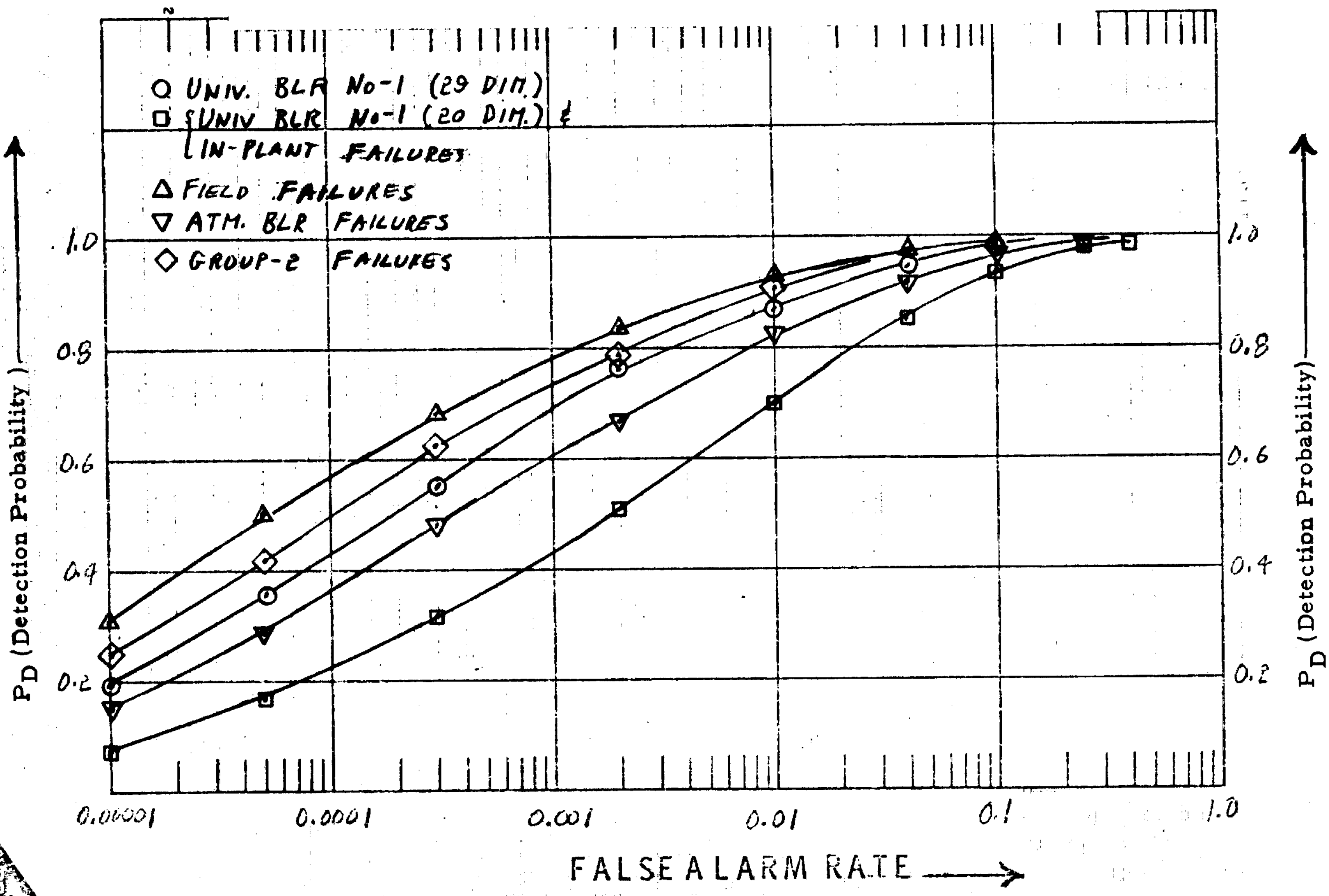
(cont'd Page 3)

IDENTIFICATION OF INDEX NOS. FOR KSC HEAT PLANT DATA VECTORS

INDEX	DESCRIPTION OF MEASUREMENT		INDEX	DESCRIPTION OF MEASUREMENT
149	Days Since PM HO9-003(E34)	Energy Makeup Feed Pump	191	Nominal Blr Outlet Temp (375 Winter, 350 Summer)
150	Days Since PM HO9-004(AHF)	Chem. Feed Pump 1	192	Nominal Supply Temp (350 Winter, 330 Summer)
151	Days Since PM HO9-004(E01)	Chem. Feed Pump 1		
152	Days Since PM HO9-005(AHF)	Chem. Feed Pump 2		
153	Days Since PM HO9-005(E01)	Chem. Feed Pump 2		
154	Days Since PM HO9-006(AHC)	Agitator		
155	Days Since PM HO9-007(AFK)	Treated H_20 Tank		
156	Days Since PM HO9-008(AAB)	WTR Softener		
157	Days Since PM H10-001(M04)	Sump Pump 1		
158	Days Since PM H10-001(E34)	" " "		
159	Days Since PM H10-002(M04)	Sump Pump 2		
160	Days Since PM H10-002(E34)	Sump Pump 2		
161	Days Since PM H44-001(AHR)	HTHW Lines Zone 1		
162	Days Since PM H45-001(AHS)	HTHW Lines Zone 2		
163	Days Since PM H46-001(AHT)	HTHW Lines Zone 3A		
164	Days Since PM H46-002(AHU)	HTHW Lines Zone 3B		
165	Days Since PM H46-003(AHV)	HTHW Lines Zone 3C		
166	Days Since PM H47-001(AHX)	Cooling Tower		
167	Days Since PM H47-001(AHY)	Cooling Tower		
168	Days Since PM H47-002(M25)	Circular Pump Cooling H_20		
169	Days Since PM H47-002(E34)	Circular Pump Cooling H_20		
170	Days Since PM H48-001(AJC)	Portable Generator		
171	Days Since PM H48-001(E02)	Portable Generator		
172	Days Since PM H49-001(AEX)	No. 3 Boiler		
173	Days Since PM H49-001(E34)	No. 3 Boiler		
174	Days Since PM H49-002(M25)	No. 3 Boiler Circ. Pump		
175	Days Since PM H49-002(E34)	No. 3 Boiler Circ. Pump		
176	Days Since PM H50-001(AGW)	No. 1 Steam Atom Boiler		
177	Days Since PM H50-001(E34)	No. 1 Steam Atom Boiler		
178	Days Since PM H50-002(AGW)	No. 2 Steam Atom Boiler		
179	Days Since PM H50-002(E34)	No. 2 Steam Atom Boiler		
180	Days Since PM H50-003(M25)	No. 1 Makeup Feed Pump Steam Atom Boiler		
181	Days Since PM H50-003(E34)	No. 1 Makeup Feed Pump Steam tom Boiler		
182	Days Since PM H50-004(M25)	No. 2 Makeup Feed Pump Steam Atom Boiler		
183	Days Since PM H50-004(E34)	No. 2 Makeup Feed Pump Steam Atom Boiler		
184	Days Since PM H51-001(AJC)	100 HP Boiler Mobile		
185	Days Since PM H51-001(E02)	100 HP Boiler Mobile		
186	Days Since PM H52-001(AJC)	Portable Boiler SN2544		
187	Days Since PM H53-001(AJC)	Portable Boiler SN2544		
188	Rainfall During Last 3 Days x 10^2			
189	Rainfall During Last 5 days x 10^2			
190	Rainfall During Last 10 days x 10^2			

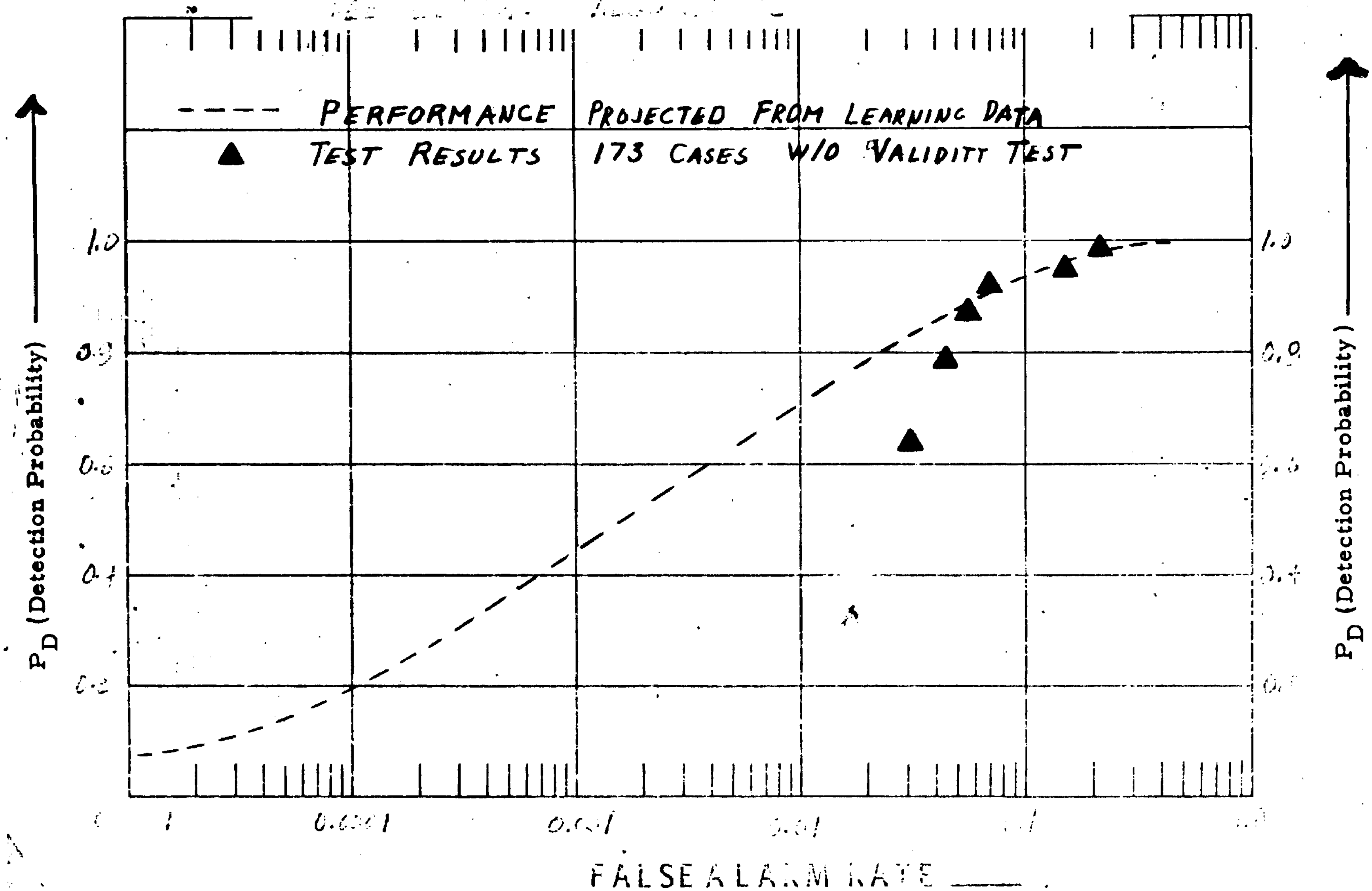
C

FIGURE 2.1
PROJECTED CLASSIFICATION PERFORMANCE TRADE-OFF CURVES FOR CANDIDATE DETECTION ALGORITHMS



NOTE: USE TYPE B PENCIL FOR VUGRAPHS AND REPORT DATA.

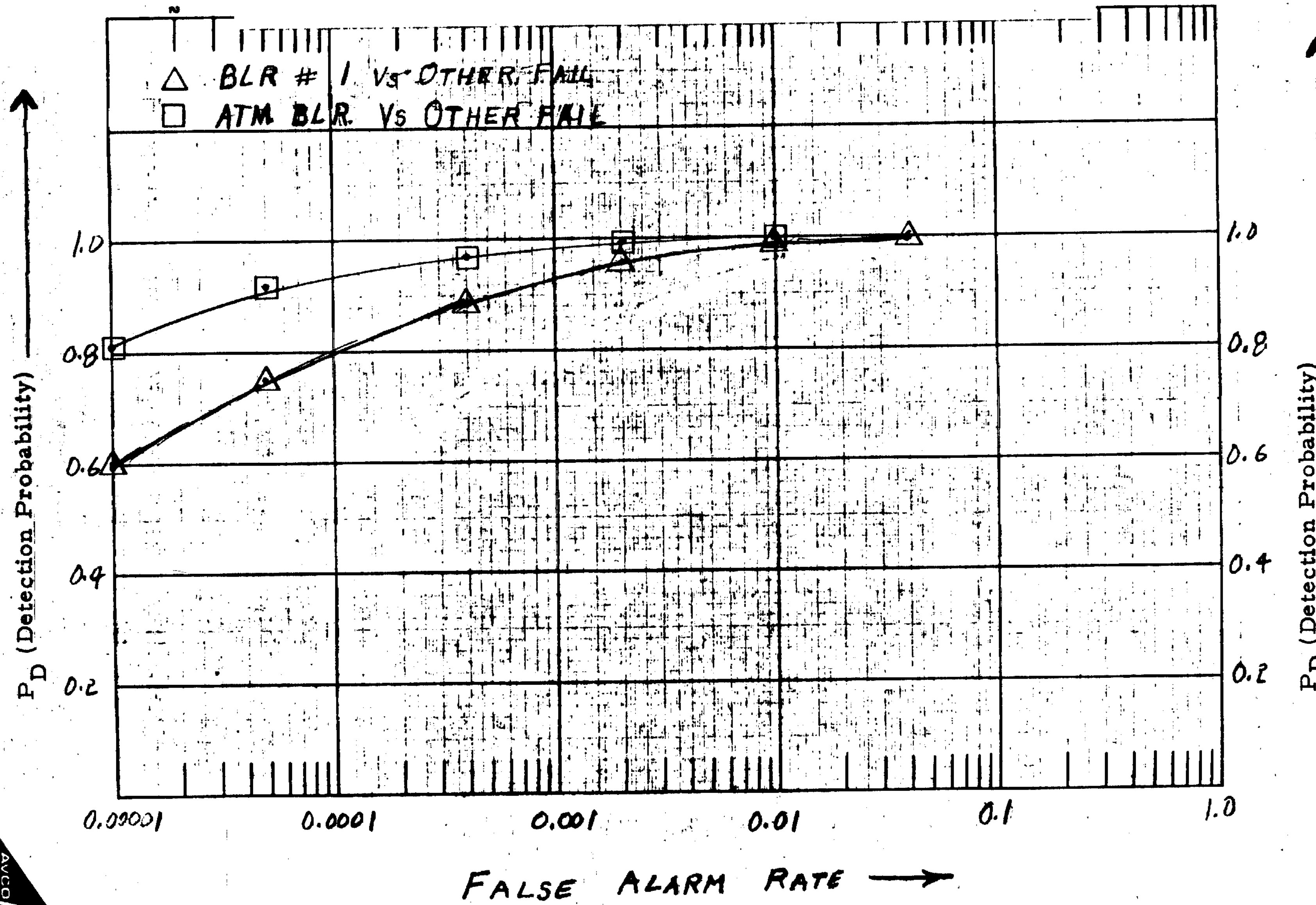
FIGURE 2.2
COMPARISON OF CLASSIFICATION PERFORMANCE TRADE-OFF CURVES FOR PROJECTED AND TEST PERFORMANCE FOR 20 DIMENSIONAL BOILER NO. 1 DETECTION ALGORITHMS



NOTE: USE TYPE B PENCIL FOR VUGRAPHS AND REPORT DATA.

FIGURE 2.3

CLASSIFICATION PERFORMANCE TRADE-OFF CURVES FOR DIAGNOSIS OF BOILER NO. 1 AND ATOMIZING STEAM BOILER FAILURES



NOTE: USE TYPE B PENCIL FOR VUGRAPHS AND REPORT DATA.

FIGURE 2.4

COMPARISON OF ESTIMATED AND ACTUAL TIME TO FAILURE FOR THE ATOMIZING STEAM BOILER

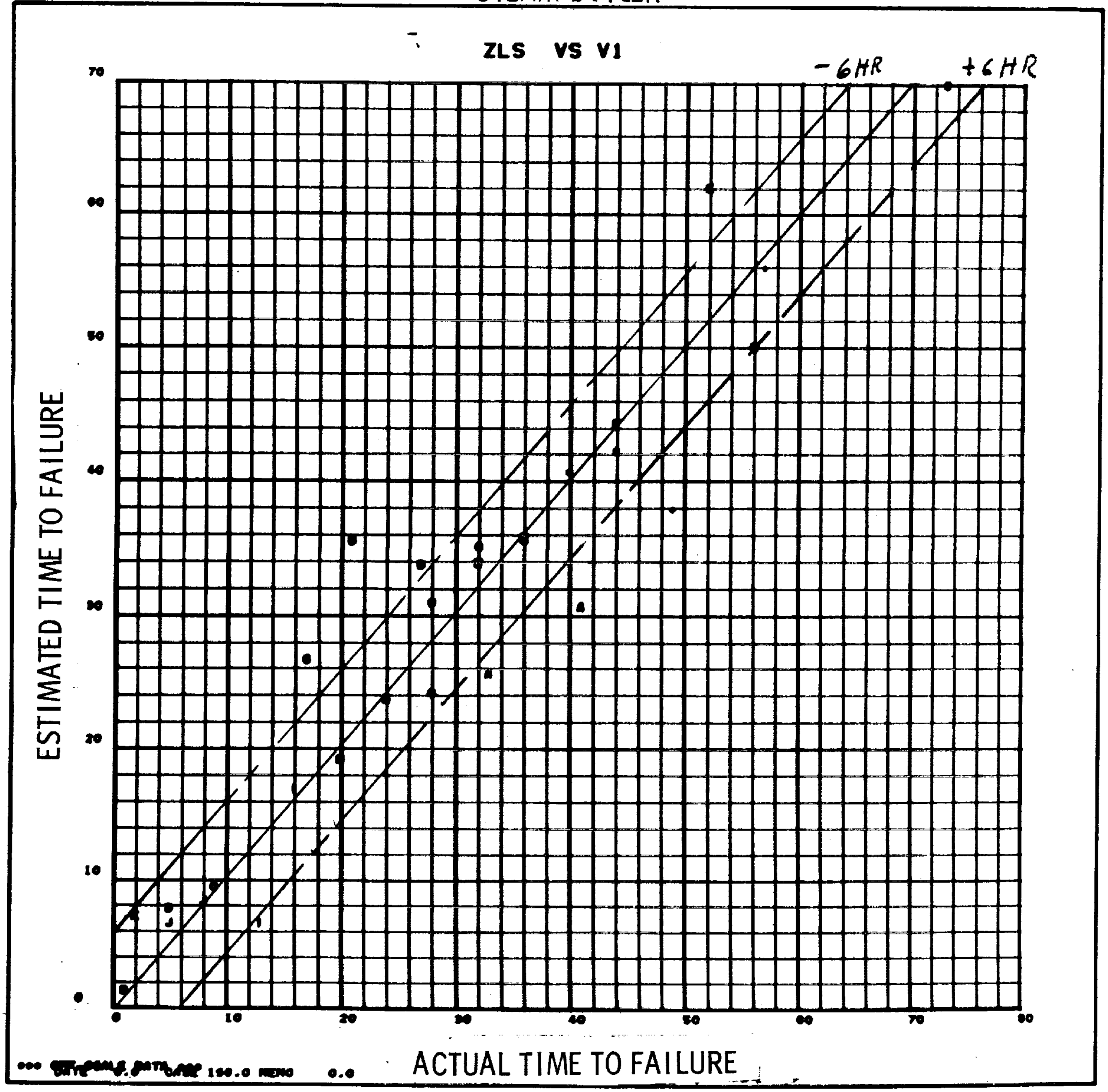
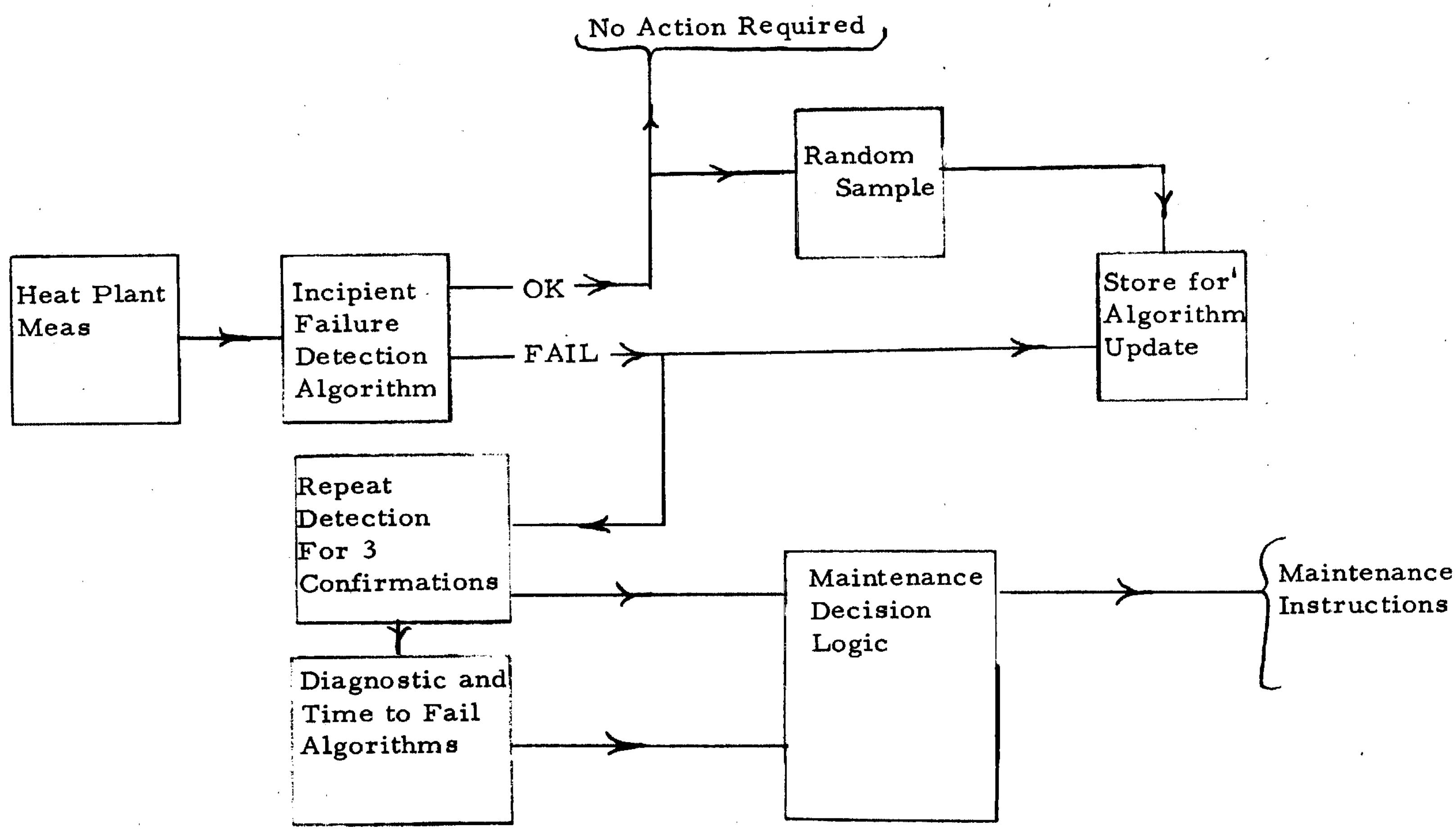


FIGURE 2.5 RECOMMENDED USE OF PREDICTIVE MAINTENANCE ALGORITHMS



3.0 APPLICATION OF MAINTENANCE ALGORITHMS TO KSC CENTRAL HEAT PLANT

Before presenting the details of the development of the maintenance algorithms, it is necessary to review the details of the KSC central heat plant and how the maintenance algorithms can be used to assist in the maintenance of this system. The next section will summarize the KSC heating plant. This will be followed by the discussion of how the ADAPT derived maintenance algorithms can be used to improve the maintenance of the current system without any modifications and then how the same types of maintenance algorithms could be used in conjunction with an automated data monitoring system to completely automate the detection and diagnosis of out of tolerance performance of the KSC heating plant.

3.1 Description of the KSC Heat Plant

Figure 3.1 presents a schematic diagram showing some of the key features of the KSC central heat plant. Although this figure does not include many of the components of the system and therefore does not adequately present the complexity of the system, it does illustrate the large amount of redundancy which exists in this heating system. It will serve as a basis for describing the application of a maintenance algorithm. For a detailed analysis of the results, it will be necessary for the reader to refer to layout drawings of the entire system.

The heat plant is basically composed of three boilers, any two of which are sufficient to carry the full load of all three zones. Boilers No. 1 and 2 are identical, and Boiler No. 3 is a different and smaller boiler. Normal operation calls for atomizing the fuel using a steam atomizer gun with the atomizing steam supplied by either one of the two atomizing steam boilers. In general, the pumps in the system have been placed such that a pump failure can be compensated for by valving out the disabled pump and allowing the other pumps to carry the load.

Figure 3.2 presents a map of the buildings which are supplied hot water by the central heat plant. This figure also shows many details of the distribution system as of mid 1971. Major changes were made in the distribution system at the end of the first quarter of 1970 and again in August 24, 1971, when flight crew training building was moved from zone 2 to zone 3. In addition, other minor changes were made periodically during the period in which the data for this study was obtained. Although the maintenance algorithms are relatively insensitive to these changes, the ability to simply update the algorithm offers an attractive solution to the problem which will be discussed further in Section 5.

3.2 Role of Maintenance Algorithms

The maintenance problem of this system is greatly simplified by its redundancy. However, it is still desirable to have prior knowledge of an impending failure and to perform the maintenance prior to the occurrence of the failure. This has the dual advantages of allowing the failing component to be removed from the system prior to doing more damage to other components in the system, and also eliminates the possibility of the failure creating an inconvenience to the user through leakage or other damage which may occur between occurrence and time of discovery of the failure.

The current approach to this problem is to perform preventive maintenance on a schedule which has been designed to minimize the occurrence of failures in the system. Although this system is far superior to simply waiting until the failure occurs and then repairing the failure, it has several major drawbacks which include: 1) a relatively high cost associated with performing the required inspections, 2) opportunity for additional faults to be introduced during the inspection process itself, and 3) the continued possibility of catastrophic failures occurring because the inspection process either did not result in detection of an impending failure or the failure developed too rapidly to be detected in the normal preventive maintenance cycle.

The present study seeks to demonstrate the feasibility of a new approach to maintenance of complicated systems based on the concept of monitoring the performance of the system continually to detect incipient out of tolerance performance so that corrective action can be initiated prior to the occurrence of the failure. Clearly, this approach should be used in conjunction with a preventive maintenance (PM) program to further reduce the number of catastrophic failures. This converts the classical preventive maintenance system to a demand preventive maintenance (DPM) system where preventive maintenance is performed prior to failure, but on only when required. The key question of feasibility is the ability to detect incipient failure sufficiently prior to the occurrence of the failure that the actual failure can be prevented. If such algorithms can be derived then their application will eliminate the requirement for disassembling the equipment to perform the inspection and thus both reduce maintenance costs and eliminate the possibility of introducing additional faults into the system during an unnecessary inspection. In addition, the capability to detect incipient failures allows one to correct the failure before it occurs even if it occurs too rapidly to be detected by a standard PM program. This will prevent further damage and its associated repair cost to both the heating system itself and/or to the customers facilities.

Three types of algorithms would be useful for implementing a maintenance system such as this. The most critical algorithm which must be developed is one to detect incipient failures. This algorithm provides the basic information which is required to implement DPM scheme. If one can detect in advance that this system is near failure, then one can initiate the appropriate corrective action. However, this task is greatly simplified if the measurements can also supply the information which is required to diagnosis where the impending failure will occur.

This will be accomplished by a second group of algorithms which will be applied after the detection has been accomplished and will re-examine the measurements and diagnose the impending failure. This set of algorithms would then define to the maintenance personnel what component must be removed from the system and overhauled to prevent the failure from occurring. This time the failure algorithm would allow the scheduling of maintenance for those failures occurring sufficiently far in the future.

The demonstration of the feasibility of applying ADAPT derived algorithms to the maintenance can be achieved if one shows that it is possible to derive an algorithm for detecting out of tolerance performance of the system. Thus, the major thrust of the present study was to investigate detection algorithms, select a detection algorithm for verification and demonstrate through independent test cases that such an algorithm can be derived for a system such as the KSC central heat plant. In addition, the expected potential performance was determined as a function of the complexity of the algorithm. The feasibility of diagnostic and time to failure algorithms was also shown by deriving demonstration diagnostic algorithms and time to failure algorithms, and projecting the potential performance of these algorithms. Since these algorithms are far less critical to the feasibility and since the performance projection has in the past proved quite indicative of the actual performance, detail proof testing with independent test data was only carried out for the detection algorithms.

Examination of Figure 3.1 shows that there are many operating options or configuration in which the KSC central heat plant can operate. The major variation in the system is probably due to different boiler combinations. Thus, for the feasibility study, it was decided to limit the investigation to consideration of only one boiler operating configuration. Since the feasibility of detecting out of tolerance behavior was demonstrated for this condition, it follows that the other boiler configurations would also be amenable to this approach. Boiler No. 3 is significantly smaller than either Boiler No. 1 or Boiler No. 2, and it is the only boiler which is too small to operate by itself. Thus, the number of combinations of boiler operations which must be considered for this particular system would be six. These six are: 1) Boiler No. 1 operating by itself, 2) Boiler No. 2 operating by itself, 3) Boiler No. 1 operating with Boiler No. 3, 4) Boiler No. 2 operating with Boiler No. 3, 5) Boilers No. 1 and 2 operating together and 6) All three boilers operating together. The configuration selected for this study is indicated by the cross hatched component shown in Fig. 3.1. This configuration allows any of the components of the system to be operating with the exception of Boilers No. 2 and 3. Clearly, this still leaves a great deal of variation in the system configuration and, if it were impossible to develop successful algorithms with this general configuration, one could still consider further reducing the number of options. However, the analysis showed that it was feasible to develop the algorithm with all of the other variations included and thus this approach was not pursued. The details of this decision will be discussed further in Section 5. 2.

3.3 Applications of Maintenance Algorithms with Present Central Heat Plant

The role of the maintenance algorithm in maintaining a system such as the KSC central heat plant can be seen by examining how these algorithms could be used with the present heat plant system. As part of the normal operation of the KSC central heat plant, a maintenance log is kept in which pressures, temperatures, and other pertinent measurements of the central heat plant are recorded hourly. Figure 3.3 shows a typical log sheet recorded for August 14, 1970. The ADAPT algorithms have been derived to make use of this data which is recorded on the log sheet as well as other pertinent data which might effect the operation of the central heat plant such as the current weather information of key items in the system. When the proposed maintenance system is implemented the time since maintenance of key items will change since the preventive maintenance will no longer be performed on a routine schedule. However, the date at which any maintenance is performed on every key item of the system should still be recorded and the time since this maintenance used where the days since PM variables appears in the data vector. There are two differences which could affect the performance of the algorithm. The first is that various groupings of maintenance items are no longer correlated and the second is that one would expect considerable increases in time between maintenance on given items. The first of these will not significantly affect the performance of the algorithm. If it has an affect it would be to make it more difficult to observe this affect when deriving the algorithm. However, this has already been accomplished and thus this aspect in the change of the character of the maintenance is insignificant. The increased length of time between maintenance of items can be significant as it results in an extrapolation of the affect of this maintenance. It will only become significant at long times and when it is significant the ADAPT validity criteria may detect the problem occurring. However, even in this case the algorithm updating procedures which have been suggested will account for the difficulty. The worst situation that could result from this change in the way the maintenance is done is that at some time period, long compared to the normal maintenance cycle, the performance of the algorithm could be slightly degraded and this degraded performance would exist until the first update of the algorithm. It should be emphasized that the degradation should be very slight since it will only occur in a relatively small number of variables which in aggregate make a small contribution to the decision.

Thus, the records that are now kept provide an excellent basis for beginning the maintenance of the KSC central heat plant with or without modification of the heat plant, the data gathering system or acquiring any new data processing equipment. The procedure consists of: 1) taking the data which is now recorded on this measurement log for the given hour during the day for which the system is being evaluated, 2) combining this with the weather and other pertinent data, 3) punching this data, and 4) processing it in a general purpose computer. This process is illustrated in Figure 3.4.

The ADAPT PPM algorithms would be stored in the computer and used to calculate a number which was indicative of the health of the central heat plant and by comparing this number to a pre-determined threshold decide if a failure would occur. If a failure is expected, the ADAPT diagnostic programs would be used in the same general purpose computer to determine where the failure will occur. These diagnostic programs would include the ADAPT diagnostic algorithms and logic required to sequentially apply all the required algorithms and print out which component will fail. As indicated in Figure 3.4, this entire process could be combined into a single program such that when the data and ADAPT maintenance program were entered into the general purpose computer, the program automatically would have applied the ADAPT detection algorithm and, if this algorithm indicated that there was no problem with the system, the computer would simply be programmed to print out that all was well. If the ADAPT detection algorithm indicated that a failure was near, the program would continue to perform the diagnostics and print out the results of the diagnosis indicating the expected location of the failure.

Clearly, the preceding discussion has been oversimplified and several important decisions must still be made. For example, the number of times which the system is examined is a parameter which must be decided based on the performance of the algorithm. If the algorithm is capable of detecting failures one or more days in advance, it would appear reasonable to apply the algorithm only once a day if the data were manually collected and key-punched. The false alarm rate can be reduced by repeating the application of the algorithm at hourly intervals after the out-of-tolerance condition is first detected and corrective action initiated only in the event that a certain number of consecutive hourly applications of the algorithm yield agreement that the system is in danger of failure. In this mode of operation, a false alarm rate as high as one in ten could be tolerated for the first application and one in three for the successive applications. This will result in an overall false alarm rate of 1 in 300, or one false alarm per year and provide significantly improved detection. It will also allow the verification of the algorithm to be accomplished with a significantly smaller number of test cases. The penalty paid for this improvement is a requirement to apply the algorithm 3 extra times every two weeks. There are many trade-offs such as to the number of cases, the false alarm rate to be set into the algorithm, the role of the validity criteria and the time of day at which the system should be evaluated which should be considered. These decisions do not bear on the feasibility of implementing the system and merely provide additional flexibility to meet the needs of the user. For the approach to be feasible, one must be able to achieve a final false alarm rate of less than the order of 1 in 100 with a 75 to 90% detection probability. This performance may be achieved either through a single application of the algorithm or through a combination of appropriately established thresholds and repetitive application of the algorithm. Each of these modes of operation will be discussed in more detail in Section 5.2 after the performance of the reference algorithm has been derived. The key result of this discussion is that the maintenance algorithms may be applied to the present KSC central heat plant without any additional hardware if manual data collection and key punch is used.

3.4 Application with Automated Monitoring System

The possibility exists that the function of keeping the KSC central heat plant maintenance log will be assumed by an automated data monitoring system which would consist of appropriate sensors to take the data and transmit it to a central maintenance computer where the information would be recorded periodically. It will be useful to consider in some detail how the ADAPT algorithm could be incorporated in this system to completely automate the entire DPM process. The procedure is illustrated in Fig. 3.5. Comparison of Fig. 3.5 and Fig. 3.4 shows that the procedure for the automated monitoring is in principal very similar to that of the manual monitoring. The measurements required are the same as in the monitoring using the manual system. However, in the automated system the measurement of the values and transmission of the values measured to the computer will be accomplished by the automated detection system. Because of the simplicity of the ADAPT algorithms, the planned maintenance computer should easily have the capability to incorporate the ADAPT detection and diagnostic algorithms within it. Thus, the maintenance computer can contain as part of its normal function the ADAPT maintenance computer program which is illustrated in Fig. 3.6. This program would take the weather data, and hourly measurements of the central heat plant and process them through the detection algorithm. One attractive way of accomplishing this would be to perform this function once a day at some specified time. The algorithm threshold could be set for high false alarm rate, say the order of one and ten. If the algorithm detects that the system is not operating normally, it will initiate continued application of the maintenance computer program until three consecutive out-of-normal measurements spaced one hour apart are reached. Note this is the same procedure suggested for the manual operation in the proceeding section. Alternatively, with the automated system, the algorithm could simply be applied every hour and "n" consecutive failure indications required to initiate action. Assuming that the one hour interval between measurements are sufficient to make the cases independent and an individual false alarm rate of one in ten, this will result in effective false alarm rate of 10^{-n}.

In addition to initiating the consecutive decision logic, the detection of an out-of-normal condition will also initiate the collection and recording of the appropriate diagnostic data. When the prescribed out-of-normal indications are given by the ADAPT classification algorithms, the computer program will instruct itself to perform the diagnostics. The diagnostics will be performed by processing the diagnostic measurements though a series of algorithms to separate each failure mode from all other failure modes, separate each possible region in which a failure could occur from all other regions, perform a nearest neighbor analysis to determine the failure most like the one presently being determined, and if the specific failure mode was successfully identified to apply the time to failure algorithm to determine when the failure can be expected.

The results of all these diagnostic algorithms will then be processed through the decision logic to evaluate the answers obtained for all of the algorithms and for each set of data. Depending on how near the incipient failure is to failures which have occurred in the past, one or more of several answers are possible. One possibility is the prediction of the failure followed by the identification and ordering of the possible failure modes. If a specified failure was identified, an estimate of time-to-failure may be possible. This type of output would be expected for the more common failures where greater learning data is available to develop the algorithms. On the other hand, for the more rare failures, there may not be enough information to develop a time-to-failure algorithm or possibly even to positively identify the failure mode. In this case, the output might be simply an indication that there was an impending failure or even just that the system was operating in unusual mode. In both of these cases, the nearest neighbor algorithm and possibly the failure region algorithms would give some indications which could be used to provide clues as to where the failure should be expected.

Returning to Fig. 3.5, we see that the results of the application of this maintenance computer program to the data collected automatically on the central heat plant will produce the maintenance instructions indicating: 1) all is well, 2) that a failure will occur and the diagnostics of the failures, or 3) that the operation is unusual and some sort of prognosis concerning this operation. In addition to deriving the maintenance instructions, the maintenance computer would still be used to produce the maintenance log which is now produced by hand and monitor the alarms to indicate catastrophic failure. Alarms for such items as the boiler being out or temperature below the minimum are no different than the alarms which are currently used and are necessary because no maintenance system, either the current PM system or the ADAPT PPM approach will be perfect.

The final function of the maintenance computer would be to select on a prescribed basis (possibly utilizing the results of the detection algorithm) cases to be used to update the detection and diagnostic algorithms. This data would be stored on tape by the maintenance computer and periodically this tape would be removed and along with an algorithm update program processed through one of the existing general purpose computers at KSC to produce a new set of ADAPT maintenance algorithms. This procedure, although not absolutely essential is highly desirable since it will significantly improve the performance of the maintenance algorithms with a very small cost in additional complexity and processing. In addition, it provides the capabilities to account for the continual changes which occur in the KSC central heat plant and distribution system. This updating capability can be provided on any general purpose computer capable of inverting an approximately 20 by 20 matrix. The equations required to update the detection and diagnostic algorithms are given in Appendix E. It will be limited to accounting for changes in the system which do not modify the form of the data history (i.e. number of types of measurements which are used in the algorithm). Examples of acceptable changes are such things as changing the buildings on a given zone to another zone, minor changes in fuel oils, etc. An unacceptable change would be such as deleting a major component in the system such as the atomizing boilers. For this latter type of change it would be necessary to rederive the optimal base functions. This capability cannot be provided on a generalized "cookbook" basis. The derivation of the optimum function requires a detail analysis of the particular problem and is different for every problem considered.

In summary, the application of the ADAPT maintenance algorithms to the KSC central heat plant can be accomplished without the addition of any hardware in either its present configuration or in the configuration with an automated data monitoring system. In both cases, a certain amount of additional software primarily the incorporation of the ADAPT maintenance program in an appropriate computer is required. As a part of this study, Avco has supplied KSC with ADAPT algorithms and these algorithms have been implemented on general purpose computers at KSC. The feasibility of the entire system rests on the ability to obtain a detection algorithm which at least when applied successively over a fraction of a day will result in an acceptable detection probability with a false alarm rate of the order of one per year or better.

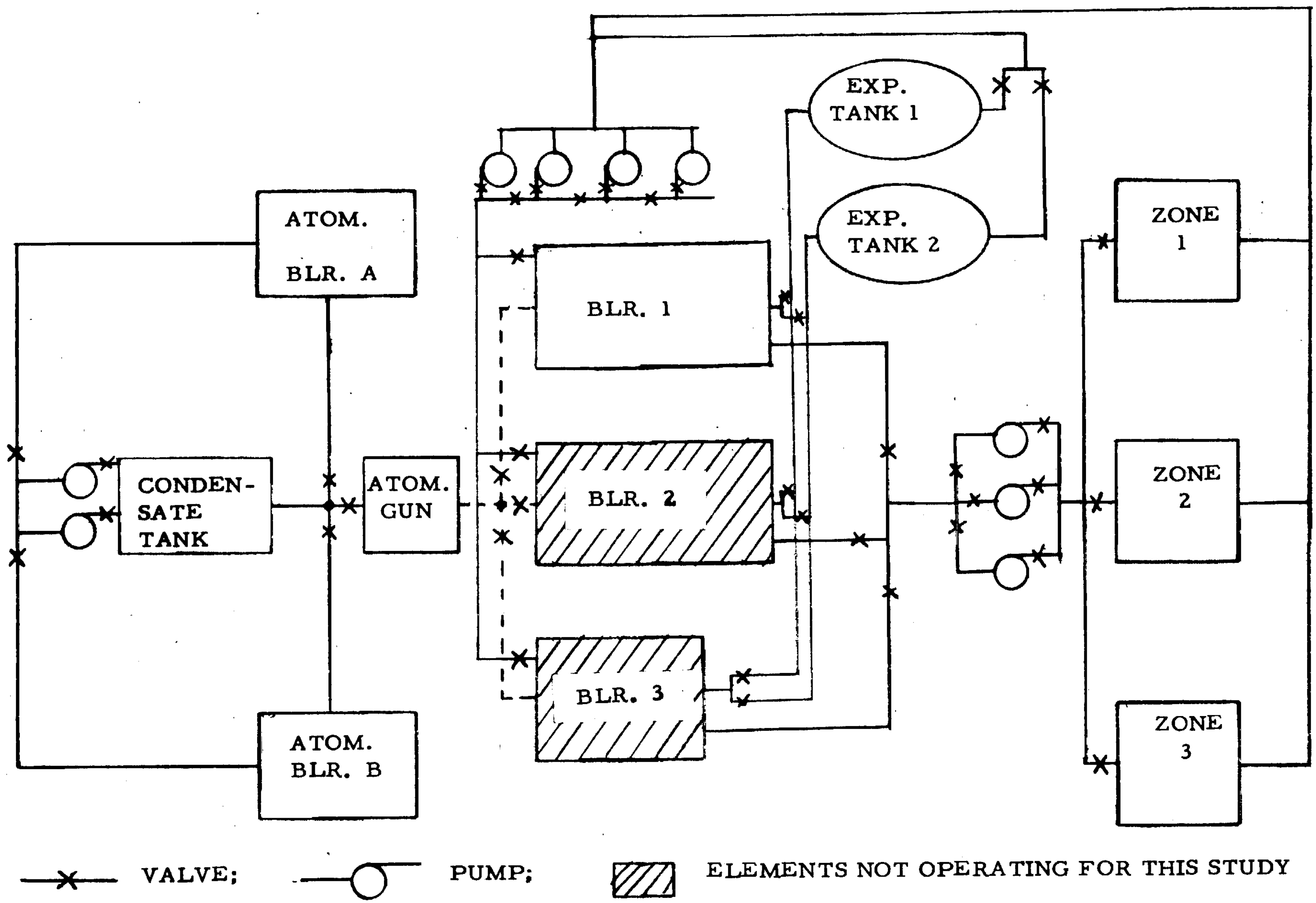
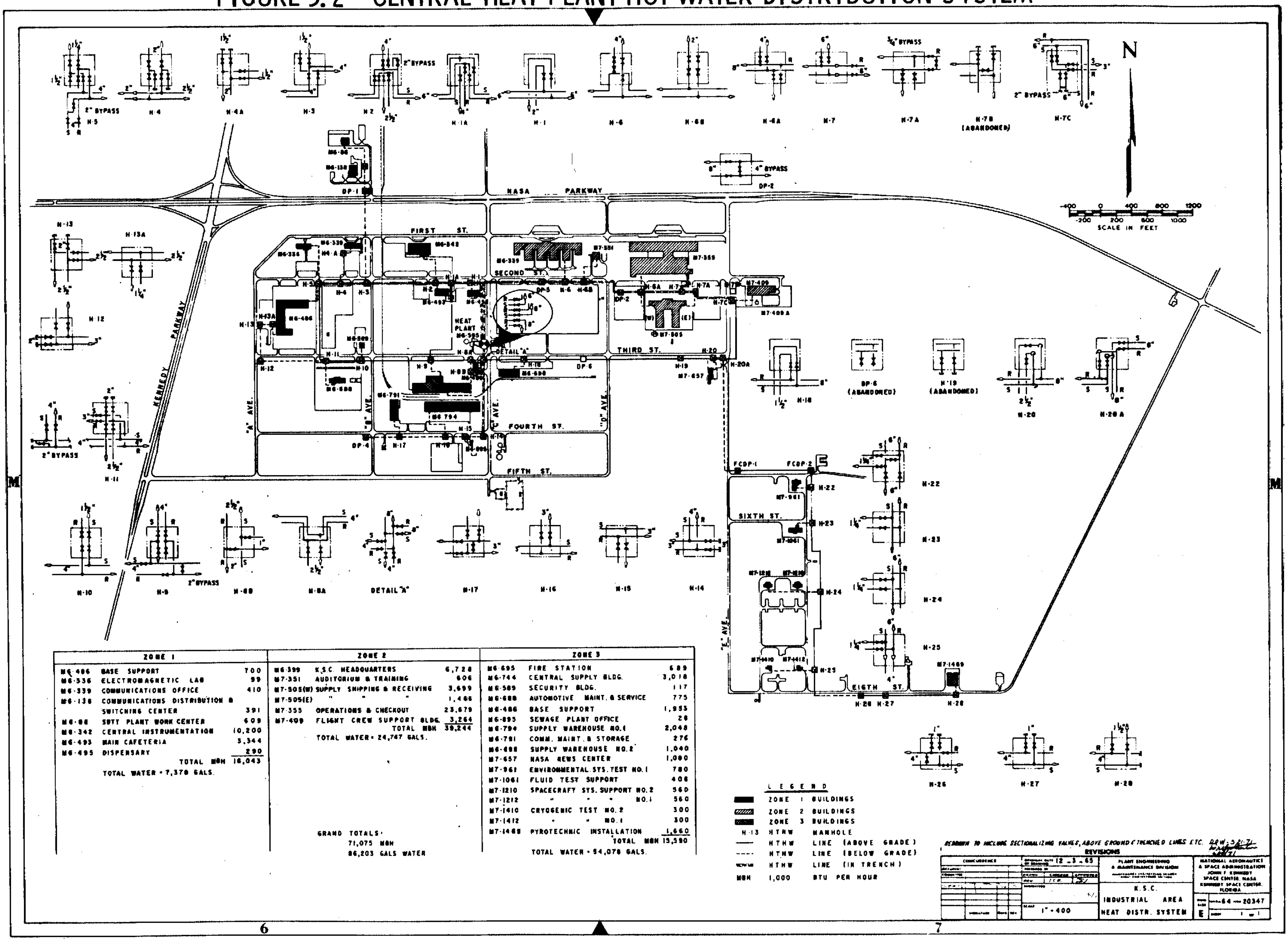


FIGURE 3.1

SCHEMATIC SKETCH OF KSC CENTRAL HEAT PLANT

FIGURE 3.2 - CENTRAL HEAT PLANT HOT WATER DISTRIBUTION SYSTEM



ZONE 1		
M6-486	BASE SUPPORT	700
M6-336	ELECTROMAGNETIC LAB	99
M6-339	COMMUNICATIONS OFFICE	410
M6-138	COMMUNICATIONS DISTRIBUTION & SWITCHING CENTER	391
M6-88	SBTT PLANT WORK CENTER	609
M6-342	CENTRAL INSTRUMENTATION	10,200
M6-493	MAIN CAFETERIA	3,344
M6-495	DISPENSARY	290
	TOTAL MBH	16,043
	TOTAL WATER - 7,378 GALS.	

ZONE 2		
M6-399	K.S.C. HEADQUARTERS	6,728
M7-351	AUDITORIUM & TRAINING	606
M7-505(W)	SUPPLY SHIPPING & RECEIVING	3,699
M7-505(E)	" " "	1,466
M7-355	OPERATIONS & CHECKOUT	23,679
M7-409	FLIGHT CREW SUPPORT BLDG.	3,264
	TOTAL MBH	39,244
	TOTAL WATER - 24,747 GALS.	

GRAND TOTALS-
71,075 MBH
86,203 GALS WATER

ZONE 3		
M6-695	FIRE STATION	689
M6-744	CENTRAL SUPPLY BLDG.	3,018
M6-589	SECURITY BLDG.	117
M6-688	AUTOMOTIVE MAINT. & SERVICE	775
M6-486	BASE SUPPORT	1,953
M6-895	SEWAGE PLANT OFFICE	28
M6-794	SUPPLY WAREHOUSE NO.1	2,048
M6-791	COMM. MAINT. & STORAGE	276
M6-698	SUPPLY WAREHOUSE NO.2	1,040
M7-657	NASA NEWS CENTER	1,080
M7-961	ENVIRONMENTAL SYS. TEST NO.1	780
M7-1061	FLUID TEST SUPPORT	408
M7-1210	SPACECRAFT SYS. SUPPORT NO.2	560
M7-1212	" " " NO.1	560
M7-1410	CRYOGENIC TEST NO.2	300
M7-1412	" " NO.1	300
M7-1468	PYROTECHNIC INSTALLATION	1,660
	TOTAL MBH	15,590
	TOTAL WATER - 54,078 GALS.	

FIGURE 3.3 - TYPICAL HEAT PLANT LOG SHEET (8/14/70)

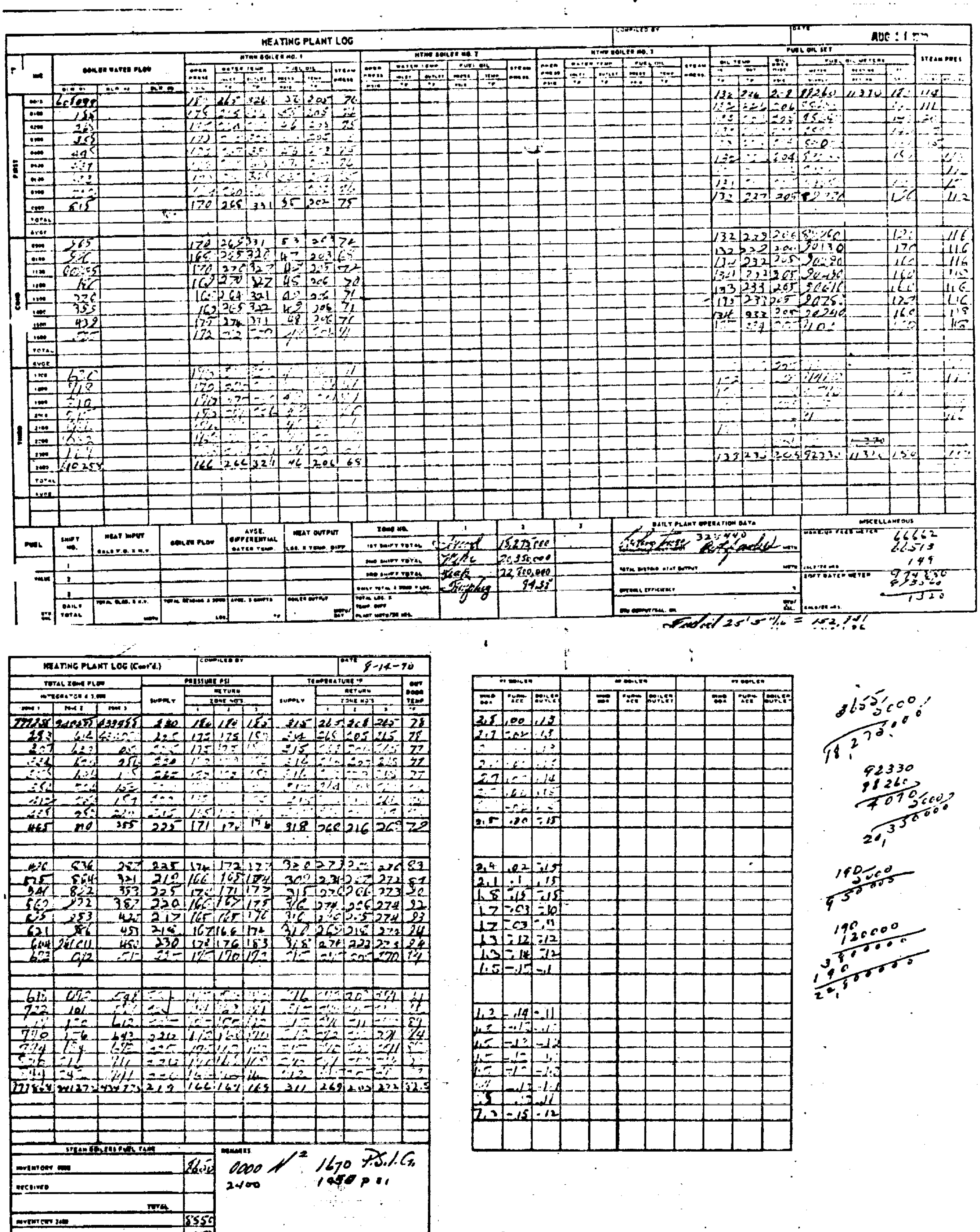


FIGURE 3.4

USE OF ADAPT MAINTENANCE ALGORITHMS
(PRESENT SYSTEM)

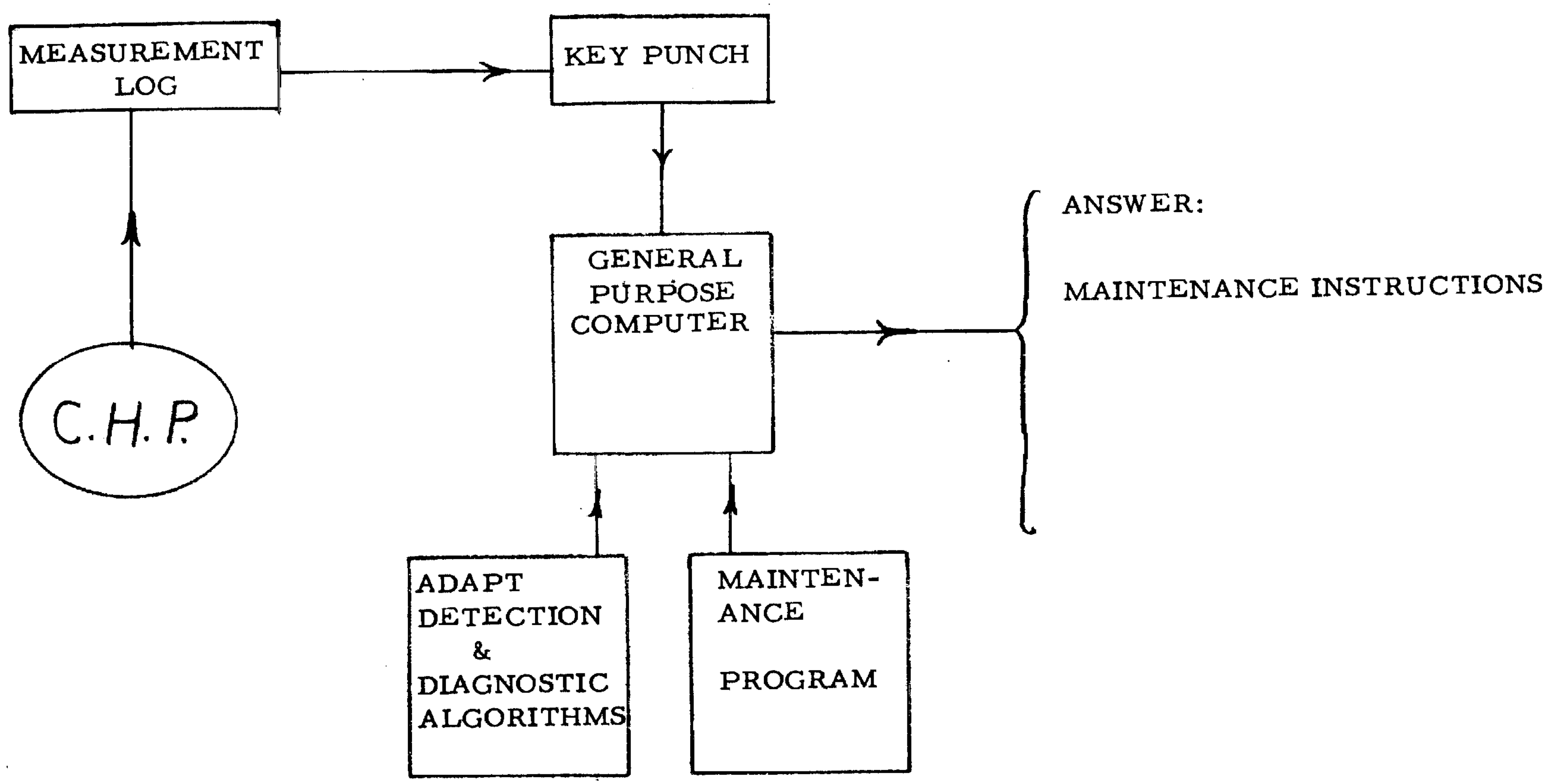


FIGURE 3.5

USE OF ADAPT MAINTENANCE ALGORITHMS

(AUTOMATED MONITORING)

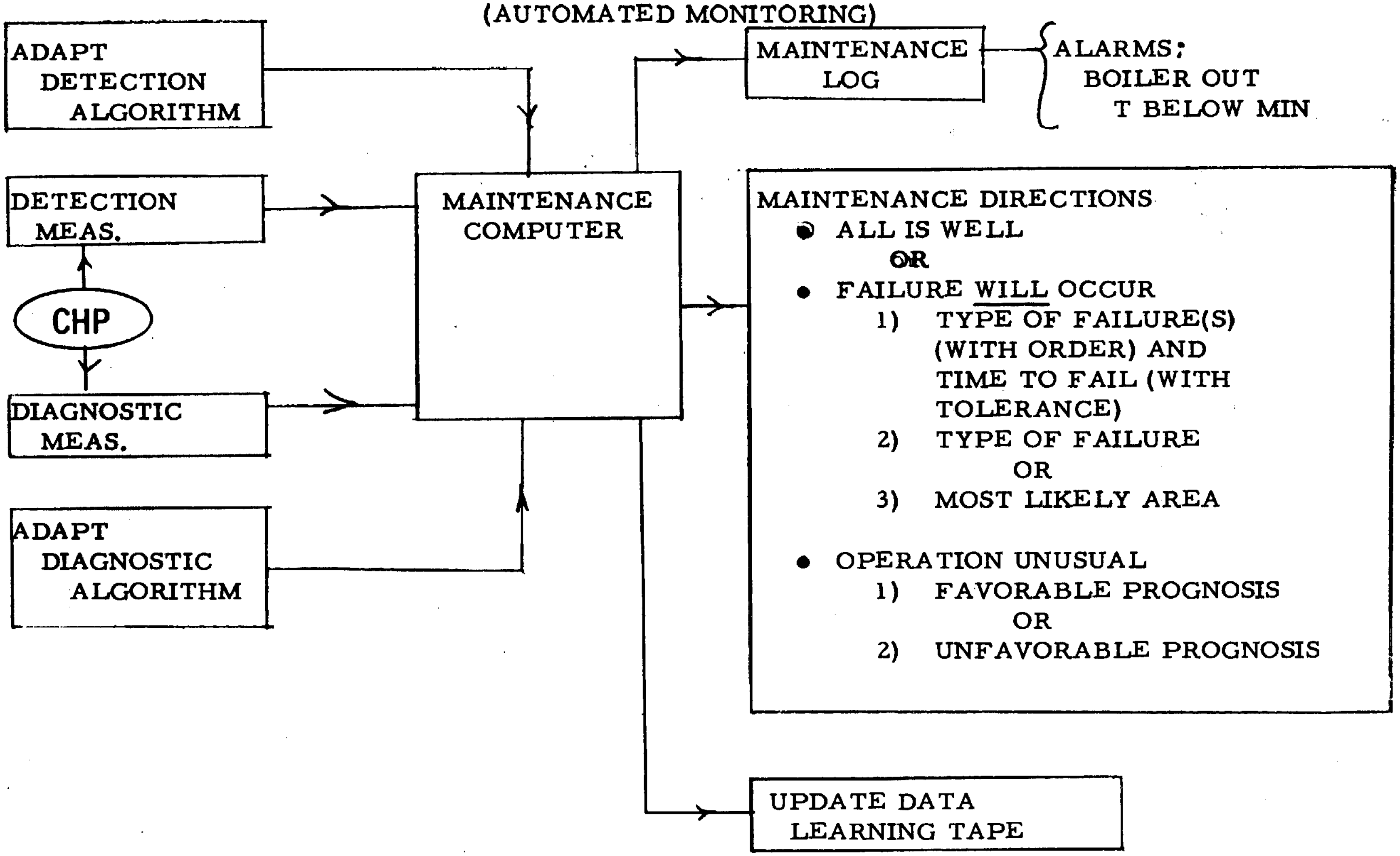
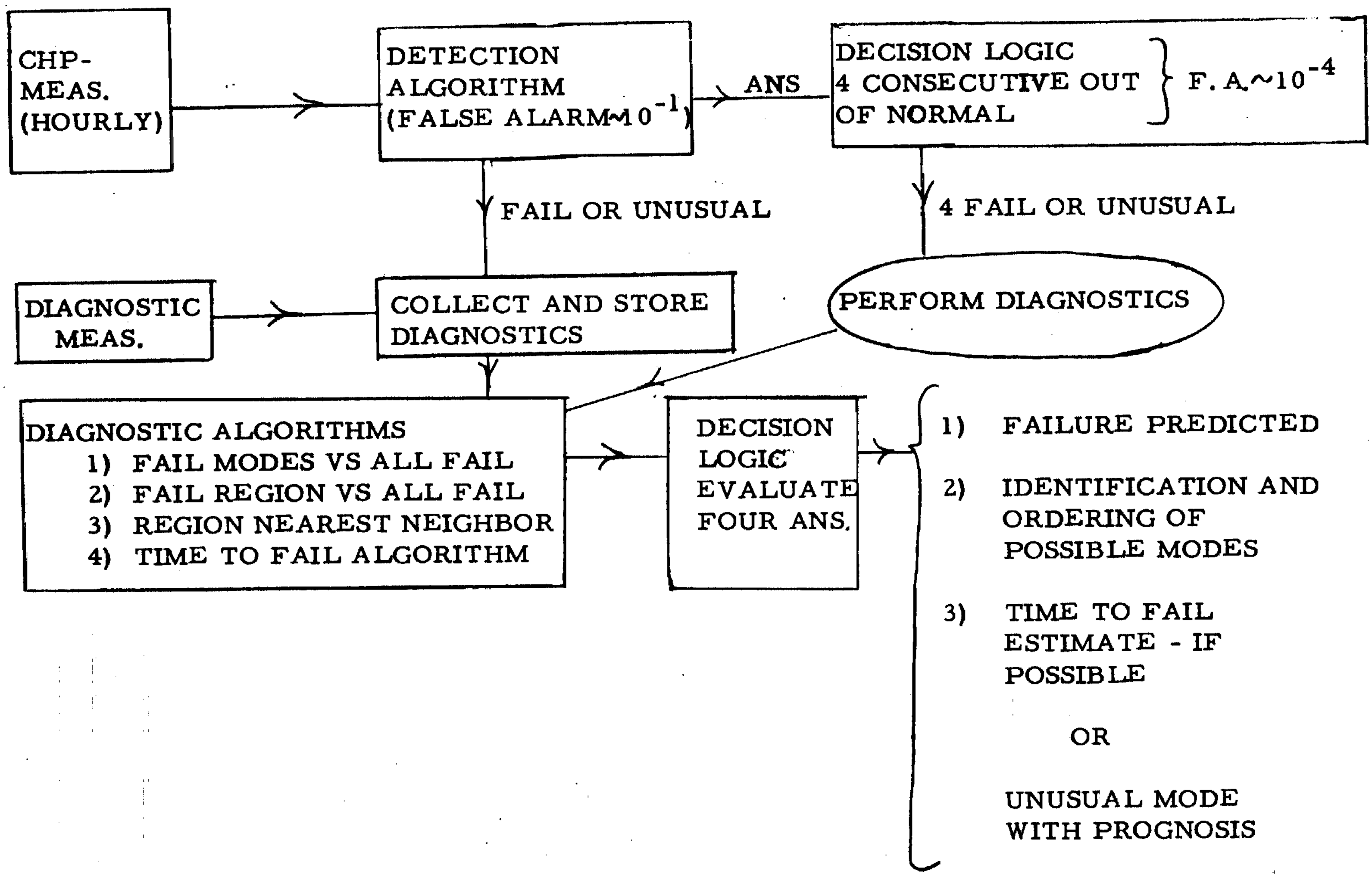


FIGURE 3.6

MAINTENANCE COMPUTER PROGRAM



4.0 DESCRIPTION OF ADAPT

4.1 Definition of Data Histories

The ADAPT techniques address themselves to the representation and empirical analysis of data which appear as data histories, i.e., an indexed series of numbers. The generality of the ADAPT programs can be seen from the variety of applications described in References 1-15. The features of ADAPT which make it advantageous for empirical analysis are reviewed in Appendix A. In the present case the indexing variable is the name of the measurement. Thus, the indexed sequence of numbers which characterize the operation of the KSC central heat plant at any given time may be viewed as a data history or plot as illustrated in Figure 4.1. Here we have plotted the value of the measurement as a function of the indexing variable which is simply a number associated with the name of the measurement. To illustrate this the name of the measurement has been included in Figure 4.1. This figure gives a portion of the data history for Aug. 14, 1970, at midnight. From this figure we see that no rain fell between 11:00 p.m. and midnight on Aug. 14. The temperature at midnight (i.e. measurement No. 2) was approximately 82° and the average rainfall for the past twelve hours was also zero. The average temperature from noon until midnight was 80°. This process is continued until a curve defining all of the measurements to be analyzed is generated. This curve is defined as the input data history for the case associated with Aug. 14, 1970, at midnight. Similar cases are generated for each of the days and times considered in the analysis.

In general, the histories may be given in continuous (analog) form or in discrete form. Since the ADAPT programs operate in digital computers, analog histories are each digitized into a finite set of N numbers, so each data history is treated as an N-dimensional vector in Euclidean space. If there are M histories, the result is an N x M matrix of numbers.

4.2 Optimal Representation of Data Histories

With the M input history vectors defined, the first step in ADAPT is to construct a set of optimum orthonormal base vectors. Since in general the number of optimum base vectors to be generated will be less than the numbered required for complete representation, there will be an error vector equal to the difference between the history vector and its representation in the new optimum base. The square and magnitude of this error vector is the measure of error for each history, and the average of these square magnitudes for all histories is the mean square error incurred in representing the data history vectors in the new base. For this process the definition of the word "optimum" in the expression, "optimum orthonormal base" is that the optimum base is that base which minimizes the above defined mean square error incurred when one represents the learning data histories using the new base vectors. The optimum base is chosen in an ordered fashion, so that the first vector is the best and so on. For example, if only one vector is used in the new base, that base vector is the one which makes the one vector representation error the smallest. If a second vector is used also, it is chosen such that together with the first vector it minimizes the two vectors representation error. This is continued for as many vectors as is necessary or desirable for the analysis to be performed.

When formulated mathematically, this criterion requires the maximization of a quadratic form whose unknowns are the components of one of the "optimum" base vectors, and whose coefficient matrix is the covariance matrix of the input histories. This problem is a classical one in linear algebra, which often appears under the names optimum empirical orthogonal functions, Karhunen-Loeve Expansion, or principal components analysis of a matrix.* The solutions for the unknown vector components are the normalized eigenvectors of the covariance matrix, and the resulting values of the quadratic form are the eigenvalues of this matrix. Once they are obtained, they are simply arranged in order of decreasing size of the eigenvalues. The largest eigenvalue gives the most reduction in mean square error that can be achieved with only one new base vector and the corresponding eigenvector is this new base vector. The next largest eigenvalue gives the most reduction in the error that can be achieved by using a second new base vector in addition to the first one found above, and this second vector is the eigenvector of this second largest eigenvalue. This process can be continued until the desired accuracy is achieved. The sum of the NR largest eigenvalues gives the maximum mean square error reduction which can be achieved with NR new base vectors; when adding additional eigenvalues does not significantly increase this sum, the use of the corresponding eigenvectors as additional base vectors does not significantly improve the representation.

A convenient measure of the degree of representation achieved with a given number of base vectors is the sum of the eigenvalues of the vectors used, divided by the average square magnitude of the original data history vectors. This represents the reduction in mean square error achieved divided by the total error reduction possible; in statistical terms this is the percent of the variation of the data explained by the representation used. Since information is only conveyed by the variation in the data and the variation has the form of an energy, the percent variation explained is also known as the information energy. A similar measure of representation which is applied to the individual data vectors is the ratio of the square magnitude of the data vector in the NR base vector system to the original square magnitude of the data vector. This provides a measure of the adequacy of the empirically derived base for representing each history, and when applied to a test history serves as the basis for

*For a detailed discussion of the Karhunen-Loeve Expansion and its advantages in empirical data analysis see Reference 16.

the apriori test of the validity of applying the empirical data analysis to the test case.

For each history the NR components in the optimal system are the optimal representation of the data in the sense described above. Alternatively, these components may be interpreted as coefficients of the Fourier series of optimal orthonormal functions representing the history. Thus, this vector analysis is equivalent to the expansion of functions in a set of orthonormal functions, of which the Fourier series is the most common example. The approach taken is analogous to the classical solution of boundary value problems in mathematical physics where the appropriate differential equation is used to define a set of orthonormal functions to satisfy a given function on the boundary. This boundary function is then expanded in the set of orthonormal functions defined by the governing differential equation. In the case of empirical data analysis, the governing differential equation is not available to define the set of orthonormal functions and instead the learning data set is used to numerically define the best set of such functions or vectors.

The optimal components are used in all further empirical analysis. Thus, the original M x N numbers representing M histories have been reduced to M x NR components, plus N x NR numbers to define the optimal vector base. Since the base system is optimal, the number of terms, NR, necessary to give a useful representation of history is small, often of the order of 10 or less, and the reduction in the number of numbers is usually large.

The ADAPT representation process just outlined can be clarified with the simple example of two input histories, which has been carried through analytically in Appendix B. For this special case the first optimal function is proportional to the average of the two history functions, the second to their difference, a result in accord with simple intuition. The relative sizes of the two eigenvalues is found to depend on the degree of correlation of the two histories. This illustrates the point that the more highly correlated information appears in the first term of the optimal representation. Thus, the last terms in the ADAPT representation are the most noise-like, and dropping of terms in the ADAPT representation results in retention of the easiest to use information.

4.3 Use of Optimal Representation for Developing Predictive Maintenance Algorithms

Having arrived at the optimal (Karhunen-Loeve) representation, attention is now turned to use of the optimal components for performing empirical clustering analysis, classification, parameter estimation, extrapolation and clutter subtraction. For clustering analysis, one represents each history by a point in optimal coordinates, and the degree of similarity of two histories can be defined as the distance between their two points. If the optimal representations are normalized, this distance is simply related to the correlation of the two histories.

Thus, the application of visual, nearest neighbor, or other cluster identification schemes to points (i.e. data histories) of the optimal space will lead to identification of natural clusters and algorithms to identify their members.

For classification (including the special classification problem of detection) the same representation of a history as point in optimal coordinates is used. A number of parametric schemes and linear non-parametric schemes which can be applied are included in the ADAPT programs. One frequently used scheme assigns a single number to each history in the following way: All the histories are divided into two classes according to the sorting desired. Then an unknown direction (vector) in the optimal space is postulated, and the projection of each history on that direction is obtained. This projection is a scalar associated with each history. The mean of this projection for each of the two classes is found, and then the difference between the two means. Also, the dispersion of the projections of each class about its own mean is found. The postulated direction of projection is determined by maximizing the ratio of the squared distance between the mean projections to the sum of the dispersions of the projections. When the direction of the projection is known, the projection of each history is determined and the range in which it falls for each class can be found. The criterion that a given new history is sorted into a given class is that its projection on the direction found in this way falls within the range of projections of the learning data of that class. This linear scheme for sorting into two classes was first suggested by Fisher, and is known as the Fisher linear discriminant.

This and other linear schemes may be extended to multi-class problems by repetitive application, separating a different class with each application. If the statistics of the learning data are Gaussian the maximum likelihood technique, which is included as an option in ADAPT, may be used for multi-class classification problems.

The ADAPT technique for constructing an algorithm to predict a physical parameter associated with each history again makes use of the components of each history in the optimal system. For every history in the learning data, the known value of the parameter is written as a linear combination of the optimal components. The unknowns are the coefficients in this linear combination, which are taken to be the same for every history. The sum, over all histories, of the square error of this linear representation is then minimized to determine the coefficients. This amounts to a regression of the parameter on the optimal components. When the coefficients are found, they can then be used with optimal components of any new history to obtain an estimate of the value of the parameter for that history.

ADAPT offers a unique approach to extrapolating data histories. The learning data used is the entire history, including the region over which one hopes to eventually extrapolate. This learning data is first used to find the optimal representation for the entire history. One then determines the best coefficients by making a least square fit of the available portion of the histories to a generalized Fourier series using the optimal orthogonal functions over the available portion of the history and these coefficients are used to reconstruct the entire history from the complete optimal orthogonal functions.

The task of clutter subtraction is accomplished by first obtaining data histories which characterize the clutter to be subtracted. These are the first few ADAPT optimal functions obtained from data histories which were produced solely by the phenomena whose characteristics are to be subtracted. These histories characterize the clutter to be subtracted and are utilized as the first directions in a Gram-Schmidt orthogonalization. The ADAPT optimization is interrupted after the Gram-Schmidt orthogonalization and the components associated with each of the directions determined by the clutter histories are set equal to zero for all data histories. The ADAPT optimization is then continued through the Karhunen-Loeve expansion, resulting in an optimal coordinate system which does not contain the directions associated with the clutter to be subtracted. When the histories are reconstructed using the series expansion in terms of these optimal functions (i.e. coordinate directions) the resulting histories no longer contain the characteristics of the clutter which was subtracted.

It is not necessary to actually find the optimal coefficients of a new history which is being investigated to apply an ADAPT derived algorithm. The transformation from the N-dimensional data vector space to the NR-dimensional optimal vector space can be inverted and incorporated into the algorithm vectors. Then the process of applying this algorithm to a new data vector involves primarily the dot product or combination of dot products of this N-dimensional data vector with an N-dimensional algorithm vector or vectors, a rather simple procedure.

The development of maintenance algorithms for the KSC central heat plant will require the use of both the classification and parameter estimation capabilities of the ADAPT programs. Two types of classification algorithms can be of use for the maintenance problem: 1) Failure detection algorithms and 2) Failure diagnostic algorithms. Failure detection algorithms are classification algorithms in which one class consists of all of those data histories corresponding to times at which the system is performing normally and the other class are those data histories corresponding to times when a failure will occur in the near future. Diagnostic algorithms are required to determine which failure mode is expected. This is obtained by performing a classification analysis to separate each of the failure modes either from all other failure modes or from all other cases. Finally, after one has determined that a failure is going to occur and has diagnosed what type of failure this will be, it would be useful to estimate the time at which the failure would occur. This can be accomplished through the application of parameter estimation algorithms where the parameter to be estimated is the time to failure. Thus, both the classification and parameter estimation capabilities of the ADAPT programs will be utilized to develop the appropriate maintenance algorithm for KSC central heat plant.

4.4 Evaluation of Performance and Validity

An objective of the ADAPT approach to empirical data analysis is to provide the analyst with information regarding both the performance and the validity of the algorithms which he develops. The performance tells the analyst how good his algorithm is when it is applied to test data belonging to the same population as the learning data used to derive the algorithm. The validity criteria is a measure of how well the test data belongs to the population of the learning data. Thus, the availability of performance data allows the analyst: 1) to select the best algorithm, 2) to verify that the performance of the algorithm is sufficient to accomplish the objectives, and 3) to insure that the algorithm is based on physics and not merely a fortuitous manipulation of the data. The validity criteria on the other hand provides the user with a measure of the applicability of the algorithm to the particular case being tested. We shall now discuss the performance measure and then the validity criteria.

Performance Measure - Fisher Discriminant

The linear discriminant used for the analysis of the KSC heat plant data was the Fisher discriminant. Similar performance measures may be developed for any linear discriminant, but many details of these performance measures will differ for the particular discriminant. Since the Fisher discriminant was the only one used for the analysis of the KSC central heat plant and the performance measures associated with the application of ADAPT programs are most highly developed for this discriminant, we shall limit the present discussion to performance measures applicable to the Fisher discriminant.

The simplest measurement of the performance of a linear classification algorithm such as the Fisher discriminant is to examine the projection values actually obtained when the learning and/or test data is projected on the optimum directions selected by the linear discriminant. The ADAPT programs present a bar chart plot of these projections for each of the learning cases, which can be used to visualize the performance of the algorithm on the learning data. Figures 4.2 and 4.3 present such bar charts comparing the performance of the universal detection algorithm derived using 192 measurements and the performance derived using 50 measurements respectively. Examination of these figures shows that although they present a detail view of the performance on a case by case basis, it is difficult to get an overall picture of how much better one algorithm is than the other algorithm. Furthermore, it is clear that if one wished to compare the more than 30 detection algorithms which were derived as a part of this study, this measure of performance would be extremely awkward to use.

The most desirable way to overcome the twin difficulties of obtaining a convenient overall measure of the algorithm performance and of comparing a large number of algorithms is to evaluate the performance of an algorithm in terms of a single number. Since the Fisher discriminant is the result of the minimization of the ratio of the sum of the squares of the within class scatter divided by the distance between the means of the two classes, the value of this parameter is an excellent measure of the performance of the Fisher discriminant. In particular, the smaller the value of this parameter the better the performance of the algorithm. For example, for the bar chart shown in Fig. 4. 2 this parameter, designated by the quantity $\Sigma\sigma / V$ has the value of .52, whereas for the algorithm presented in Fig. 4. 3 this parameter has the value of .41. In Appendix C it is shown that for the special case of equal standard deviations of the projections of each of the classes, this parameter is uniquely related to the probability of making an error. The corresponding values of probability of error for the bar charts shown in Figs. 4. 2 and 4. 3 are approximately .05 and .005 respectively.

It is of interest to plot the performance of an algorithm as a function of the ratio of number of cases to number of dimensions used to develop the algorithm. A plot such as this is called a performance map and Fig. 4. 4 illustrates this performance for the cases shown in Figs. 4. 2 and 4. 3. The solid symbols in each case represent the actual algorithm illustrated by the bar chart in Fig. 4. 2 and 4. 3. The open symbols represent other algorithms derived using the same data and a different number of dimensions. This curve is particularly useful because it allows the analyst to decide whether he may have confidence that the algorithm is based on physics or is "overdetermined" and merely represents a mathematical manipulation of the data with no physical meaning. For example, consider the situation of fitting 3 points to a third order polynomial. The third order polynomial represents a three dimensional space. Fitting 3 points to this third order polynomial is always possible and normally these "overdetermined" coefficients have no physical basis. However. if a significantly larger number of cases, say 30, is fitted to this third order polynomial then one knows that there must be some physical relationship embodied in the polynomial which allows one to fit 30 cases to a third order polynomial. The same is true in any empirical analysis and in general, this phenomenom is a function of the performance of the algorithm. This is illustrated in Fig. 4. 4 by the cross hatched area which separates random separations from good separations. The random separations represent Avco's experience with a great number of problems and show the region in which the algorithm can perform even if there is no physical basis for the separation. Thus, the location of an algorithm on the performance map immediately tells whether this algorithm displays the overdetermined character or not. It is also clear that as one decreases the number of dimensions one moves vertically on the performance map. However, at some point the decrease in dimensions will also eliminate some information which is useful to the separation. When this occurs the performance of the algorithm will decrease, the algorithm will move both to the right and upward on the performance map. The objective is to get as far to the left on the performance map as possible while satisfying the requirement of remaining a significant distance away from the random separation's region. Thus, the performance map has been a useful tool for comparing different algorithms and for carrying out the analysis required to determine the dimensionality at which an algorithm should be produced.

Although the value of the Fisher parameter or the performance map make excellent performance measures for comparing different algorithms, they do not directly display the trade-off between detection probability and false alarm rate. It is shown in Appendix C that for the special case where the standard deviation of both classes are equal, the performance map can be related directly to the trade-off curve between detection probability and false alarm rate. However, this trade-off curve produced for any given algorithm is another excellent way to compare algorithms since it provides a pictorial display of this trade-off. Figure 4.5 presents these trade-off curves for the same algorithms shown in Figures 4.2 thru 4.4. The ordinate on this plot is the detection probability, that is probability that an out of tolerance condition of the KSC plant will be detected by the algorithm. The abscissa is the false alarm rate for the probability that the normal period of operation will be called abnormal. It is clear from examination that Figure 4.5 that the trade-off between the detection probability and false alarm rate for each of the algorithms is shown very clearly. In addition, this presentation clearly shows the relative merits of the algorithms being prepared. Thus, once the dimensionality of an algorithm has been selected, this detection probability versus false alarm rate curve provides the most convenient method of comparing algorithms. In general, through the remainder of this report when an algorithm's development is being discussed its performance will be displayed on a performance map. When an algorithm has been developed and is being discussed for use and testing, its performance will be displayed on a detection probability versus false alarm curve.

Performance Measure - Parameter Estimation

The problem of evaluating the performance of a parameter estimation or regression algorithm is quite similar to that of estimating the performance of a classification algorithm. The simplest display in the performance of estimation algorithm is a plot of the estimated value of the parameter versus the actual value of the parameter. Fig. 2.4 shows such a plot for the estimated time to failure for the atomizing steam boiler. Thus, the functional role of this presentation of the regression results is very similar to that of the bar chart for the classification results. It shows the performance of the algorithm on each case extremely well. It also gives an excellent graphical qualitative picture of how well the algorithm is working. But like the bar chart it is an awkward presentation for comparing a large number of algorithms or for the analysis of dimensionality to be used in developing the algorithm. These two functions are again better performed on a performance map.

As in the case of classification, the development of a performance map for parameter estimation requires that one have a single number to evaluate the performance of the algorithms including such things as correlation coefficient or standard deviation of the error. In the ADAPT programs the measure of performance is the ratio of the standard deviation of the error resulting from the application of the algorithm divided by the variance (i.e. the standard deviation of the error when one uses the mean as the estimate of the parameter). This ratio is designated by the symbol σ_{RAT}. Again, the smaller the value of this ratio, the better the performance of the algorithm. Thus, the performance map shown in Fig. 4.6 for the time to failure algorithm illustrated in Fig. 2.4 is a plot of σ_{RAT} versus the ratio of the number of cases to number of dimensions. Here again, the ratio of the number of cases to number of dimensions plays the same role as it did in the classification algorithms. Again, experience with previous empirical problems has allowed the inclusion of an experience factor for the probability that the algorithm will be based on the physics of the problem and not merely a random separation. Thus, the regression performance map shown in Fig. 4.6 can again be used both to compare the performance of algorithms and as a tool for the analyst while developing the algorithms.

Validity Criteria

The ADAPT programs also provide validity criteria which are based on the ability of the optimal functions derived from the learning data to represent the test data. These validity criteria are identical for and applicable to all ADAPT classification prediction and clustering algorithms. The validity criteria essentially makes use of the data vector's geometric property of length. The length of the learning data vectors may be calculated in the original data space and then compared with the new length when the learning data is represented in the optimal ADAPT space. The ratio of these two lengths is defined as the validity parameter (Q). The validity parameter can be calculated for the test data vector by computing its length in the original data space and the optimal ADAPT space. If the test data vector's length is reduced significantly more than that of the learning data vectors when it is represented in the optimal space, this is indication that the test data is from a different population than the learning data used to develop that algorithm.

The major problem in applying this validity criteria is that of establishing the threshold between valid and invalid cases. The correct way to establish this criteria requires the knowledge of the distribution function of the validity parameter

E

for both the population of valid cases and invalid cases. It is clear from some rather obvious limits such as the fact that the validity parameter must lie between zero and one and that its standard deviation at both of these end points must be zero that the distribution function is definitely non-Gaussian and, in fact, is not satisfied by any of the well known classical distribution functions. Thus, one must experimentally develop the appropriate statistical properties of each of these populations. It is relative easy to get a reasonable approximation to this distribution function for the population of the valid cases by plotting up and examining the validity parameters for the learning data. However, it is considerably more difficult to find an estimate of the statistics for the validity parameter of the invalid cases. In fact, the only approach available for this at present is to make some reasonable assumption for a threshold such as the minimum value observed in the learning data or the mean value in the learning data minus some quantity such as the standard deviation, evaluate a series of test data against this criteria and then re-examine the performance on the test data and determine the conditions for which the results are consistent. This will be illustrated in more detail in Section 5.4 where the heat plant failure detection algorithm performance is evaluated.

The validity criteria for the ADAPT extrapolation of data histories is based on the fact that the learning data is now identical to the first portion of the data histories and was not used to make the data base. However, the data which was used to make the base also contains the portion covering the identical range of the indexing variable as the learning portion of the data history to be extrapolated. Thus, one may compute the RMS error for the first (i.e. known) portion of all the learning data histories. One may then take the average of this, finding the average RMS error for all the learning data histories and also the standard deviation $\sigma_{\bar{E}}$ of these RMS errors. One may then compare the RMS error of this known range of the test case with the average and standard deviation of the RMS error for the corresponding region of the learning data and calculate the confidence in the validity of the extrapolation. For example, if the RMS error of the test cases falls outside of the range of the average RMS error for the learning data plus or minus its two-sigma value, one has only 5% confidence that the extrapolation will be accurate to the degree indicated by the performance estimate based on the learning data.

The next sections of this report will present the detailed results of the representation, detection, diagnostics, and time to failure estimates derived for the KSC central heat plant using the methods which have been outlined above.

FIG. 4.1 - CONSTRUCTION OF DATA HISTORY FROM DISCRETE MEASUREMENTS

History No. 2005 = Data History for Aug. 14, 1970 @ 2400 Hrs.

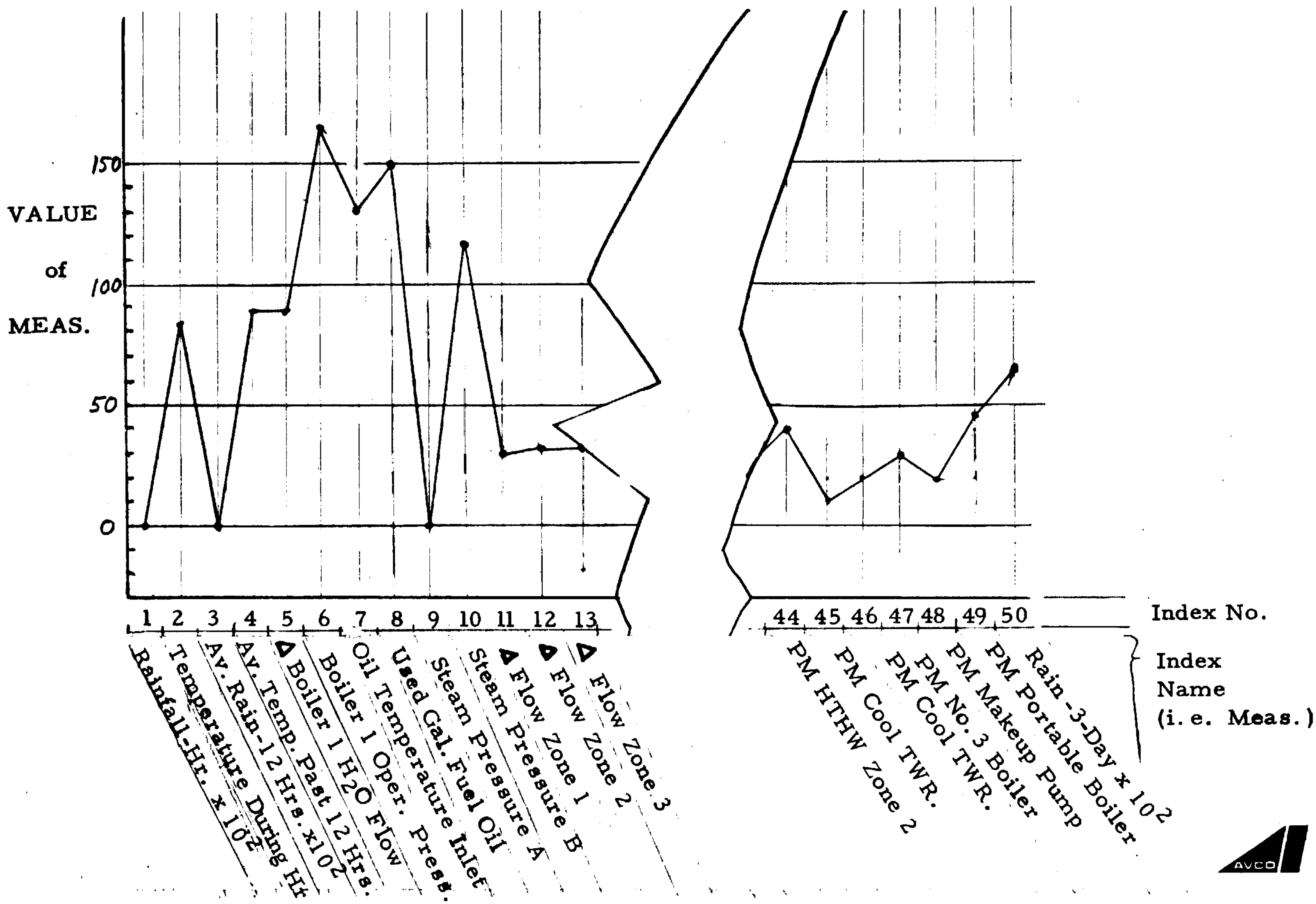


FIGURE 4.2 - PROJECTION OF LEARNING DATA ON SEPARATION DIRECTION FOR SEPARATING INCIPIENT FAILURES FROM GOOD CASES USING 192 MEASUREMENTS

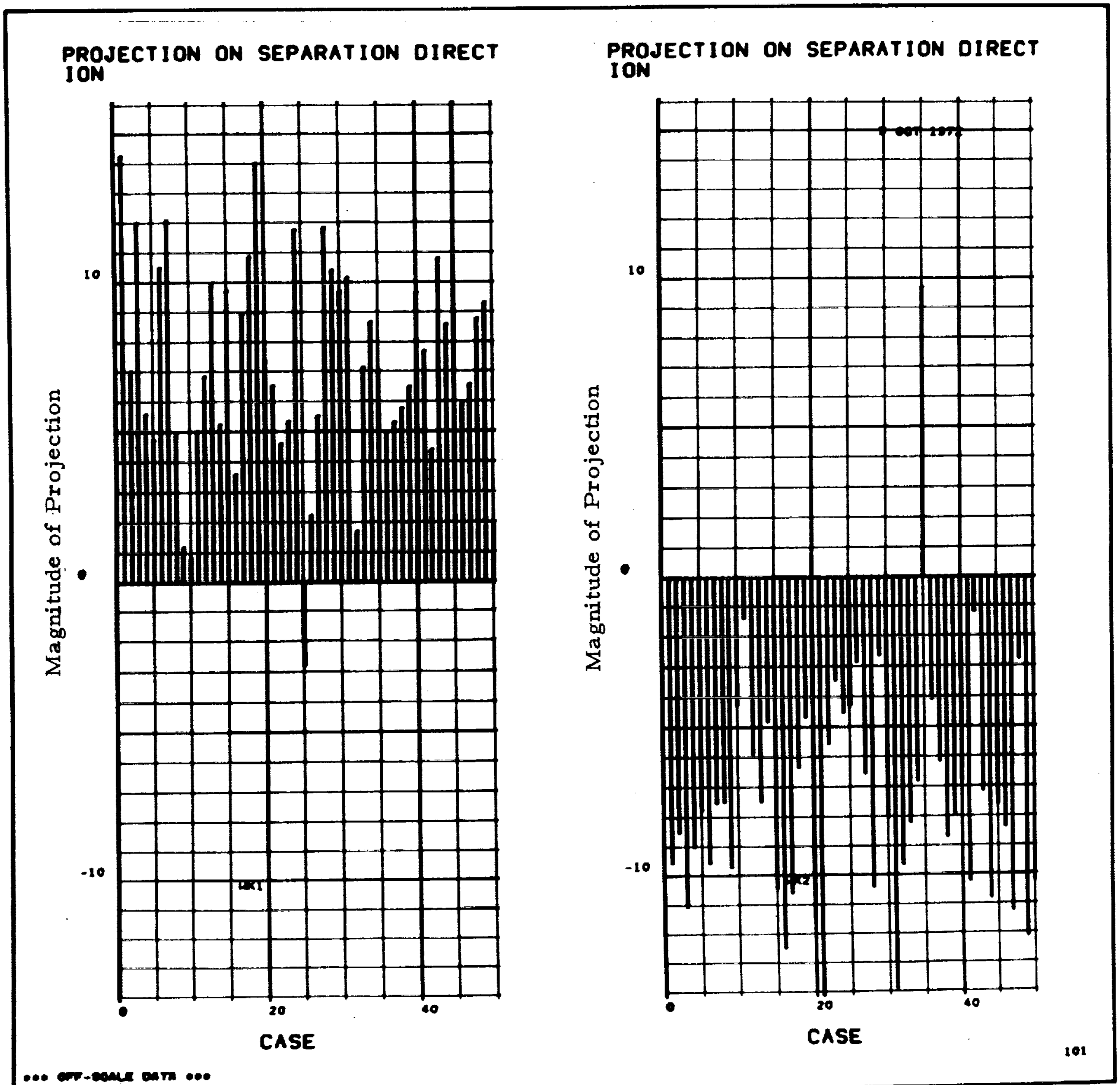
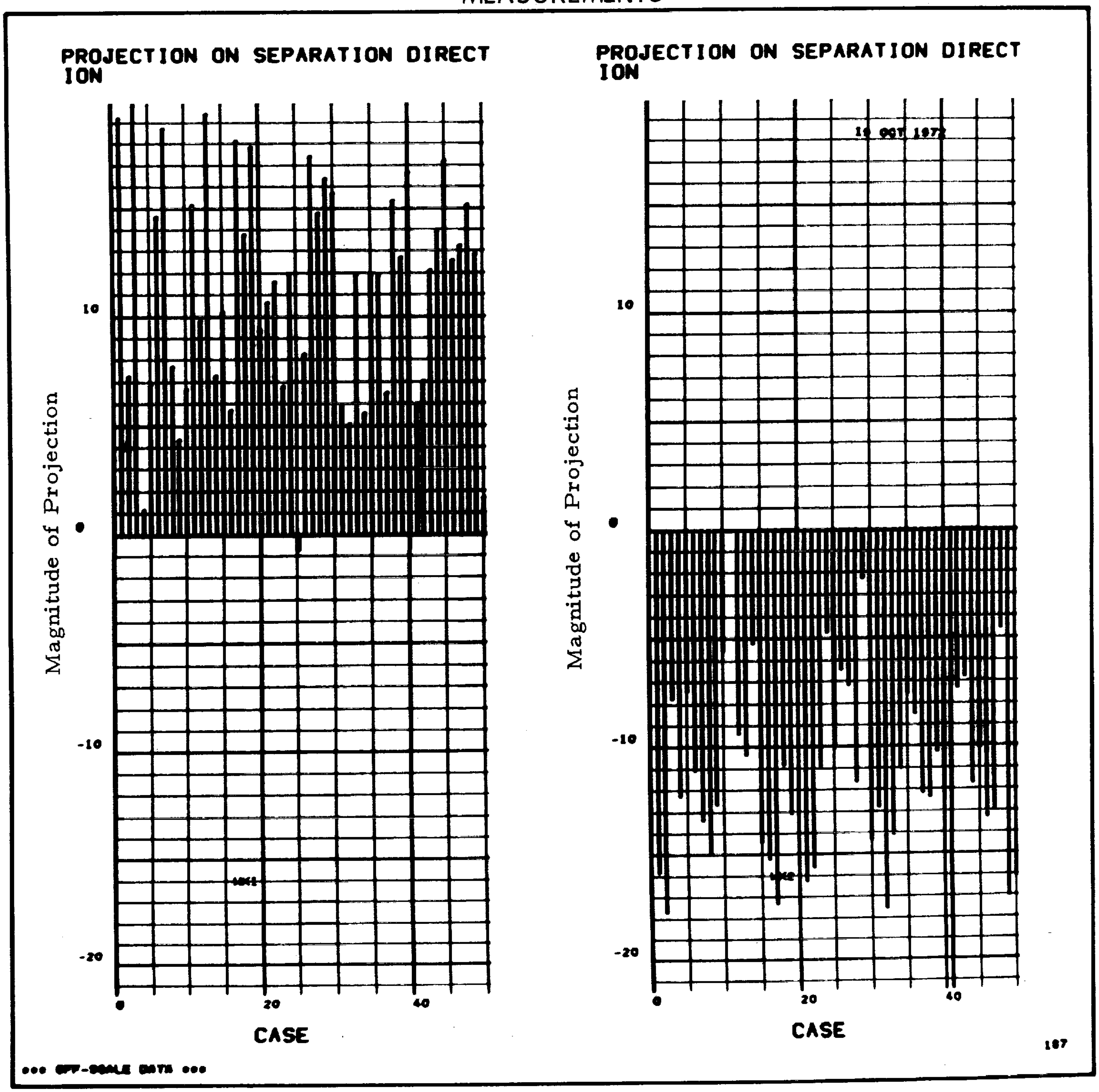


FIGURE 4.3 - PROJECTION OF LEARNING DATA ON SEPARATION DIRECTION FOR SEPARATING INCIPIENT FAILURES FROM GOOD CASES USING 50 MEASUREMENTS



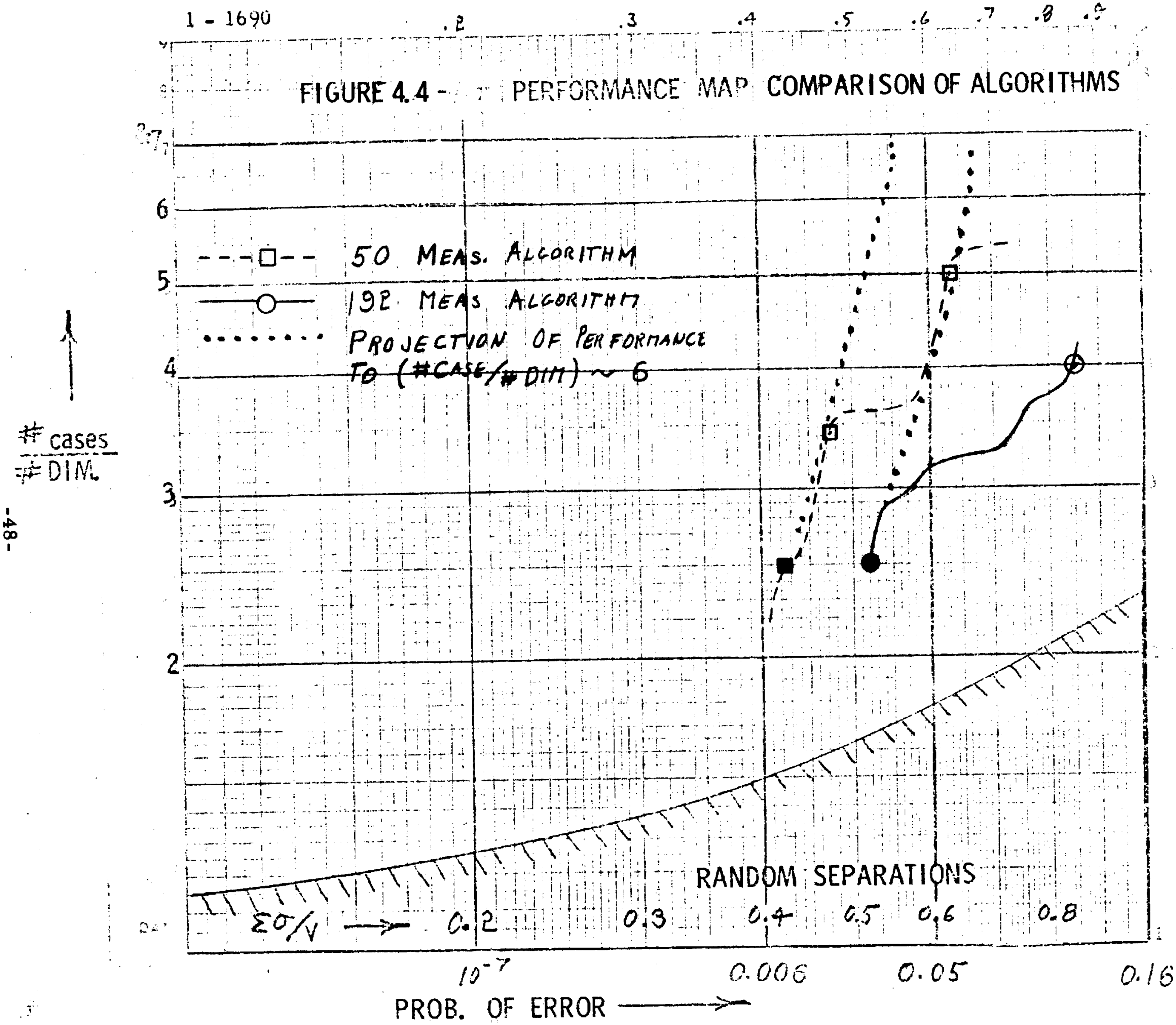
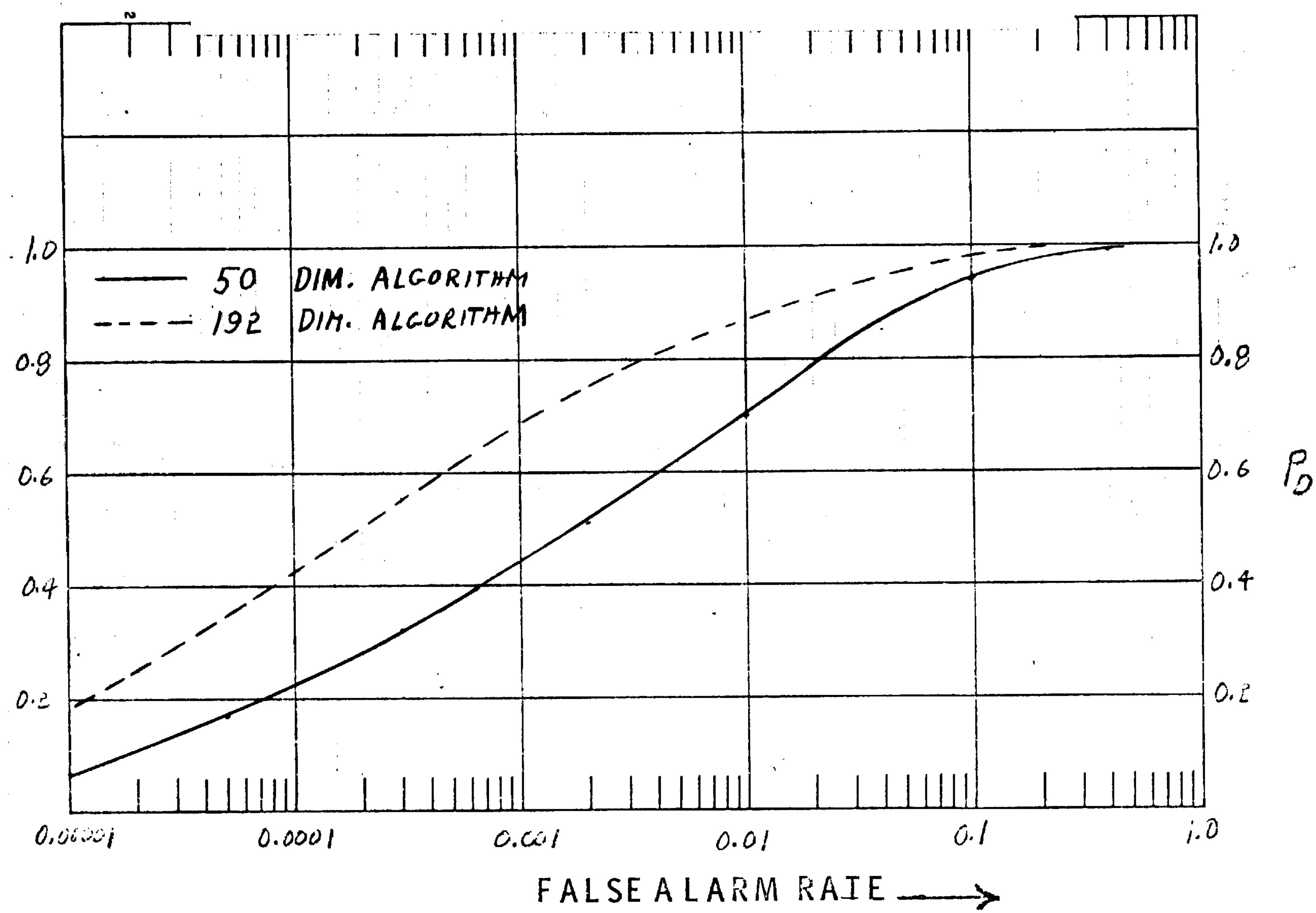
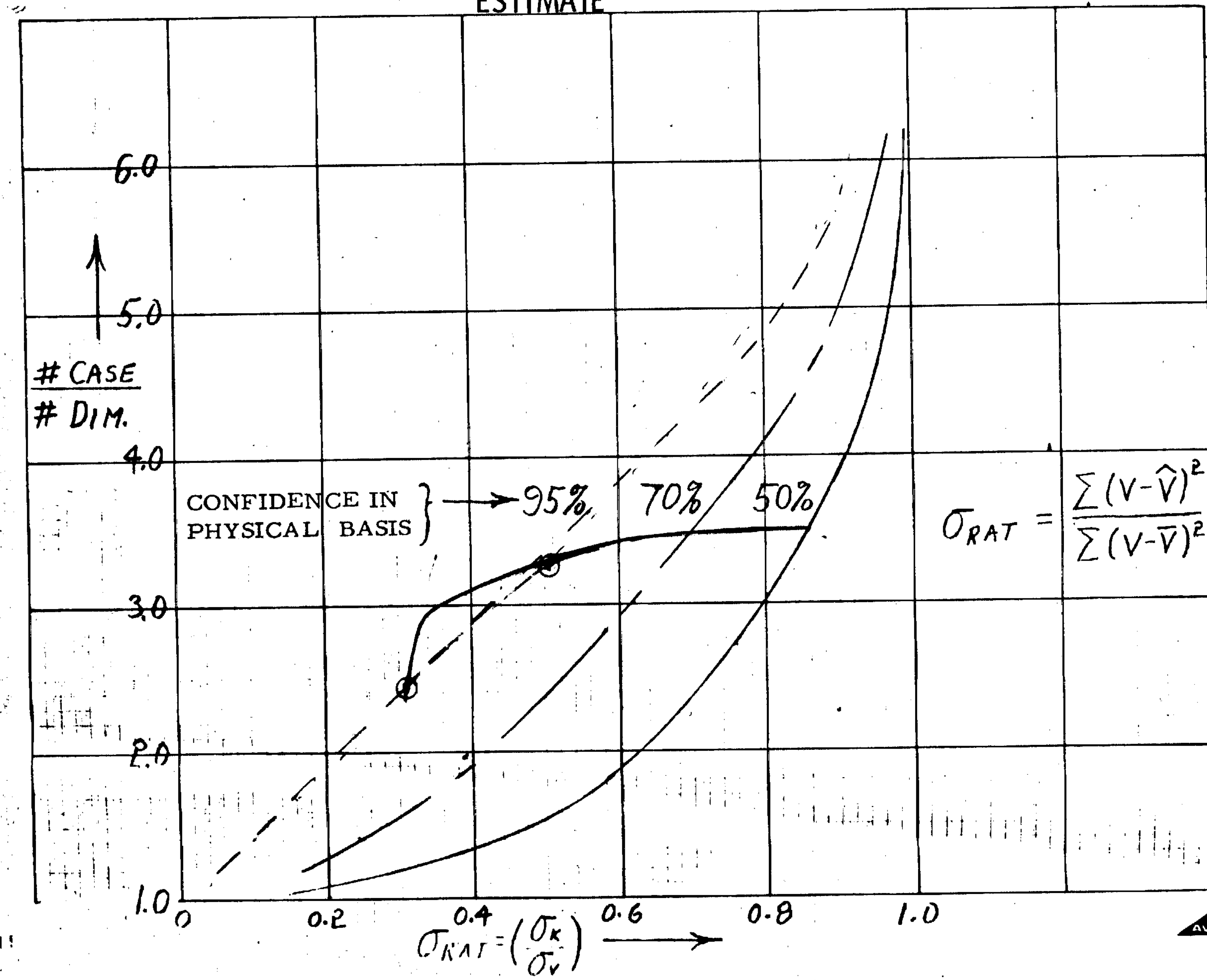

1 - 1690
.2
.3
.4
.5
.6
.7
.8
.9
FIGURE 4.4 - PERFORMANCE MAP COMPARISON OF ALGORITHMS
50 MEAS. ALGORITHM
192 MEAS ALGORITHM
PROJECTION OF PERFORMANCE TO (#CASE/#DIM) ~ 6
7
6
5
4
3
2
cases / # DIM.
RANDOM SEPARATIONS
Σσ/V
0.2
0.3
0.4
0.5
0.6
0.8
10^{-7}
0.006
0.05
0.16
PROB. OF ERROR

FIGURE 4.5
PROJECTED CLASSIFICATION PERFORMANCE TRADE-OFF CURVES FOR COMPARING ALGORITHMS



NOTE: USE TYPE B PENCIL FOR VUGRAPHS AND REPORT DATA.

FIGURE 4.6 - REGRESSION PERFORMANCE MAP FOR TIME TO FAILURE ESTIMATE



5.0 DETECTION ALGORITHMS

The implementations of the ADAPT derived maintenance algorithms into a demand preventive maintenance (DPM) system described in Section 3 is only feasible if one can derive an algorithm which will detect incipient failure of the system. This section will present the demonstration of the feasibility of accomplishing this by applying the ADAPT programs to the KSC central heat plant data. To insure success in obtaining detection algorithms several different avenues were explored and each of these avenues shall be reviewed in this section. Many successful detection algorithms were derived. The selection of the algorithm to be used for the final demonstration was based on the algorithm performance, complexity and the complexity of the software required to implement a complete set of algorithms of the selected type. The algorithm selected was the universal detection algorithm for boiler No. 1. This algorithm was then optimized for maximum performance and tested using approximately 200 independent test cases. All of these results are discussed in Sections 5.1 thru 5.4. Section 5.5 presents a discussion of the implications of these results to preventive maintenance.

5.1 ADAPT Representation of Heat Plant Data

The power of the ADAPT approach to data analysis is primarily due to the derivation of the optimum representation for any given set of data prior to developing the empirical algorithms. The representation obtained by processing through the ADAPT programs is correct for any set of data having the same number of independent variables or indexing points as the data histories used in the learning data. It is an optimum representation for that subset of this data for which the learning data is a good sample of the population statistics. Thus, it is necessary to develop a new base whenever the number of index variables used for the analysis is changed and it is desirable to develop a new base whenever the distribution of subclasses in the learning data set is drastically changed. Every time one changes the number of measurements used in the analysis, it is necessary to develop a new base to use the smallest number of measurements in the ADAPT processing. Furthermore, if one drastically changes the approach to achieving the detection algorithm, such as changing the approach from deriving a universal detection algorithm to deriving an algorithm for subgroup on the scatter plot, it is desirable to develop a new base. It is also desirable to use a specific diagnostic base when developing the diagnostic algorithms. For these reasons a relatively large number of bases were developed in the course of this study both for the detection algorithms and the diagnostic algorithms.

The methodology of developing the base, the general results displayed by the base and the methods of using these results are essentially identical regardless of the details of the base. This section shall present a detailed review of only one base. The initial base developed in this study was the 29 case exploratory base using 190 independent variables. This base was used for the initial exploratory analysis which will be described in Section 5.2. At the conclusion of the exploratory analysis it was decided to add two more variables to the candidate variables and to use a basic 100 case data set as the learning data for the detection algorithms. This data set was then used to develop a new base with 192 measurements and initial studies were performed on this base which eventually lead to the reduction from 192 variables to the 50 most important variables or indexing points for detecting incipient failure. This 50 point base is the base which was used in the final analysis and is most relevant to the universal detection algorithm which has been selected for detail evaluation. The details of this base will be used to illustrate development of the ADAPT representation.

The 50 variables selected for the final detection algorithms are listed in Table 2.5. The scheme for configuring the value of the measurements associate with each of these variables or measurements into a data history suitable for processing in the ADAPT programs is presented in Section 4. This scheme was schematically illustrated in Fig. 4.1. Figure 5.1 presents the corresponding data vector for August 14, 1970, as plotted out in its entirety by the ADAPT programs. Figure 5.1 presents the data history for all 50 of the variables presented in Table 2.5. Although the names are not specifically listed on Figure 5.1 as they were on Figure 4.1, they correspond to the numbers or index shown in Table 2.5. For example, referring to Table 2.5 we see that indexing variable No. 30 is the number of gallons of oil used in the atomizing boilers. Referring to Figure 5.1 we see that in August 14, 1970, the number of gallons of fuel used in the atomizing boiler was 100.

The usual procedure in deriving the optimum representation with the ADAPT programs is to first subtract the average of all the data histories from each of the data histories to provide data histories having a zero mean. Figure 5.2 presents the average data history for all 100 learning cases used to develop this base. When these data histories were processed through the ADAPT programs, it was found that all the information obtained could be represented by 50 optimum functions. Figure 5.3 presents the amount of information presented as a function of the number of optimum functions used. For example, if one uses 5 optimum functions, Figure 5.3 shows that the fifth optimum function contributes almost 5% of the information contained in the total data set and the upper or cumulative curve on Figure 5.3 shows that the first five optimum functions taken together provide approximately 88% of the information contained in the data set.

Figures 5.4 and 5.5 present the first two optimum functions for representing this data. The indexing variable for these optimum functions is again defined by Table 2.5. Thus, examination of the first optimum function shown in Figure 5.4 shows that variables No. 6, 9, 10, 15, 16, 20, 21, and 25 dominate the variation. In fact, reference to Figure 5.3 shows that these eight variables account for almost 57% of the variation in the data. Reference to Table 2.5 shows that these variables are: Boiler No. 1 operating pressure, steam pressure for atomizing boiler A, steam pressure for atomizing boiler B, return temperature for zones 1 and 2, the 12-hr. average of the steam pressure in atomizing boiler A, the 12-hr. average for the steam pressure in atomizing boiler B, and the 12-hr. average of the supply pressure, respectively. Since variables 9, 10, 20, and 21 are dominant, one may interpret the first optimum function as being dominated by the definition of which of the two atomizing steam boilers (see Figure 3.1) are operating at any given instance. The next most important factor in determining the first optimum function is clearly related to the load on the system since it appears to be dominated by the boiler operating pressures, supply pressures, and the zones 1 and 2 return temperatures.

Examination of the second optimum function shown in Figure 5.5 shows that considerably more variables are important to this optimum function. The most important variables for defining the second optimum function are the boiler operating pressure, the amount of fuel oil used, the return temperature of the three zones, the 12-hr. average of the fuel oil used, the 12-hr. average of the supply temperature and pressure and the rainfall over the past three days. Realizing that the collection of the rainfall in the various drainage ditches and manholes throughout the distribution system is a major contributor to the load on the system, we see that the second optimum function is almost entirely determined by how hard the central heat plant must work.

Before continuing with a physical interpretation of the representation, it will be useful to clarify the meaning of these optimum functions by illustrating their use in a generalized Fourier series representation. Consider the reconstruction of the data history shown in Figure 5.1 using these first two optimum functions. Since all of the optimum functions are orthogonal functions, the coefficients for the generalized Fourier series may be obtained by the classical formulation which is simply a dot product of the corresponding optimum function with the data history. Given the set of coefficients corresponding to any data history, one can reconstruct the data history as follows. The first step is to take the first coefficient and multiply it times the first optimum function and add this on a point by point basis to the corresponding value of the average input vector (Figure 5.2) to obtain the one term reconstruction. One then takes the second coefficient and multiplies it times the second optimum function and adds the result again in a point by point fashion to the one term reconstruction to obtain the two term reconstruction. The two term reconstruction is shown in Figure 5.6, when it differs by more than the line thickness from the original reference presented in Figure 5.1 the reference history has been included on the figure as a dotted line. Comparison of the two term reconstruction (the solid line) with the reference history shows that for this case the reconstruction matches extremely well with only two terms. For this particular case, 62% of the information contained in the original data history is contained in this reconstruction. Comparing this with the information energy plot shown in Figure 5. 3, we see that this history is somewhat below the average of 72% and thus the two term representation represents a below average reconstruction. The process may of course be continued by using higher order terms in the series and the results of this continuation for the five term and ten term representations are presented in Figures 5. 8 and 5. 9 in a format similar to that of Figure 5. 6. Representations for the five and ten term reconstructions are 85% and 95%, respectively. Comparison of these representations with the average representation for five terms show that this reconstruction is slightly less accurate than the average reconstruction. In the case of the ten term reconstruction, it is a very typical match. As would be expected, the agreement between the reconstruction and the actual history improves as the number of terms used is increased. Furthermore, it is interesting to note that the most difficult portion of the history to reconstruct appears to be variables from approximately 31 thru 49. Reference to Table 2. 5 shows that these variables are the maintenance records.

Since the optimum functions used to reconstruct all of the data histories are identical, they can contain no information regarding the differences between any of the cases. Thus, the entire physics of the problem must be included in the coefficients of the optimum representation. Thus, Figure 5. 3 states that the two numbers corresponding to the first two coefficients for each of the histories in the learning data set represent 72% of the information which can be learned from this data set. Since two numbers can be conveniently presented in a two dimensional presentation, it is useful to examine this best possible two dimensional presentation. Figure 5. 9 presents a scatter plot of these two numbers. The abscissa on Figure 5. 9 is the coefficient of the first term in the generalized Fourier series representation of each of the cases shown. The ordinate is the coefficient for the second term in the Fourier series representation. The one's on this figure represent those cases taken during normal operation of the central heat plant and the two's represent those cases taken just prior to a failure in the central heat plant. Consider the case represented by the two located in the upper right hand corner of this figure. For this case, one would reconstruct the two term history by multiplying 205 times each of the values in Figure 5. 4 adding these numbers to 53 times the value of each of the indexing variables in Figure 5. 5 and sum the result of these two products with the average vector shown in Figure 5. 2. Examination of this figure immediately shows that in general the easiest to represent 72% of the information does not contain enough information to make the desired classification algorithm to separate incipient failures from non-failure cases.

Three groupings of cases can be seen on the scatter plot shown in Figure 5.9. The meaning of these groupings can be understood by recalling the physical interpretation of the optimum functions presented in the discussion of Figures 5.4 and 5.5. Since the first optimum function is primarily concerned with which atomizing steam boiler is operating, the separation in the first coefficient is due to this factor. That is, those points with positive values of the first coordinate (abscissa) are operating on steam boiler A and those with negative values of the first coefficients are operating on steam boiler B. Figure 5.3 shows that approximately 50% of the entire variation in the data is due to the inclusion of both steam boiler A and steam boiler B cases in the learning base. Thus, the problem of deriving the detection algorithms could be greatly simplified by only using one of the two steam boilers. Including all seven of the possible boiler configurations in the analysis would add a great deal of additional irrelevant variation to the problem and therefore the division of the problem into seven similar problems to cover the various configurations of the boilers has probably improved the algorithms performance at the expense of requireing additional algorithms to implement the system. Although further simplification would be introduced by dividing it into 14 instead of 7 problems and only considering a single steam boilers operation, this further reduction would probably limit the number of learning cases for any configuration to the point that for the present data set, it would be extremely difficult to derive the diagnostic algorithms, although it appears that one should be able to derive detection algorithms. Since the performance of the algorithm including the variation from steam boiler A to steam boiler B will be shown to be satisfactory, it is recommended that the detection algorithm be developed including this additional variation.

The variation in the second coefficient (i.e. the ordinate) was due almost entirely to how hard the system was working. The larger the value of the second coefficient the harder the system is working. Thus, points located near the bottom of Figure 5.9 represent cases where the load is very light and points located near the top of Figure 5.9 represent cases where the load is very heavy. It should be recalled that the first optimum function presented in Figure 5.4 is dominated by the definition of which atomizing steam boiler is operating; however, that function also contains information regarding the load. In fact, those parameters regarding the load such as the return temperature of zones 1 and 2 and the operating pressure of the atomizing boiler had opposite signs to the corresponding values in the second optimum function. Thus, as the load is increased a data point moves rapidly to the top of Figure 5.9 and slightly to the right. The two groupings of learning data having negative values of the first coefficient represent data histories having different loads but operating with atomizing steam boiler A. Such groupings would be created by separation between cold and hot days or between data histories taken in the rainy season and not during raining season.

In addition to providing the scatter plots of the first two coefficients, the ADAPT programs also provides scatter plots of any two coefficients desired by the analyst. It is standard procedure not only to examine these first two plots, but to examine the remaining combinations. This is done to reveal serious errors in the data recording or keypunching and the presence of any unusual case or group of cases. An example of this occurred on the base which has just been described. Figure 5.10 shows the scatter plot of the coefficients of the fourth and fifth optimum functions for this base. Examination of this plot shows that all of the cases are grouped in the upper right hand corner except for a single case which is in the lower left hand corner. When this phenomenon occurs it is an indication that the isolated case is a very unusual case and, therefore, investigation is required to determine whether the unusual nature of the case is based on a physical characteristic of the case or is due to an error in data recording or keypunching. This review can be assisted by the examination of the optimum functions associated with the coefficients used on the scatter plot in which the case is unique.

In this case, the two optimum functions of interest are the fourth and fifth which are shown in Figures 5.11 and 5.12. Examination of these two figures shows that the important variables to the scatter plot shown in Figure 5.10 are the number of gallons of fuel used, the return temperature for zone 2, the 12-hour average of the supply temperature and the three-day rainfall. The strongest of these is the 12-year average of the supply temperature. Thus, the first step is to review these variables for the case represented by the point in the lower left hand corner of Figure 5.10. Review of the data runs which produced this base shows that point to be the case associated with 2400 hours on April 22, 1971. This was the case associated with the incipient failure of the atomizing steam boiler which occurred at 9:00 on April 23, 1971. Careful examination of this case shows that in the process of transforming the data from the original data sheets to the keypunch instruction sheet, a subtraction was not carried out. This resulted in an error of a factor of approximately 30 in the value of this measurement. Clearly, this case must be corrected. Therefore, this case was deleted and replaced by a correct case, and the base was rederived.

A single erroneous case does not drastically effect the optimum base for representing the data. This can be seen by comparing the information energy and the optimum functions associated with the corrected base with the corresponding information energy and optimum functions for the original base (Figures 5.3 thru 5.5). The corresponding information for the new base is presented in Figures 5.13 thru 5.15. The information energy and first optimum function are essentially identical and thus the physical interpretation discussed earlier does not change. Although at first glance the second optimum function presented in Figure 5.15 looks different from that presented in Figure 5.5, this is not the case. Careful examination of these two figures will show that Figure 5.15 is actually the mirror image of Figure 5.5. Since the ADAPT programs are dealing with directions in the optimum space, the sign associated with these directions has no effect on the optimality of the base. The apparent difference between Figures 5.5 and 5.15 is simply due to the program arbitrarily selecting different sign for the second function. The only effect of this change is to make a corresponding change in the coefficient associated with each history. Thus, in the new base the second coefficient of a history will be the negative of its coefficients in the original base. Clearly, the degree of representation as well as any of the physical information contained has not been altered by this change.

This change manifests itself in the scatter plot by reversing the sign associated with the physical interpretation. Thus, the scatter plot for the new base shown in Figure 5.16 is essentially identical to that shown in Figure 5.9 except the sign of the second coefficient has been changed. Thus, the two groups occurring near the top of Figure 5.9 now occur near the bottom of Figure 5.16. The one group occurring near the bottom of Figure 5.9 occurs near the top of Figure 5.16. The reader can verify that the discussion of the physics associated with the optimum function presented in Figure 5.5 is identical to the physics which would be inferred from Figure 5.15 except that the effect of the sign of the ordinate of the scatter plot is reversed.

The effect of correcting the erroneous case is fairly significant for the fourth and fifth optimum functions which are strongly influenced by the 12-hr. average of the supply temperature. The effect on the higher numbered optimum functions has simply been to shift these functions down by one position in the series since with this error corrected the variation which previously required both the third and fourth optimum functions for representation now requires only a single optimum function. This is illustrated by Figures 5.17 and 5.18 which present the seventh optimum function for the original base and the sixth optimal function for the new base, respectively. Examination of these two figures shows them to be essentially identical. This result is typical of all the optimum functions beyond the fifth term. Figures 5.19 thru 5.22 presented third, fourth, eighth, and nineteenth optimum functions associated with the corrected base. The reader can easily verify that the third optimum function is essentially associated with fuel consumption, the fourth optimum function corrects the fuel consumption for the rainfall, sixth optimum function deals with the oil consumption of the atomizing steam boiler, the eighth optimum function deals with the difference between the 12-hr. average and the instantaneous fuel rate, and the nineteenth optimum function defines a low load condition.

Considerable analysis such as that which has been discussed above can be carried out on each base developed in support of this program. This analysis would lead to considerable understanding of the important factors and mechanisms controlling the operation of the KSC central heat plant. The preceding discussion has given an example of how the representation can be used to get a better understanding of the system, and to provide an insight to assist in making decisions regarding the development of the detection algorithm. It is only this latter result

F

of this analysis of the representation which is pertinent to the specific objectives of this study. Thus, a complete analysis of the representation considering the higher order terms for the base presented here, and the other bases developed as a part of deriving detection and diagnostic algorithms is beyond the scope of this study.

The major conclusions with respect to developing the detection algorithms resulting from this analysis of the representation are: 1) more than two terms will be required to develop a detection algorithm, 2) if the derivation of the detection algorithm proves too difficult, it can be significantly simplified by considering only one of the atomizing steam boilers in the derivation of the detection algorithm with the penalty of increasing the number of detection algorithms required, 3) further reduction in the irrelevant variation in the data can be achieved by limiting the algorithm development to certain load conditions, 4) it is probably necessary to develop separate detection algorithms for each of the seven boiler configurations, and 5) the case associated with April 22, 1972, at 2400 hours should be omitted due to an error in keypunching the data.

5.2 Exploratory Analysis

Exploratory analyses were carried out to: 1) select preprocessing to be used for the remainder of the detection algorithm development, 2) project the expected performance of a final algorithm, 3) illustrate the effect of the reduction of the number of measurements on the performance of the algorithm, and 4) to estimate the performance which could be obtained from each of several approaches to deriving the detection algorithm. This section will review the primary results of each of these exploratory investigations.

Preprocessing

The ADAPT programs offer the user several options for preprocessing the data prior to selecting the optimum representation. Certain preprocessing options can be selected based on knowledge of the problem. The preprocessing options considered in developing maintenance algorithms for the KSC central heat plant include the subtraction of the average data vector prior to processing through the representation program and the equalization of the variation in each of the measurements. The subtraction of the average from each data history has the advantage of producing data histories with zero means and of minimizing the irrelevant variation. Except for problems involving extremely unusual situations such as clutter subtraction, a subtraction of the average vector from each data history normally results in easier derivation and therefore better algorithms.

The variation associated with each of the measurements used for the present analysis was approximately equalized by multiplying the measurement value by a constant such that the maximum values of the measurement fell within the same order of magnitude. In some cases, this was also achieved by subtracting appropriate constants from each measurement. Although programs are available within the ADAPT system to provide exact equalization of the variation, the application of these programs can only be justified relative to the simple approach used here when it is known that each variable will have approximately equal influence on the representation. Since this is not the case for the present analysis, the more approximate equalization approach was utilized.

The ADAPT programs also provide options to increase or decrease the significance of spikes in the data histories by preprocessing each variable by either taking its logarithm or raising ten to the power of that variable. These processes may also significantly affect the number of terms required to achieve a given representation. The ADAPT programs also include preprocessing options to carry out a normalization such that the magnitude of each data history, i.e. data vector, is unity.

Raising each variable to the power of ten would accentuate the uncertainty associated with the lack of knowledge concerning the proper variation which should be associated with each of the measurements. Thus, this preprocessing option can also be rejected. However, the log preprocessing and the normalization preprocessing cannot be rejected on an apriori basis. For this reason the initial 29 case exploratory data set made up of 190 variables was utilized to investigate the effect of the log and normalization preprocessing on the performance of the algorithm. The results of this investigation are summarized in the performance map presented in Fig. 5.23. The reference processing without normalization or log preprocessing is shown by the solid line connecting the circles. The effect of normalization is indicated by the solid line passing through the square symbols. As can be seen the effect of normalization can be expected to be very small, and for this particular case, a slight reduction in performance is observed. Based on these results it can be concluded that normalization will probably have an insignificant effect on the performance of the detection algorithms. Noting that the normalization has the disadvantages of slightly increasing the complexity of applying the algorithm and significantly increasing the complexity of interpreting the physics associated with the algorithm, the decision was made not to normalize the data.

The effect of taking the logrithm of each measurement before processing it through the ADAPT programs is shown by the solid line passing through the triangles. This algorithm has significantly poorer performance than the reference case, and in fact, has such poor performance that one cannot have high confidence in the physical basis of that algorithm. For this reason the log preprocessing was also rejected. Thus, for the remainder of this study, the data used will neither be normalized or nonlinearly distorted and the only preprocessing used will be to subtract the average data history from each data history prior to processing and to approximately equalize the variation associated with each of the measurements by multiplying them by an appropriate constant.

Projected Performance

The performance map shown in Figure 5.3 can also be used to estimate the expected performance of the algorithms on test cases. The process of estimating this performance is illustrated by the dash line which proceeds nearly vertically from the position of the algorithm on the performance map. The slope of this line is based on experience and represents the decrease in performance which will be observed due to the partial overdetermined nature of any algorithms which are near the cross hatched line on the performance map. Experience has shown and analysis has confirmed (see Appendix C) that when the ratio of number of cases to number of dimensions exceeds approximately six, one can expect that there will be no further significant degradation and performance as one moves vertically on the performance map without rejecting useful information. Thus, these exploratory studies indicate that it should be possible to develop a universal detection algorithm with a performance parameter ($\Sigma\sigma/v$) of approximately .67. This corresponds to a probability of error between .05 and .1.

Effect of Number of Measurements Used

The same exploratory data base which was used to evaluate preprocessing was used to illustrate the effect of reducing the number of variables on the performance of the algorithm. The reduction in variables is achieved by examining the relative importance vector for the algorithm and only retaining the most important variables as defined by this relative importance vector. The results of this analysis are also summarized in Figure 5. 23.

The solid symbols show the performance of algorithms developed on the original 190 dimensions, the most important 74 of these 190 measurements, the most important 10 of the 74 measurements, and the most important 5 of the 10 measurements. As can be seen by examination of this performance map these algorithms were developed at considerably different values of the ratio of the number of cases to the number of dimensions, and therefore, a direct comparison of their performance might be misleading. In this particular case, the direct comparison of the performance would give the same qualitative results; however, to obtain meaningful quantitative results the comparison should be between the projected performance of these algorithms. Thus, following the dash lines for the algorithms being considered, one may calculate the projected performance in the same manner as was done for the reference 190 dimensional algorithm. Note that for this 190 measurement algorithm, the projection from either the 13 or 14 dimensional cases is identical. This simply means that in reducing the number of dimensions from 14 to 13 no significant information was discarded. The fact that the 10 measurement algorithm projects to the same point is purely coincidental. Thus, the projected performance parameter for these four algorithms are: .67, .48, .67, and 1.2, respectively. These values correspond to probability of error of approximately .05, .01, .05, and .2.

As pointed out in Section 4, any value of this performance parameter can be associated with a performance trade-off curve of detection probability versus false alarm rate. This curve is also convenient for illustrating the effect of the number of measurements used and is shown in Figure 5.24. Symbols on this figure are identical to those used on Figure 5.23 and correspond to the same algorithms. Figure 5.24 also lists the ten most important variables which were used in the ten variable algorithm. The first five of these ten are the variables which were used in the five variable algorithm. Figure 5.24 clearly shows that initially the reduction in the number of measurements used results in significant improvement in the performance of the algorithm. The improvement in the performance of the algorithm when reducing from 190 to 74 variables is most likely due to the fact that the majority of 116 variables which were deleted apparently contributed very little information to the classification problem and a great deal of confusion to the data analysis. This process was continued and the relative importance vector for the 74 algorithms was examined and the 10 most important of these variables retained and used for the 10 variable algorithm. The performance of the 10 variable algorithm is significantly reduced relative to the 74 variable algorithm. This is due to the fact that the 64 variables which were discarded in this reduction contained enough significant information relative to separating failed from unfailed cases to overcome the confusion loss resulting from using the larger number of variables. This implies that the optimum number of variables to be used for the analysis of the heat plant lies somewhere between 10 and 190. Continuation of the process to 5 variables clearly would be expected to lead to further reduction in performance and Figure 5.24 verifies this further reduction.

Type of Detection Algorithm

There are several types of detection algorithms which might be considered as the basis for the demand preventive maintenance system for the KSC heat plant. The easiest algorithm to use but most difficult to achieve would be a universal detection algorithm which would predict the presence of an incipient failure regardless of which boilers were operating, the type of failure, load on the system or mode of operation. The next easiest algorithm to implement would be effectively a series of universal detection algorithms limited to a specific boiler configuration. For example, the universal boiler-1 detection algorithm which is valid for any operating condition or load provided only boiler No-1 is operating. Clearly, a similar algorithm must exist for boiler 1 operating in conjunction with boiler 2, etc. For the KSC central heat plant it was shown in section 3 that there are a total of seven such detection algorithms required. It might also be possible to develop algorithms which would detect only specific failure modes. For example, a detection algorithm might be developed to detect incipient failures of the atomozing steam boiler. Again the variation should be reduced and the data used to derive this algorithm and thus the derivation should be somewhat simplified. If the algorithm worked exclusively with respect to detecting only the failure on which it was developed, there would not be any need for diagnostic algorithms with this type of detection algorithms. However, it is unlikely that this algorithm would work in this manner since many failures look sufficiently similar that the algorithm developed for detecting one particular type of failure is very likely to detect other types of failures also.

A final method of subdividing the cases to reduce the amount of variation which must be handled in deriving the detection algorithm is to use the ADAPT scatter plot outputs to define natural groupings which have relatively small amounts of variation within the groups. Figure 5.25 illustrates such a selection. This figure presents the scatter plot of all of the available 81 measurement cases. This scatter plot was produced by projecting this data on the first two optimum directions of a 74 measurement base. Group 2 which is included in the solid box on this figure was then selected and used as the learning data set.

If the same base had been used to develop the group 1 detection algorithm as was used for the scatter plot on which the group was selected, one would expect the saving in variation to occur primarily by removal of variation associated with the first two terms of the optimal space for which Figure 5.25 is the projection. However, since this grouping was used on a different base, this reduction of variation occurred not in the first two optimum directions but in the higher ordered optimum directions. This can be seen by comparing Figures 5.26 and 5.27. Figure 5.26 presents the information energy as a function of number of terms used for the base made up of the cases in group 1. The corresponding information energy for all the cases used in the reference 81 measurement algorithm is presented in Figure 5.27. Comparison of these figures shows that the amount of information contained between 2 and 20 terms is significantly greater for the base developed on the group 1 cases.

In order to investigate which of these four types of algorithms were feasible, exploratory algorithms were developed on the 81 measurement base. The 81 measurements used in this base were selected by considering the relative importance vector for the initial exploratory studies of each of these algorithms. The five algorithms which were prepared are: an example of a universal detection algorithm for a single boiler, (i.e. 1) the universal boiler No. 1 detection algorithm); an example of an algorithm to detect a specific type of failure, (i.e. 2) the atomozing boiler failure detection algorithm); two examples of algorithms for detecting failures in certain portions of the system, (i.e. 3) an algorithm for detecting failures occurring in the central heating plant, 4) an algorithm for detecting failures occurring outside of the central heating plant, and 5) the algorithm for detecting the failures occurring in group 2 as defined in Figure 5.25. The performance of each of these algorithms as a function of the dimensionality is summarized in the performance map presented in Figure 5.28. The solid lines through the appropriate symbols represent the actual patch of a given algorithm on this performance map. The performance of the algorithm developed for a ratio of number of cases to number of dimensions between 2.5 and 3 has been projected to determine the expected performance of these algorithms on the test data. Although in a few cases this does not represent the best performance that one can anticipate achieving by full development of the algorithm, it does approximate the performance that one can expect given approximately an equal amount of development effort for each algorithm.

The projected performance of each of the algorithms shown in Figure 5.28 is summarized as a trade-off curve between detection probability and false alarm rate in Figure 5.28. In order to determine whether an algorithm is shown on Figure 5.29 will be useful in the predictive preventive maintenance system, we must establish acceptable limits and false alarm rates and detection probability. In Section 3 a scheme for requiring a total of four indications of a failure before initiating action is outlined which was applicable to both the manual and automated implementations systems. This scheme required that after the first failure was detected no action will be initiated until three more consecutive detections occurred. The false alarm probability ($P_{F.A.}$) for a scheme such as this is given by

$$P_{F.A.} = (P_{F.A.1}) \; (P_{F.A.2})^3 \tag{5.1}$$

and a corresponding detection probability (P_D) is given by:

$$P_D = (P_{D1}) \; (P_{D2})^3 \tag{5.2}$$

where: P_{D1} = the detection probability for the initial detection of a fault

P_{D2} = the detection probability for each of the subsequent evaluations

P_{FA1} = false alarm probability for the initial detection

P_{FA2} = false alarm probability for the subsequent detections

Thus, if one considers that the initiation of the further analysis required to evaluate three consecutive faults is acceptable once every ten days, one may select P_{FA1} = to 0.1. If we desire approximately one false alarm per year with respect to initiating maintenance action, one should select P_{FA2} = .3, yielding an overall false alarm rate of .003 or approximately one per year. Examining Figure 5.29 we see that P_{D1} corresponding to P_{FA} = .1 is .94 for the universal detection algorithm and the P_{D2} corresponding to P_{FA} of .3 is .98. Substituting these into equals 5.1 and 5.2 we find that the overall detection probability is .88. That is the scheme outlined will detect 88% of the failures with only one false alarm per year.

A similar analysis for the other three algorithms shown in Figure 5.29 yields a detection probability of approximately .97 for either the algorithm to detect field problems or the algorithm for detecting scatter plot Group 2 failures. The same scheme applied to the algorithm for detecting failures in the atomizing boiler gives a detection probability of .93. Examining Figure 5.29 we can see the advantage of this multiple application of the algorithm. In order to achieve the same false alarm rate with a single application of these four detection algorithms, one would have detection probabilities of approximately 0.55, 0.87, 0.85, and 0.73. If one uses the multiple application result any of the four detection algorithms give performance which would be satisfactory for implementation of the predictive preventive maintenance approach. Since the development of universal detection algorithms for the seven boiler configurations is considerably less expensive than any of the other approached tried, it is believed that for the KSC heat plant application this is the best algorithm. There are other algorithms with better performance and for other applications these might be desirable. However, if the expected performance of the universal boiler No. 1 detection algorithm can be achieved and proven, feasibility of the predictive preventive maintenance approach will have been established in a mode utilizing a relatively straight forward detection scheme.

The relative importance vectors which go with these five algorithms are presented in Figures 5.30 to 5.34. These relative importance vectors show the importance of each of the 81 measurements to the decision being made by each of the five algorithms. The importance is measured by the absolute magnitude of the relative importance vector corresponding to the index number associated with the measurement as defined by Table 5.4. Thus, examining Figure 5.30 we see that measurements No. 3, 15, 24, 43, 48, 49, and 71 are quite important to the universal detection algorithm for Boiler No. 1 failures. Reference to Table 5.4 shows that measurement No. 3 is the rainfall during the past hour. Measurement No. 15 is steam pressure and atomizing Boiler A, measurement No. 24 is the supply temperature, etc. Examination of Figures 5.30 thru 5.34 show that amount of rainfall during the past hour is important to the universal and field problem detection algorithms, whereas the longer period rainfall is more important for the Group 2 detection algorithm. The change in the water flow through Boiler No. 1 is important to the universal Boiler No. 1 detection algorithm, the algorithm for detecting in-plant failures, extremely important to the algorithm for detecting field failures. It is insignificant for the algorithm for detecting failures in Group 2. Clearly, analysis of these relative important vectors can provide a basis for understanding of how each of these algorithms works and how one should approach the problem on imporving the algorithm. Since the universal detection algorithm has been selected as the recommended approach for the KSC heat plant, the further development of these detection algorithms will be illustrated in the next sections using this algorithm as the example.

5.3 Optimization of Universal Detection Algorithm

The exploratory studies have answered the question as to what preprocessing should be used and which algorithm should be developed. We have also seen as part of these exploratory analysis that the number of measurements used can have a significant effect on the performance. One of the steps is optimizing the algorithm is to selectively reduce the number of measurements used by examination of the relative importance vector. Another major step is to select the number of dimensions to be used for the algorithm. Decision must also be reached on exactly how the algorithms are to be applied and as to whether the validity criteria should be applied with the algorithm. Some of the process of reducing the number of measurements was accomplished in the exploratory analysis. As pointed out in Section 5.1 initial analysis was performed on 192 measurements. This 192 measurement base was reduced to an 81 measurement base for the exploratory analysis. The 81 measurements were selected as 81 measurements pertinent to all five of the algorithms investigated in the exploratory analysis. Now that the analysis has been reduced to a single algorithm, one can be more selective and select only measurements pertinent to this single algorithm. This was done and the resulting 50 measurements which were selected were used to formulate the base which has also been presented in Section 5.1. It is instructive to compare the effect of this reduction from 192 measurements to 50 measurements on such things as the variation, the scatter plot, and the optimum functions. The information energy for the 50 variable base is presented in Figure 5.13. The corresponding 81 measurement base was presented in Figure 5.27. Figure 5.35 presents the information energy for the original 192 measurement base. This figure confirms the behavior that as one decreases the number of measurements, the representation becomes easier and the number of dimensions required to explain a given amount of the information decreases.

The simplification of the representation is displayed dramatically by the corresponding scatter plots. Figure 5.36 presents the scatter plot for the original 192 measurement base. This should be compared with the scatter plot for the 50 measurement base presented in Figure 5.16. The reader notices immediately that in the 50 measurement base there are three very tight distinct groups as compared to a relatively large scattering of groups occurring on 192 measurement base. Thus, we see as the number of measurements have been reduced we have been able to find a representation in which the definitions of the natural groups have become more precise. Comparison of the first and second optimum functions given in Figures 5.37 and 5.38 for the 192 measurement base with the optimum functions presented in Figures 5.14 and 5.15 for the final 50 dimensional base again shows that the reduction of the measurements has modified the representation. In 192 dimensional base the first optimum function is effected by a great number of variables. The first variable in Figure 5.37, the day of the year contributes considerable variation to the data; but when the relative importance vectors were analyzed, it was shown to be relatively insignificant to the detection problem. The day of the year was therefore omitted from the 50 measurements selected for use in the final algorithm development. Thus, although the first optimum function is dominated by the atomizing steam boiler for both the 50 and 192 dimensional algorithms, the domination is more complete for the 50 measurement base.

The second optimum function shows that in the 192 measurement base the atomizing boilers are still significant contributors indicating that the first optimum function was not adequate to completely explain the interrelations of this characteristic with the other measurements. Figure 5.15 implies that the atomizing steam boiler measurements are no longer the dominant measurements for the 50 measurement base. Thus, for the 50 measurement base the first optimum function was able to explain a much greater percentage of the interaction of the atomizing steam boiler with the other variables.

Examination of the relative importance vectors has allowed us to reduce the number of variables used in deriving the algorithm. The next question is that of dimensionality which should be used for the detection algorithm. Since the number of learning cases was limited to approximately 100 by the availability of usable data, the maximum dimensionality which one can consider will be of the order of 40 to 50. Thus, the initial processing was performed using 40 dimensions and the resulting relative importance spectra is shown in Figure 5.39. This figure shows that dimensions 38 and 40 made significant contributions to the performance of the algorithm. The 28th dimension was the next significant measurement and the next was the 19th dimension. As discussed in Section 4 this effective dimensionality displays itself dramatically on the performance map. The performance map for this algorithm is shown in Figure 5.40.

The trace of this algorithm on the performance map passes through the three points designated by the squares representing dimensionalities of 40, 29, and 20 at which the algorithms were developed. The 29 and 20 dimensional cases were selected by examination of Figure 5.39 which indicated that these two algorithms would be near break points in the path of the algorithm along the performance map. Comparison of Figures 5.39 and 5.40 illustrates that the quantitative significance of the relative importance vector on the performance map is greatly distorted by the nonlinearities involved. Thus, even though the greatest importance fell in the 38th optimum direction one sees little difference in the projected perperformance of the 40 dimensional and 29 dimensional algorithm. On the other hand, there is considerable difference between the projected performance of the 29 and 20 dimensional algorithms as indicated by the dash lines passing through these algorithms. As previously pointed out once the projected performance in terms of the performance parameter, $\Sigma\sigma/V$ has been determined, one can make an estimate of the expected detection probability versus false alarm rate. This has been done for the 20 dimensional algorithm and for a compromise between 29 and 40 dimensional algorithms. This compromise was used since it is felt that the projection is not sufficiently accurate to account for the differences between the 29 and 40 dimensional algorithms. The resulting performance trade-off curves are presented in Figure 5.41.

It is interesting to note that the projected performance for the 20 dimensional algorithm is exactly the same as that which was projected for the 81 measurement algorithm in the preceding section. This implies that the 20 dimensional algorithm is based on the same physical principals as the algorithm which was being investigated in the exploratory studies using 81 measurements and 30 dimensions. Thus, we see that the reduction of the number of measurements from 81 to 50 has allowed us to achieve the same performance with approximately 10 less dimensions.

The projected detection probability versus false alarm rate again allows us to evaluate the expected detection probability using the multiple application schemes which are being considered for this demand preventive maintenance approach. Applying equations 5.1 and 5.2 to the data presented in Figure 5.41 we see that since the predicted performance for the 20 dimensional algorithm is identical to the performance for the previous 81 measurement algorithm, the detection probability remains at 88%. For the higher dimensional algorithm the projective detection probability associated with a false alarm rate of approximately one per year is 96%. As before both of these detection probabilities are acceptable. The lower the dimensionality the less likely will it be to employ the validity criteria successfully use the algorithm. For this reason, there is some advantage in not having to use the higher dimensional algorithms. However, the determination of this requires an analysis of the proof test cases which will be presented in the next section. Therefore at this point we shall consider both the 29 and the 20 dimensional algorithms as candidate algorithms for the detection required to implement the demand preventive maintenance scheme for the KSC central heat plant. It should also be pointed out that once the specific false alarm rate has been selected, the Fisher weighting parameter, see Appendix C, provides a way in which one may increase the algorithm performance even more. Since the present performance of both algorithms is adequate and the final selection of the false alarm rate is not advisable at this stage of the program, this additional optimization was not employed. Its potential effect on the shape of the trade-off curves is illustrated in Section 6.0.

Figures 5.42 thru 5.44 present the relative importance vectors for the 40, 29, and 20 dimensional algorithms, respectively. Examination of these three relative importance vectors shows that there is a great deal of similarity between the 40 and 29 dimensional algorithms. This would have been expected by inference from their similar performance as shown on the performance map. Although there are some significant differences between the 20 dimensional algorithm and the 29 dimensional algorithms there are also many significant similarities. For example, measurement No. 29, the average fuel temperature is important to all three algorithms. In contrast measurement No. 49, days since preventive maintenance of the portable boiler only appears significant to the 40 dimensional algorithm. The fact that this variable is not significant for the 29 dimensional algorithm is a reasonably strong indication that this is a fortuitious match rather than a physically connected to the detection of incipient failures. The agreement of this result with physical intuition is further justification for considering the 29 dimensional algorithm rather than the 40 dimensional algorithm.

5.4 Algorithm Evaluation

The discussions presented in the preceding sections have shown that the ADAPT programs provide a relatively complete evaluation of the algorithms as a by-product of their derivation. In this section we shall present a more conventional evaluation of the detection algorithm which will show that the estimate of the expected performance of the algorithm provided by the ADAPT learning process is valid. This evaluation will also provide additional confidence in the ability of the ADAPT derived failure detection algorithm to perform the detection of incipient failures required to insure the feasibility of the implementation of the predictive preventive maintenance system.

The evaluation presented in this section will consist of the results of testing approximately 200 independent test cases against the universal detection algorithm. These independent proof test cases may be considered as belonging to one of three groups: 1) test cases expected to be similar to the learning cases, 2) test cases obtained under significantly different conditions than the learning cases, and 3) those cases containing errors. The proof test cases obtained under essentially the same conditions as the learning cases include all those cases obtained for the same operating configuration, i.e., boiler #1 operating by itself, and over a time period during which the design of the heating plant and distribution system was the same as during the learning period. The learning data was obtained essentially between May of 1970 and the end of 1971. The minor changes in the distribution system which occurred about August 1971 should not invalidate any test cases. On the other hand, the major changes in the distribution system which occurred early in 1970 can be expected to have a significant effect on the performance of the algorithm. Similarly, the change in configuration from boiler #1 operating alone to boiler #2 operating alone would also be expected to have a significant effect on the performance of the algorithm. The testing which will be presented in this section will resolve the questions associated with the impact of these variations on the performance of the algorithm.

One of the major problems associated both with obtaining adequate learning data and with performing proof test evaluation of this data is the availability of high confidence truth data. In this case, the truth data is the actual identification of the date and time of each of the failures and the insurance that those cases selected as failure free are indeed free of failures. This information was generally obtained by examination of the P.M. and work order records, a summary log kept by Mr. Guggenheim, and the plant log. The most useful single piece of information was the summary log kept by Mr. Guggenheim, which appeared quite adequate for identifying the date on which failures occurred. The determination of the time of failure often required additional detective work. One technique that proved quite effective was to examine the plant log for the time at which the failed component was replaced by a redundant system.

In general, errors in the truth data are more critical with respect to the evaluation than to the derivation of the algorithms. Considerable experience with ADAPT programs have shown that a few incorrect learning cases usually have a relatively minor effect on the derivation of the algorithm. However, a few incorrect learning cases can have a significant effect both on the performance evaluation resulting from the ADAPT analysis of the learning data and on the performance analysis resulting from the evaluation of independent proof data. In general errors where the existance of a failure is not detected are more likely. This might occur if: 1) the failure were considered sufficiently insignificant that it would not be eported to Mr. Guggenheim for inclusion in his log, 2) the maintenance would be carried out by the maintenance crew without the preparation of the work order records, and 3) the performance of the maintenance is omitted from the plant log. If these three events occurred, there would be no way to determine that there had been a failure on that day. The other type of truth data failure, (i.e. a day is identified as having a failure when a failure did not actually occur) should only result when the occurrence of a failure is recorded on the wrong day. In this case it is likely that this particular failure would be recorded in an inconsistent manner between the three sources of failure information. Several cases of such inconsistency were noted and these cases were not used either in the learning data or in the evaluation.

The algorithms were applied to the test data using the procedures recommended for the manual application of the algorithms. The universal detection algorithm tested are the algorithms presented in Table 2.1, 5.3 and 5.4. These tables list the equations for the dot product of a data vector with the corresponding algorithm values which have been associated with the index parameters listed in the tables. The name of each of the index parameters is summarized in Table 2.2 for the 50 component data vector. The names for the 192 component variables which include the 50 variables are summarized in Table 2.6. To simplify the relationship of the names in Table 2.2, Figure 5.45 presents a heating plant log sheet with the hourly value spaces replaced with the associated name and number of the variable.

In some cases it is desirable not only to apply the algorithm but also to apply the validity criteria which were discussed in Section 4. The validity criteria consist of a series of dot products of the same format as the algorithm itself which result in values that must be squared and summed and then compared with the minimum acceptable value. The appropriate vectors to use for the application of these validity criteria as well as a more detail description of the procedure to be used are summarized in Appendix D.

The proof testing consisted of the application of the universal detection algorithms and their associated validity criteria to the 200 test cases. The two algorithms which should perform best are the 29 and 20 dimensional universal detection algorithms. The performance of these two algorithms on the 207 test cases is summarized as a function of the particular type and use of test case in Table 5.6. The detail implication of these results may be considered for each of the two types of testing performed.

Proof Testing

The proof testing of the algorithm on cases obtained under essentially the same conditions as the learning data was performed on a total of 79 good cases and 94 cases of failures including failures in the atomizing steam boiler, sludge in the fuel tank, boiler #1, flue gas leaks, and combinations of these with field failures. The projection of these 173 cases on the scatter plot of the first two optimal directions for the 50 variable base is presented in Figure 5.46. Comparison of this figure with Figure 5.16 shows that all of the test cases fall within the region of the scatter plot of the learning data. Furthermore, the test cases can be associated with the same three groups of data which were observed in Figure 5.16. This proof test data included variations in both time of day and day of year relative to the learning data. However, there were no variations in either the operating or design configurations of the heating plant.

Figure 5.47 shows the projection of the 94 failed cases on this same scatter plot. Again, one observes that all of the failure cases fall within the region defined by the learning data, and also conform to the same three groups. In Figure 5.47 the atomizing boiler failures are designated by symbols 1 and 5, the sludge in the tank by the symbol 2, the boiler #1 failures by the sumbols 3 and 6, the flue gas leak by the symbols 7 and 4 and the field failures and combination field and other failures are the symbols 8 and 9.

Table 5.6 tabulates the performance of the algorithm on these test cases as a function of failure mode for a threshold of zero. This zero threshold as discussed under algorithm design has been set for a false alarm rate of one in ten. Both algorithms approximate this performance with the 20 dimensional algorithm missing approximately 5 of the good cases to give a false alarm rate of 0.6 out of 10 and the 29 dimensional algorithm missing 9 of the good ones for a false alarm rate of approximately 1.1 out of 10. The detection of the various failures varies both between algorithms and between failure types. Both of these variations are significant.

The performance tradeoff curves which were discussed as part of the description of the learning data performance evaluation which ADAPT performs can also be used to evaluate the performance on test data. The use of these performance tradeoff curves for this evaluation shows how the algorithm will perform over a wide range of false alarm rates and detection probabilities. Figure 5.48 presents such a curve for the testing of the 173 proof test cases on the 20-dimensional algorithm. The dashed line presented on Figure 5.48 is the projected performance taken directly from Figure 5.41 and represents the performance estimated by the ADAPT programs from the learning data.

There are several ways in which the test data may be presented on a detection probability versus false alarm rate curve. The simplest and most common is to select a threshold, count the number of detections and the number of false alarms and calculate the detection probability and false alarm rate. One may then change the threshold and repeat the procedure obtaining another point on the curve. The entire tradeoff curve may be constructed in this manner. To evaluate arbitrarily small false alarm rates would require an infinite number of cases. This is impossible and the range over which the detection probability and false alarm rate curve can be compared with actual test data by the simple method is limited by the number of test cases. This s illustrated by the cross hatched regions in Figure 5.48. The data points surrounded by a square in Figure 5.48 were obtained as just described for different thresholds. The cross hatched region represents the area over which it would be possible to place a given data point due to the uncertainty in the detection probability and false alarm rate resulting from a finite number of cases being used in the evaluation. Clearly this region becomes larger as one approaches the smaller false alarm rates or as one approaches the detection probability of 1. In the particular presentation of Figure 5.48, this uncertainty is not visible near detection probabilities of 1 due to the fact that there is very little space involved in the region between a detection probability of .99 and 1.0. Examination of the cross hatched area shows that for the 173 test cases the regions of uncertainty become quite large for false alarm rates, less than approximately .03. Comparison of both the cross hatched regions and the data points surrounded by a square with the dash or projected performance shows that in the region that one has certainty in the test performance and also in the region of interest which for the present application has a false alarm rate of .1, the agreement between the ADAPT projected performance and the actual test cases is excellent.

Another way in which the test data may be used to project the performance is to assume that the distribution of the values produced by the algorithm is Gaussian. This assumption is also made in the ADAPT projection of the learning data to a performance tradeoff curve. The form of the ADAPT algorithms, see reference 9, provide an argument for the applicability of the central limit theorem and therefore for the existence of a Gaussian distribution of the algorithm values. If one assumes this distribution, then one may use the test data to calculate the mean and standard deviation. The tradeoff curve for detection probability versus false alarm rate is then constructed from this information. This curve is represented by the line interrupted by plus signs on Figure 5.48. If the algorithm value had a Gaussian distribution, it is also possible to estimate the confidence level in any detection probability or false alarm rate as a function of the number of test cases used. If these confidence levels are applied to the false alarm rates associated with the test data curve shown in Figure 5.48, one obtains the 95% confidence limits shown by the solid lines passing through the circles on this figure. Again we see that the performance of the 95% confidence band and the projected performance from the ADAPT learning data are in reasonable agreement. The conclusion that the central

6

limit theorem suggests a Gaussian distribution cannot be supported at false alarm rates which are of the order of the reciprocal of the number of test cases or smaller. Thus, we must conclude that one cannot evaluate the performance of the algorithm for false alarm rates of the order of the reciprocal of the number of test cases and less. The evaluation of these algorithms can only be considered firm for the region of false alarm rates of approximately 0.1 to 1.0. However, for the application to the KSC heat plant a false alarm rate of 0.1 combined with multiple applications will yield false alarm of the order of 1 per year. Thus, the number of tests carried out are sufficient to verify both the ADAPT projected performance based on the learning data and the feasibility of the predicted preventive maintenance system.

The projection of the test results on Figure 5.48 under the assumption of a Gaussian distribution is also suspect due to the non-uniformity of the performance of the algorithm as a function of particular failure type. For example, the failures associated with combinations of problems are detected significantly less well than the failures associated with boiler #1 or failures in the atomizing boiler. This phenomena will result in a multi-model distribution function which cannot be described accurately by a Gaussian distribution function. Thus, the best method of evaluating the performance is the direct calculation of the detection probability and false alarm rate by counting the false alarms and detections actually observed on the test data as one changes the threshold. This information which was presented as the data point surrounded by a square and the cross hatched area on Figure 5.48 has been summarized in Figure 2.2 where it is again compared with the projection from the learning data. In Figure 2.2, the triangular symbols for false alarm rates greater than 0.1 correspond to the square symbols on Figure 5.48. The triangles for false alarm rates between .03 and 0.1 are the centroid of the corresponding cross hatched areas shown on Figure 5.48.

Figure 5.49 shows a figure corresponding to Figure 2.2 for the 29 dimensional algorithm. Again the dashed line represents the projected performance as taken from Figure 5.41 for the 29 dimensional algorithm and the solid triangels represent the performance of the 29 dimensional algorithm on this test data. Comparison of this projected performance and the actual performance shows that the 29 dimensional algorithm performs considerably poorer than was projected from the learning data by the ADAPT programs. However, when the ADAPT validity tests were utilized in conjunction with this algorithm, its performance proved to be quite similar to that which was projected using the learning data. Thus a comparison of Figures 2.2 and 5.49 suggests that the 20 dimensional algorithm can be used at least for cases obtained on the conditions similar to the learning data without the necessity of performing the validity criteria test where as the higher performing 29 dimensional algorithm should only be used in conjunction with the ADAPT validity criteria. This conclusion is further supported by the remaining evaluation tests carried out using the data obtained on conditions which differ significantly from those under which the learning data was obtained. Thus, this aspect of the performance will be discussed after the discussion of the remainder of the evaluation tests.

Evaluation Testing

Fifteen cases were obtained and tested in which the boiler #1 algorithm was applied to the boiler #2 operation as if the boiler #2 had been boiler #1. In addition, 19 cases were obtained from the period prior to May 1970 when the hot water distribution system was considerably different from that used when the learning data was obtained. Table 5.7 shows the average representation or validity criteria for these two cases as compared to both the learning and proof data. We see that in the case of all three dimensionalities considered, the test cases obtained before May of 1970 are represented more poorly than either the learning data or the test cases obtained assuming that boiler #2 is essentially the same as boiler 1. The representation, assuming that boiler 2 is essentially the same as boiler 1, is also reduced somewhat from the proof test and learning data. It is considerably better than for the cases before May of 1970. This indicates that the assumption that boiler 2 is essentially equivalent to boiler 1 is less severe than the assumption that the configuration before May 1970 was the same as after May of 1970. This is entirely consistent with the performance as indicated by Figure 5.7 or the applicability of the algorithm summarized in Table 5.5. Table 5.5 shows that the assumption that boiler #2 was essentially the same as boiler #1 for the limited number of 15 test cases had no significant effect on the performance of the algorithms since it identified all 15 cases correctly. However, in the case of the assumption that the configuration prior to May of 1970 was identical to the configuration after this time period resulted in only 13 of the 19 being correctly identified for the 20 dimensional algorithm and only 10 of the 19 for the 29 dimensional algorithm. Furthermore, the majority of the errors were in the direction of predicting failures when there was no failure. In other words, the radical change in this configuration of the distribution system made the data appear as if there were a failure in this system. This is entirely consistent with the physical results that since one considered leaks in heat exchangers and other field problems as failures that any unusual (relative to the learning cases) change in the configuration which would radically affect the load for a given operating condition should actually appear as a failure. Thus, the results illustrated by Table 5.7 are not surprising.

In addition to these 34 cases there were 3 cases in which data recording and punching errors were made. All 37 of these cases were represented in the 50 measurement base in order to perform this testing. Eight other cases which were obtained and later found to correspond to inconsistencies in the truth data were also tested. These 45 cases are shown on the scatter plot presented in Figure 5.50. The symbol 1, 3 and 6 represent boiler #2 cases, symbols 2 and 4 represent cases obtained prior to May of 1970, symbols 5 and 7 represent those cases containing errors for which no proof data was available.

Comparison of Figures 5.16 and 5.50 show that symbols 1 and 4 are considerably outside of the range of the scatter plot obtained on the learning data. This again amplifies the fact that these cases are significantly different from the learning data and adequate testing is necessary before one can presume that the algorithm will perform on these cases.

Validity Criteria

In Section 4.4 it was shown that the major problem associated with applying the ADAPT validity criteris is the determination of the threshold to be applied to the representation. Some estimate of a usable threshold for the algorithms discussed thus far can be obtained by analysis of the performance of these algorithms as summarized in Tables 5.5 and 5.6 as compared to the representation summarized in Table 5.6. Changes relative to the learning data as great as that resulting from the assumption that the data obtained before May 1970 was the same as the data obtained after May of 1970 were sufficient to cause significant errors in all algorithms. On the other hand, the assumption that Boiler 2 and Boiler 1 were identical was not sufficient to cause large reductions in the performance of the algorithm. Furthermore, the performance of the algorithm on the proof test data was significantly reduced for the 29 and 40 dimensional algorithms but not for the 20 dimensional algorithms.

Table 5.6 shows that for the 20 dimensional algorithm the representation of the before May 1970 cases had a mean value of 83-1/2% with the standard deviation of 6%. This implies that if one were to choose a validity criteria that the representation must exceed 83.5%, 50% of the invalid cases would pass the validity criteria. This is too great a percentage of invalid cases. If the distribution function for the representation (Q) is Gaussian, then the addition of one standard deviation (i.e. 6% to this validity criteria yielding a value of 89.5% would imply 30% of the invalid cases would pass. However, examination of a properties of the representation shows that the distribution function must be very different from Gaussian, and in fact, one standard deviation for this value of an average representation will allow even less than 30% of the invalid cases to pass. Thus, a reasonable validity criteria for the 20 dimensional algorithm is a representation of 89.5%. Examination of the average representation (Q) for both boiler 2 and the proof test data on the 20 dimensional algorithm shows that both of these significantly exceed this minimum requirement. Thus, one would expect the performance on both the proof test and the boiler 2 data to be approximately as good as on the learning data. A similar analysis can be made for the 29-dimensional algorithm. In this case, a representation of 98.5% should insure that the performance will be similar to that of the learning data. Examination of the results for boiler 2 shows that less than ;30% actually meet this requirement and thus the relatively good performance on boiler 2 indicates that the particular out-of-normality of boiler 2 is not of a type which causes significant problems for this particular algorithm. Approximately 30% of the proof test cases do not pass this validity test. This provides reasonable explanation for the degradation in performance of the proof test cases. Apparently, these differences from the learning data were significant for this particular algorithm. Similar analysis on the 40 dimensional algorithm shows that even perfect representation is not adequate for this high dimensional algorithm. This actually is not the case due to the fact that as one approaches representation of 100%, the distribution function approaches the delta function and thus more and more non-Gaussian so that there is always some validity criteria less than 100% which is acceptable; however, in this case it is clear that the validity criteria probably should be of the order of 99.95% and thus very little of the proof test data and almost none of the boiler 2 or before May 1970 data would pass this validity criteria.

These general results are in good agreement with the performance map presented in Figure 5.40. Since in this figure both the 29- and 40-dimensional algorithms are still quite near the random separation region, there is a portion of these algorithms which is not based on the physics of the problem. Therefore, any significant reduction in representation will make some random contributions to the decision statistic. Only when the algorithms approach a ratio of number of cases to number of dimensions of approximately six can one expect a large tolerance on the representation. Thus, the performance illustrated by Figure 5.49 as compared to Figure 2.2 could easily have been anticipated from examination of the performance map of Figure 5.40.

5.5 Implications to Preventive Maintenance

The major implication of the successful incipient failure detection algorithm is that one now has the option to perform maintenance on a demand as well as a schedule basis. The advantages of performing maintenance on a demand basis have been reviewed in Section 3. There are also conditions under which one might find it advantageous to implement both the schedule and demand maintenance systems to complement one another. Examples of cases such as this are those cases where failure during operation is extremely costly such as applications to spacecraft. In this case, one would still desire to perform schedule preventive maintenance to minimize the number of failures occurring during the operation of the system while retaining the demand preventive maintenance system to allow one to switch away from components which are about to fail during the operation. The combination of the scheduled and demand preventive maintenance systems will probably result in a more expensive maintenance system than a simple demand system, but will provide a system with even less likelihood of a catastrophic failure.

If the scheduled maintenance is to be performed either by itself or in conjunction with the demand preventive maintenance, examination of the relative importance vectors such as that presented in Figure 5.44 can assist in improving the scheduling of the maintenance. Referring to Table 2.5, we see that variables 31 thru 49 are the time since the last scheduled preventive maintenance was performed. Noting that the maximum interval between a scheduled preventive maintenance on any item in the learning data is that interval which is currently being used, it follows that the lengthening of any of the intervals for those scheduled preventive maintenance operations which show little importance in failure detection might lead to increased failures.

Thus, it is recommended that those PM items which do not appear significant in the relative importance of the detection algorithms should not be changed. As would be expected, the great majority of the schedule preventive maintenance items are in this category indicating that the existing PM schedule is an effective one.

When a particular item in the PM schedule appears in the relative importance vector as significant with a position value, (i.e. the time since the last preventive maintenance inspection is greater than average), the system has less likelihood of failure. Thus, positive relative importance implies that the schedule should be reduced. Similar reasoning indicates that if the schedule maintenance has a significant negative value, maintenance is not being performed frequently enough and the frequency of the schedule maintenance should be increased. Examination of Figure 5.44 shows that those schedule maintenance items which have positive relative importance values and, therefore, are apparently being performed too frequently include the electrical and mechanical PM on the fuel pump, on the LTW pump, the electrical PM on the chemical pump, and the electrical PM on the cooling tower. This same figure shows that those PM's having significant negative contributions to the detection algorithm include the electrical PM on the makeup feed pump, PM on the sump pump, and the mechanical PM on the cooling tower. This indicates that the frequency of these preventive maintenance operations should be increased.

Examination of both the optimum functions and the relative importance vector suggests that it is desirable to provide additional automatic sump pumps in all manholes in which water collects after rains and to modify the high temperature hot water lines so that they are either insulated from any rain which may fall around them or are removed from areas in which rain water can flow over them. This follows from the fact that the load on the system as indicated by the first optimal function is dominated to a large extent by the rainfalls. This is reinforced by the fact that the one-hour rainfall appearing in the relative importance vector also correlates with the variables defining how hard the heating plant is working. Thus one might conclude that reduction of the heat loss due to the flow of rain water in the system will have two benefits: 1) a significant reduction in fuel consumption due to a decrease in the load on the system and 2) a reduction in the major variation of the data set thus allowing further improvement in the performance of the predictive preventive maintenance algorithms.

Detail examination of the relative importance vector shown in Figure 5.44 also provides insight as to the manner in which this algorithm is working and why it has capabilities to detect incipient failures which are often better than the capability of the operator. The first point to notice is that instead of just using the performance measures available in the instrumentation of the system the algorithm is also making use of external influences such as rainfall and of the time since the most recent maintenance. The consideration of all fifty items simultaneously is something which is beyond the capability of the human mind unless the consideration has been formalized into a specific procedure or at a maintenance algorithm. For example, one might consider that the increase in zone flows would have the same influence regardless of the zone. However, examination of the relative importance vector shows that a decrease in the flow in zone 3 can be compensated for by a corresponding decrease in the flow of zone 2. However, an increase in the flow in zone 2 at the same time that the flow in zone 3 decreases is a strong indication that the system may soon have a failure. The reader is cautioned that this conclusion is only one of a great many combination of events which are considered simultaneously by the ADAPT algorithms and cannot be used by itself as a detection scheme.

The analysis of the effect of the number of measurements presented in Section 5.2 showed that when the number of measurements used drop below from 10 to 74, there is a significant decrease in the ability to predict incipient failures. Even as small a set of measurements as 10 will lead to a great many interactions when one considers that the threshold on each measurement should be different depending on the value of each of the other measurements. It is this complex interaction between these measurements that requires an algorithm such as those derived by the ADAPT programs to insure proper interpretation of past experience.

TABLE 5.1

LEARNING CASES FOR UNIVERSAL BOILER 1 DETECTION ALGORITHM

Failure Free Cases

Date	No. Cases
May 1970	2
July 1970	1
August 1970	4
December 1970	1
April 1971	3
May 1971	17
August 1971	2
October 1971	10
November 1971	10

Failure Cases

Mode	Failure Date	No. Cases
Field Failures	5/29/70	3
Field Failures	12/4/70	5
Field Failures	6/1/71	2
Field Failures	10/12/71	2
Field Failures	11/9/71	1
Field Failures	12/4/71	3
Distribution Pump	10/27/71	2
Atomizing Boiler	8/11/70	2
Atomizing Boiler	4/23/71	2
Atomizing Boiler	10/21/71	2
Cooling Tower	4/25/71	3
Cooling Tower	5/6/71	2
Sludge in Oil Tank	5/3-5/70	6
& Fuel Valve Prob.	-	-
Forced Draft Fan	5/9/71	3
Boiler No. 1	6/1/71	3
Boiler No. 1	10/12/71	2
Boiler No. 1	10/25/71	1
Flue Gas Leak	11/11/71	5
Plant Shut Down	12/6/70	1

TABLE 5.2

IDENTIFICATION OF 81 INDEX NOS. FOR KSC HEAT PLANT DATA VECTORS

INDEX	DESCRIPTION OF MEASUREMENT	INDEX	DESCRIPTION OF MEASUREMENT
1	Day of Year	35	Oil Inlet Temp. 12 Hr. Avg.
2	Hour of Day x 10^{-1}	36	Oil Gallons Used 12 Hr. Avg.
3	Rainfall During Hour x 10^{2}	37	Steam Pressure 1 12 Hr. Avg.
4	Temperature During Hour	38	Steam Pressure 2 12 Hr. Avg.
5	Av. Rainfall During Past 12 Hrs x 10^{2}	3	Zone Flow Zone 1 12 Hr. Avg.
6	Av. Temp. During Past 12 Hrs.	40	Zone Flow Zone 2 12 Hr. Avg.
7	Days Since Apr. 1, 1971	41	Zone Flow Zone 3 12 Hr. Avg.
8	Boiler A H_2O Flow	42	Supply Pressure 12 Hr. Avg.
9	Boiler 1 Operating Pressure	43	Supply Temperature 12 Hr. Avg.
10	H_2O Inlet Temperature	44	Return Temp. Zone 1 12 Hr. Avg.
11	Fuel Oil Pressure	45	Return Temp. Zone 2 12 Hr. Avg.
12	Oil Temperature Inlet	46	Return Temp. Zone 3 12 Hr. Avg.
13	Oil Discharge Pressure	47	Boiler A Wind Box x 10^{1} 12 Hr. Avg.
14	Used Gal. of Fuel Oil	48	Boiler A Furnace x 10^{2} 12 Hr. Avg.
15	Steam Pressure A	49	Avg. Oil Temp.
16	Steam Pressure B	50	Make-up Feed Meter x 10-^{1}
17	Zone Flow Zone 1	51	Gal. Oil Used in Atom Boiler
18	Zone Flow Zone 2	52	Days Since PM on No. 1 Boiler
19	Zone Flow Zone 3	53	Days Since EPM on No. 1 FD Fan
20	Supply Pressure	54	Days Since PM on Fuel Oil Oper. Tank
21	Return Pressure Zone 1	55	Days Since PM on No. 1 F.O. Pump
22	Return Pressure Zone 2	56	Days Since EPM on No. 1 F.O. Pump
23	Return Pressure Zone 3	57	Days Since PM on Tank Exp. #1 Boiler
24	Supply Temp. (Deviation from Nominal, Index 192)	58	Days Since PM on Pump LTW 1
25	Return Temp Zone 1	59	Days Since PM on Pump LTW 2
26	Return Temp Zone 2	60	Days Since EPM on Pump LTW 2
27	Return Temp Zone 3	61	Days Since EPM on Make-up Feed Pump
28	Boiler A Wind Box x 10^{1}	62	Days Since EPM on Energy Make-up Feed Pump
29	Boiler A Furnace x 10^{2}	63	Days Since PM on Chem. Feed Pump 1
30	Boiler A Outlet x 10^{2}	64	Days Since EPM on Chem. Feed Pump 1
31	Boiler A H_2O Flow 12 Hr. Avg.	65	Days Since PM on Chem. Feed Pump 2
32	H_2O Inlet Temp. 12 Hr. Avg.	66	Days Since EPM on Chem. Feed Pump 2
33	Fuel Oil Pressure 12 Hr. Avg.	67	Days Since PM on Treated H_2O Tank
34	Steam Pressure 12 Hr. Avg.	68	Days Since PM on WTR Softener

TABLE 5.2 (cont'd)

IDENTIFICATION OF 81 INDEX NOS. FOR KSC HEAT PLANT DATA VECTORS

INDEX	DESCRIPTION OF MEASUREMENT
69	Days Since PM on Sump Pump 1
70	Days Since EPM on Sump Pump 1
71	Days Since PM on Sump Pump 2
72	Days Since EPM on Sump Pump 2
73	Days Since PM on Cooling Tower
74	Days Since EPM on Cooling Tower
75	Days Since EPM on Portable Generator
76	Days Since EPM on No. 3 Boiler
77	Days Since EPM on 100 HP Boiler Mobile
78	Days Since PM on Portable Boiler SN2544
79	Rainfall During Last 3 Days x 10^2
80	Rainfall During Last 5 Days x 10^2
81	Rainfall During Last 10 Days x 10^2

TABLE 5.3

ALGORITHM FOR CLASSIFICATION OF NORMAL AND OUT OF TOLERANCE OPERATION OF THE KSC HEAT PLANT (40 DIMENSIONS)

STEP 1 CALCULATE W_T FROM THE VALUES OF THE MEASUREMENTS, V_i^T, AT TIME "T" USING

$$W_T = A_i^{Cl} V_i^T + 139.0$$

WHERE A_i^{Cl} ARE GIVEN AS A FUNCTION OF THE MEASUREMENT INDEX, i, BELOW

i	A^{Cl}	i	A^{Cl}
1	0.4081D-01	26	-0.7716D-01
2	0.2169D-01	27	-0.3733D 00
3	0.1974D-01	28	-0.2763D 00
4	-0.3222D-01	29	0.4600D 00
5	-0.1730D 00	30	-0.1249D 00
6	-0.2802D 00	31	-0.7393D 00
7	-0.2411D-01	32	0.3033D 00
8	0.1290D 00	33	-0.4983D 00
9	-0.3227D 00	34	-0.1790D 00
10	0.4907D-01	35	0.3743D 00
11	-0.1334D 00	36	0.2883D 00
12	0.2467D-01	37	0.3738D 00
13	0.3266D 00	38	-0.2277D 00
14	0.4969D 00	39	0.5692D-01
15	-0.1165D 00	40	-0.6591D 00
16	-0.1937D 00	41	-0.1463D 00
17	0.2690D 00	42	0.2946D-01
18	-0.1713D 00	43	-0.6264D-01
19	-0.9714D-01	44	0.1807D 00
20	0.3145D 00	45	-0.1093D 00
21	-0.2063D 00	46	0.5044D 00
22	-0.1254D 00	47	0.3273D-01
23	0.2621D-02	48	-0.2951D 00
24	0.2307D 00	49	0.2624D 00
25	0.2125D 00	50	-0.9435D-01

STEP 2 - $W_T > 0$ Heat Plant Operation Normal

$W_T < 0$ Heat Plant Operation Out of Tolerance - Apply Diagnostic Algorithms

TABLE 5.4

ALGORITHM FOR CLASSIFICATION OF NORMAL AND OUT OF TOLERANCE OPERATION OF THE KSC HEAT PLANT (29 DIMENSIONS)

STEP 1 CALCULATE W_T FROM THE VALUES OF THE MEASUREMENTS, V_i^T, AT TIME "T" USING

$$WT = A_i^{Cl} \quad V_i^T + 150.0$$

WHERE A_i Cl ARE GIVEN AS A FUNCTION OF THE MEASUREMENT INDEX, i, BELOW

i	A^{Cl}	i	A^{Cl}
1	0.2343D-01	26	0.2232D 00
2	0.1462D 00	27	-0.2135D 00
3	0.2001D-01	28	-0.2196D 00
4	0.4919D-01	29	0.2252D 00
5	-0.1191D 00	30	-0.1043D 00
6	-0.1857D 00	31	-0.3370D 00
7	-0.1552D 00	32	-0.3136D-01
8	0.9544D-01	33	-0.1106D 00
9	-0.2960D 00	34	0.1828D 00
10	-0.1189D 00	35	0.1808D 00
11	0.6109D-02	36	0.2760D 00
12	-0.7488D-01	37	0.2220D 00
13	0.1986D 00	38	-0.6854D-01
14	0.1631D 00	39	0.1979D 00
15	-0.1132D 00	40	-0.1725D 00
16	-0.1935D 00	41	-0.2003D 00
17	0.1076D 00	42	-0.9752D-01
18	-0.5286D-01	43	0.6160D-01
19	-0.1161D 00	44	0.7915D-01
20	0.1659D 00	45	-0.2839D 00
21	-0.1165D 00	46	0.8828D-01
22	0.2184D-01	47	-0.1904D 00
23	-0.1169D 00	48	0.8937D-02
24	0.1844D 00	49	-0.3208D-01
25	0.9435D-01	50	-0.9437D-01

STEP 2 $W_T > 0$ Heat Plant Operation Normal

$W_T < 0$ Heat Plant Operation Out of Tolerance - Apply Diagnostic Algorithms

TABLE 5.5 USE OF 207 TEST CASES

NO CASES	USE
173	PROOF TEST-VARIATION OVER TIME OF DAY AND DAY OF YEAR
15	EVALUATE BOILER 1 ALGORITHM ON BOILER 2
19	EVALUATE EFFECT OF MAJOR CHANGES IN DISTRIBUTION SYSTEM
207	

TABLE 5.6

SUMMARY OF TEST CASES

No. Cases	Use	Class	No. Correct -0 Threshold 20 Dim.	29 Dim
79	Proof Test -	Good	74	70
27	Variation over	Fail - Atom Boiler	25	25
1	time of Day and	Fail Sludge in Tank	1	1
14	Day of Year	Fail Boiler No. 1	13	14
7		Fail Flue Gas Leak	6	7
45		Fail Multiple Probs.	41	35
94		Fail		
10	Evaluate	Good Boiler B	10	10
5	Sensitivity	Fail Boiler B	5	5
4	to	Good Prior 5/70	0	0
15	Change	Fail Prior 5/70	13	10
3	Data Recording & Key Punch Error		-	-

TABLE 5.7

SUMMARY OF REPRESENTATION AND PERFORMANCE

Case	No. Dim.	40	29	20
LEARN				
	$\overline{Q}$	.99977	.9976	.988
	σ_Q	.00018	.0016	.008
	$\Sigma\sigma/V$	.42	.48	.63
PROOF				
	$\overline{Q}$	.9982	.990	.963
	σ_Q	.0011	.007	.027
	$\Sigma\sigma/V$	1.0	.91	.80
BLR 2				
	$\overline{Q}$	.989	.973	.94
	σ_Q	.005	.01	.02
	$\Sigma\sigma/V$	.32	.31	.36
BEFORE 5/70				
	$\overline{Q}$	.980	.925	.835
	σ_Q	.03	.06	.06
	$\Sigma\sigma/V$	.81	1.20	.93

C2

FIGURE 5.1 - TYPICAL DATA HISTORY (AUGUST 14, 1970)

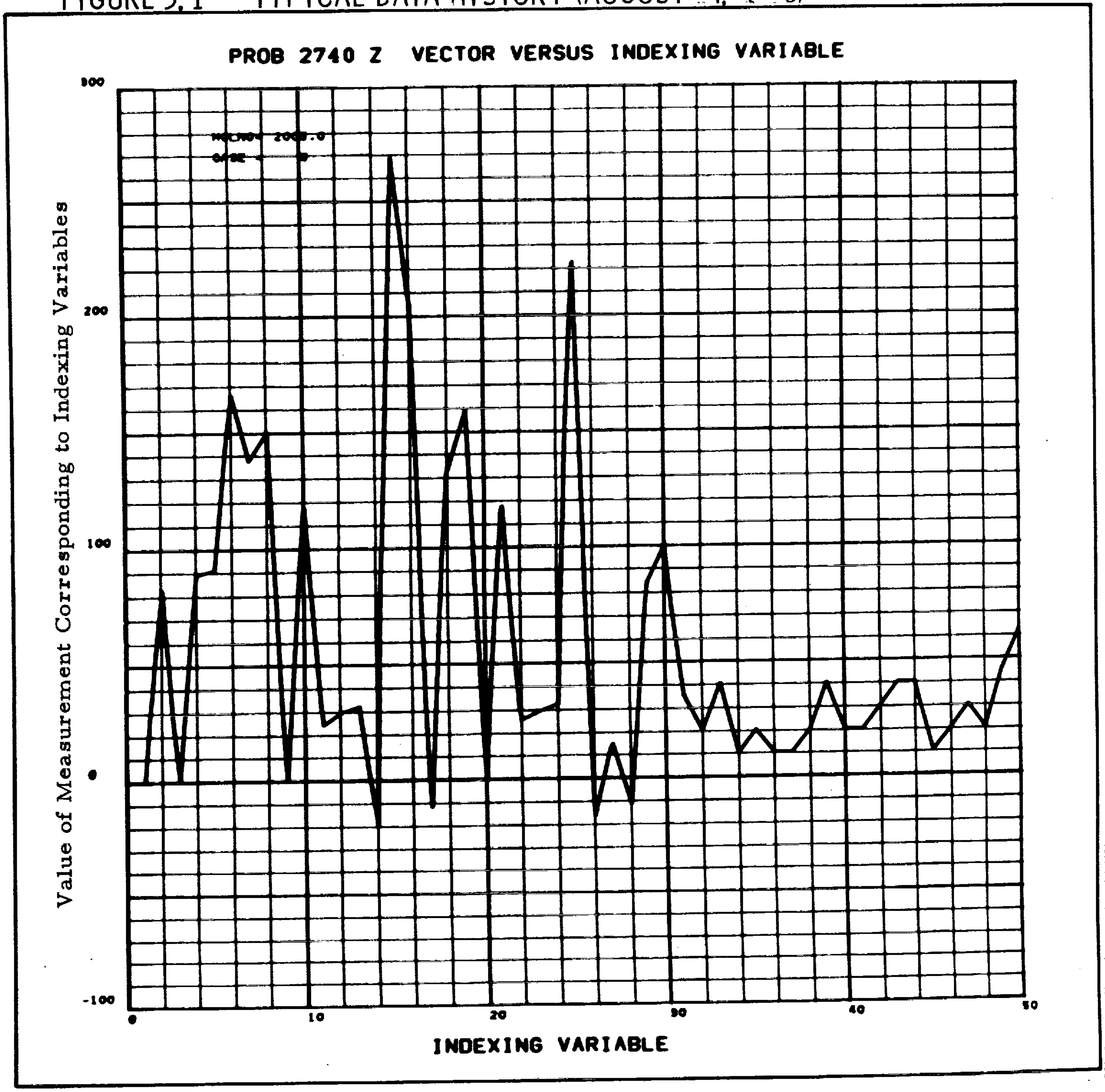


FIGURE 5.2 - AVERAGE INPUT VECTOR FOR 50 MEASUREMENT BASE

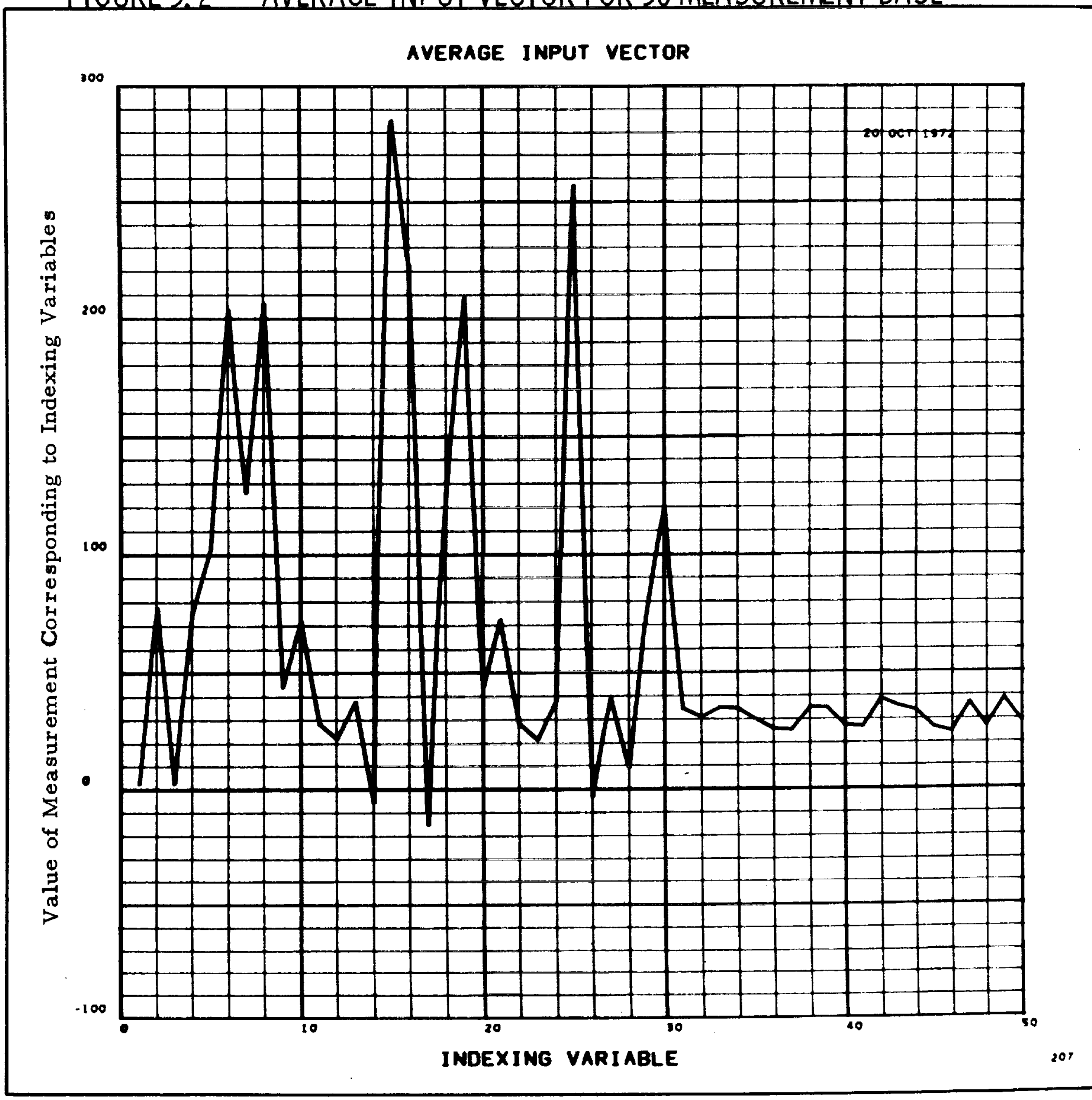


FIGURE 5.3 - VARIATION OF INFORMATION RETAINED WITH DIMENSIONALITY FOR 50 MEASUREMENT BASE

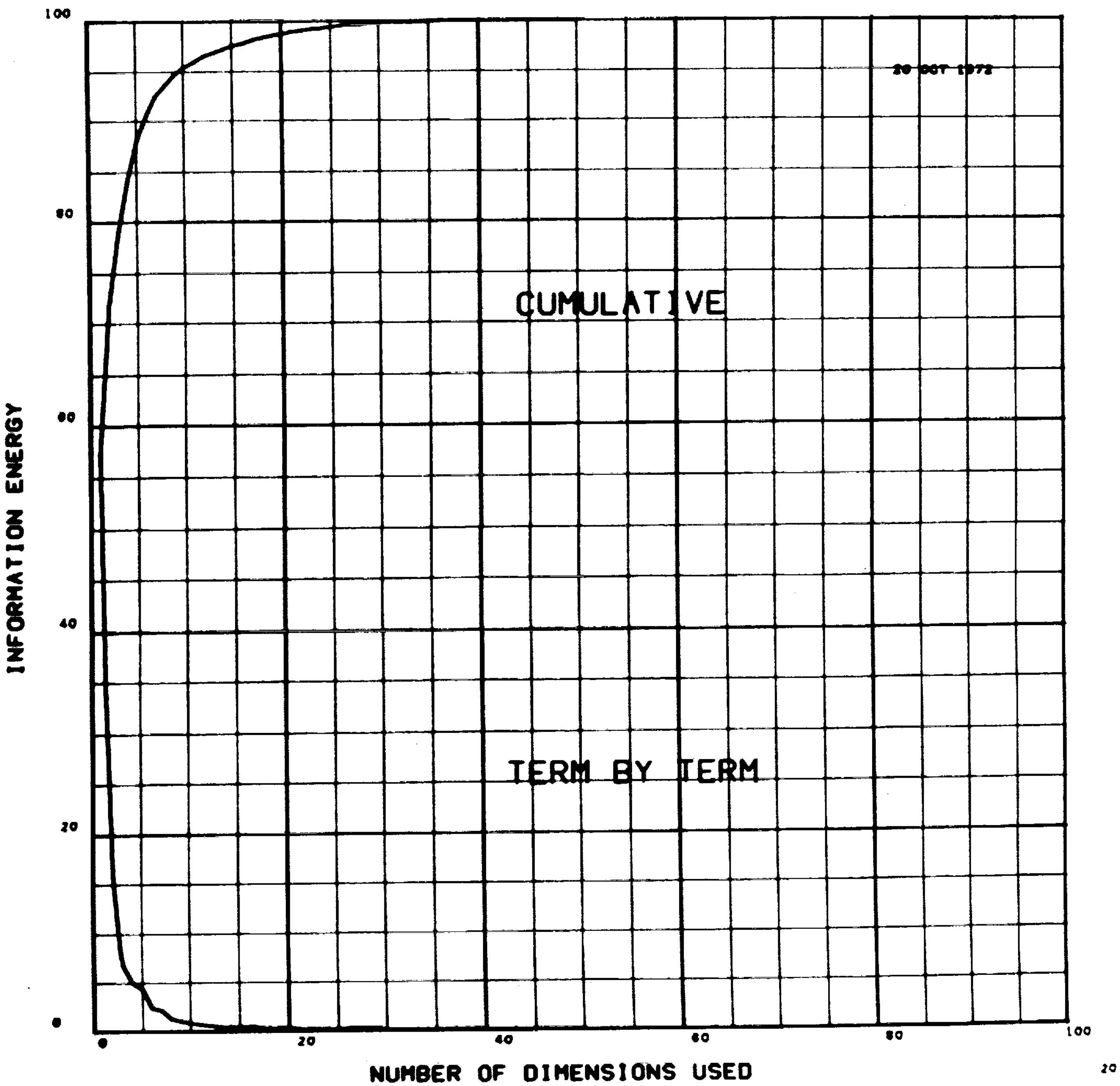


FIGURE 5.4 - FIRST OPTIMUM FUNCTION ORIGINAL 50 MEASUREMENT BASE

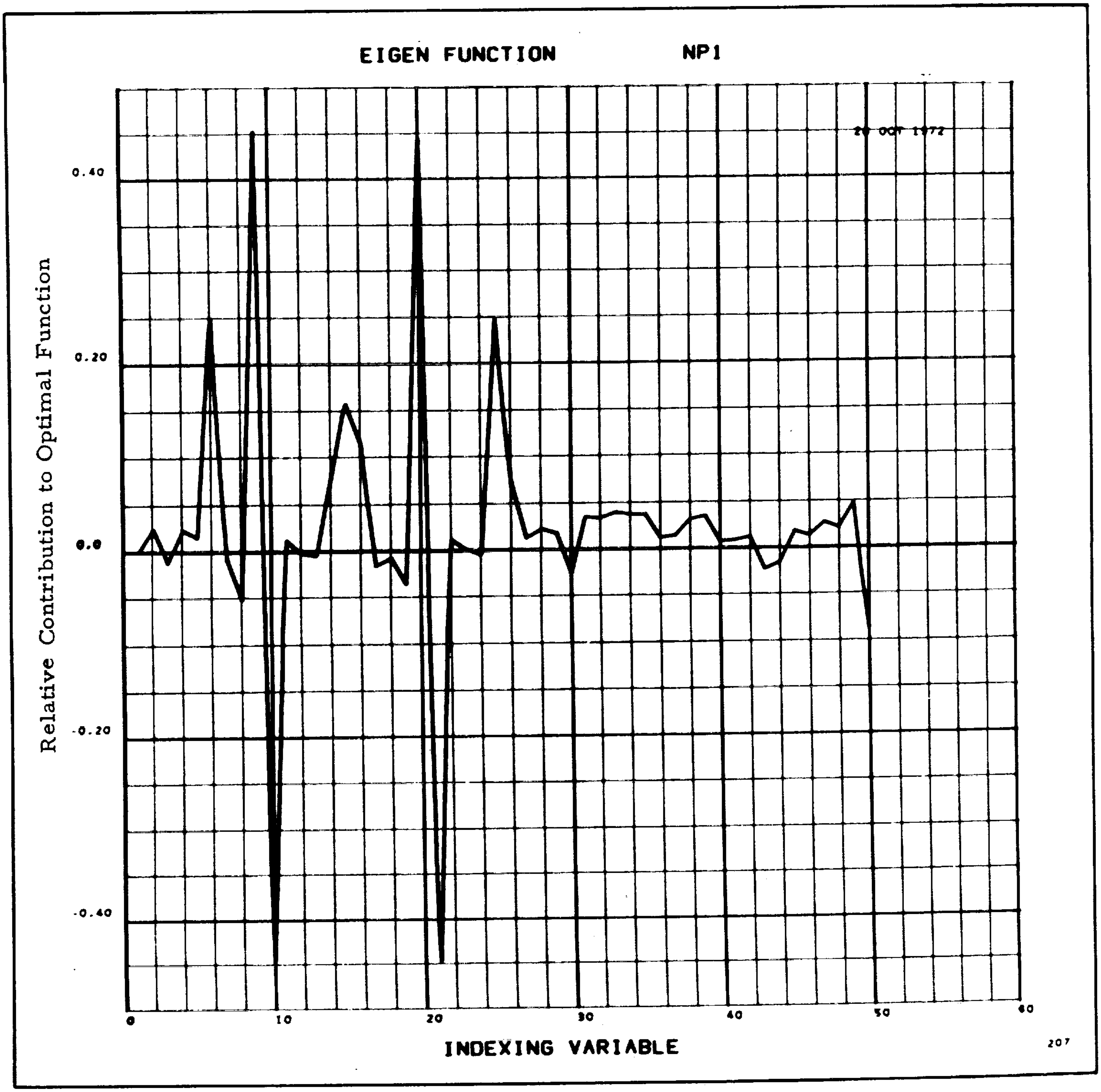


FIGURE 5.5 - SECOND OPTIMUM FUNCTION ORIGIGINAL 50 MEASUREMENT BASE

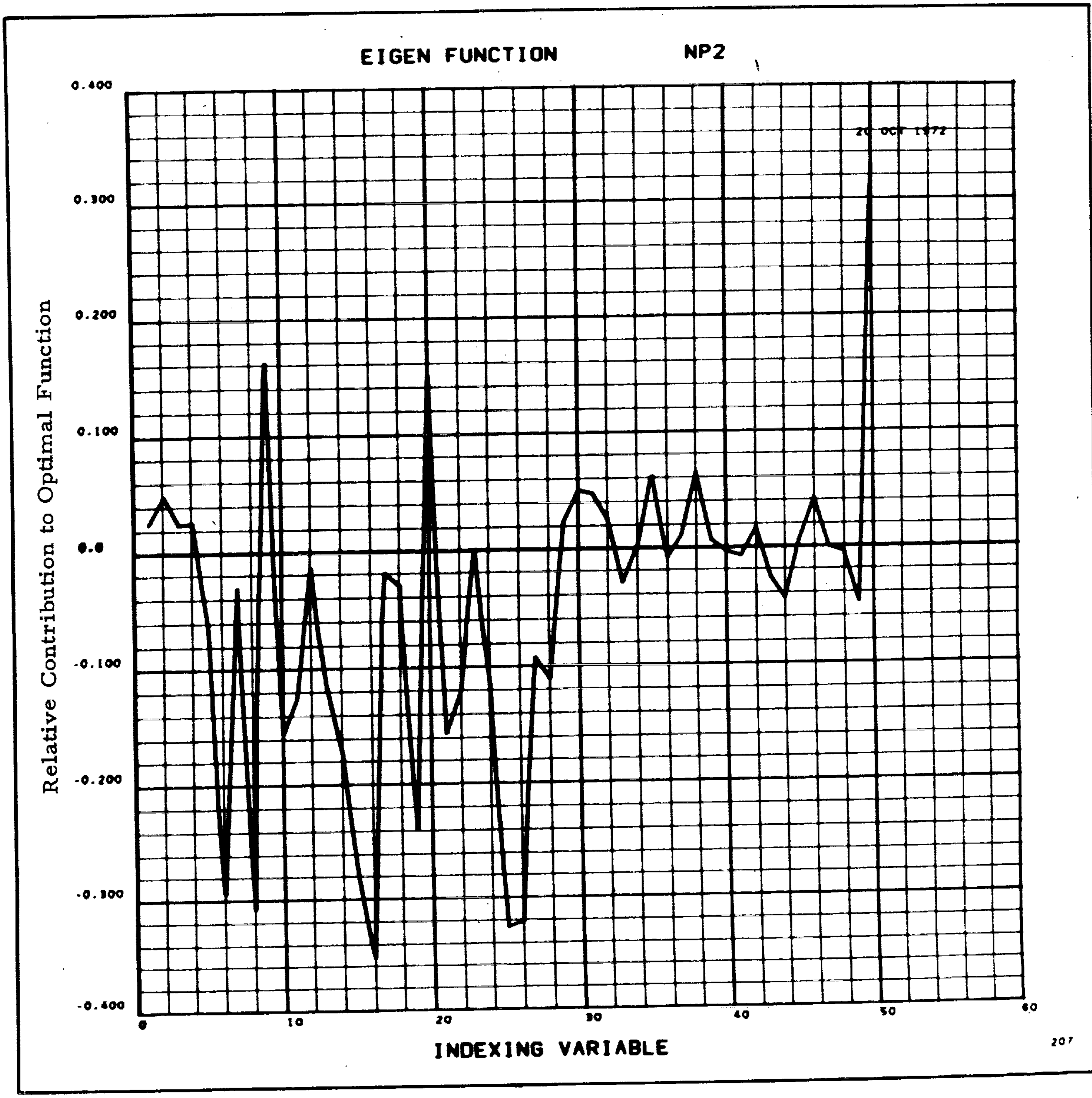


FIGURE 5.6 - COMPARISON OF TYPICAL DATA HISTORY AND TWO-TERM RECONSTRUCTION

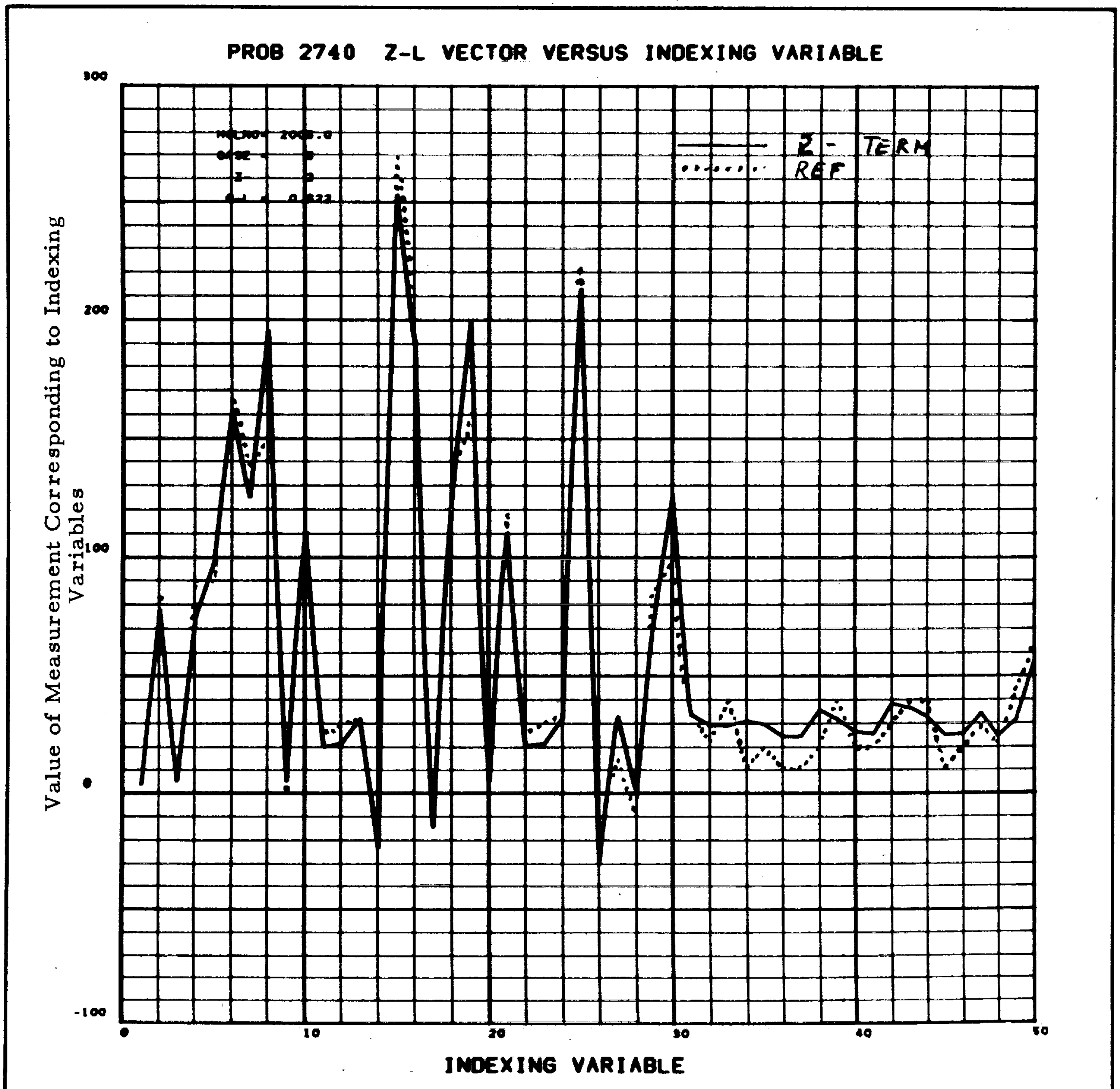


FIGURE 5.7 - COMPARISON OF TYPICAL DATA HISTORY AND FIVE-TERM RECONSTRUCTION

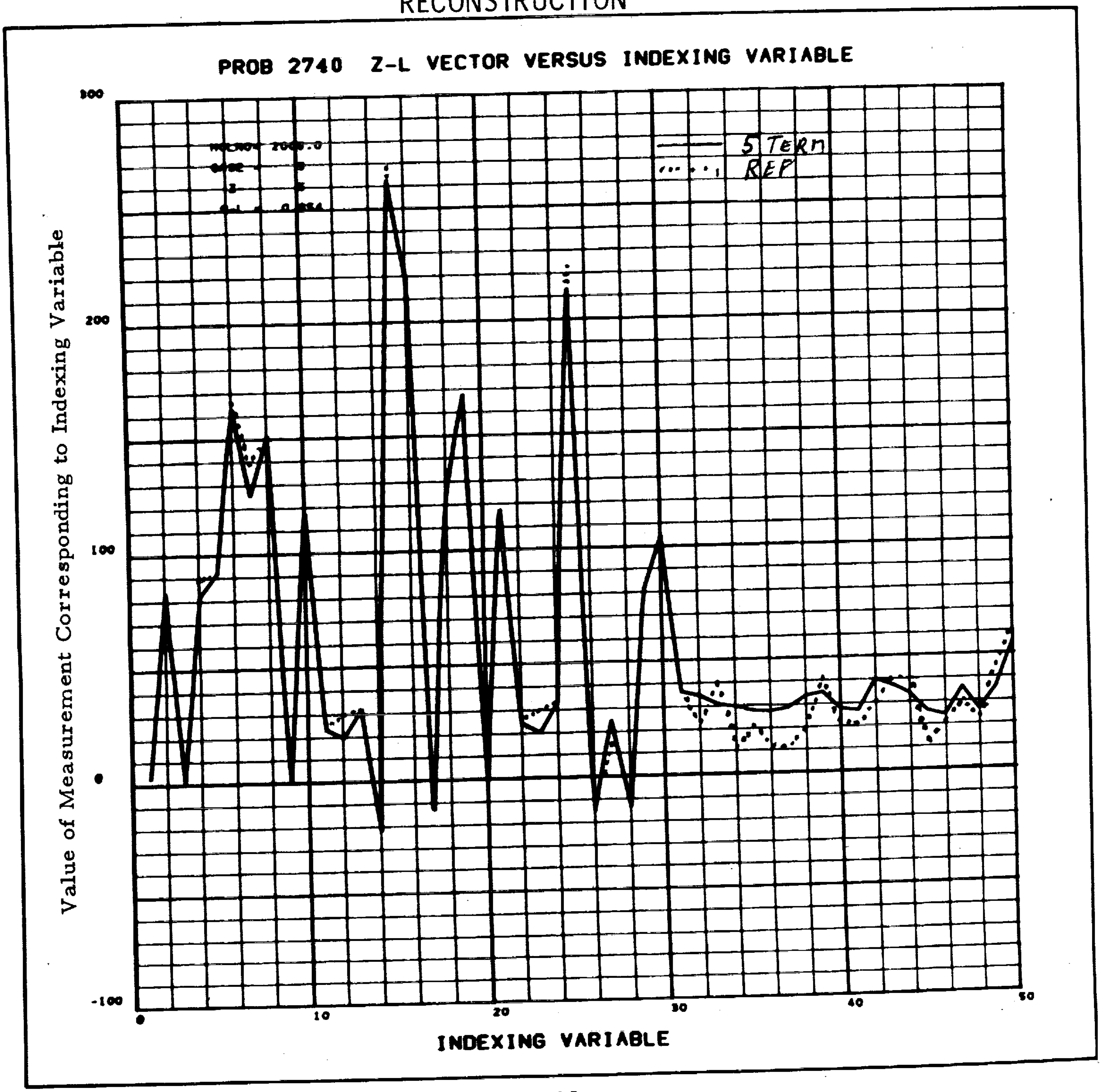


FIGURE 5.8 - COMPARISON OF TYPICAL DATA HISTORY AND TEN-TERM RECONSTRUCTION

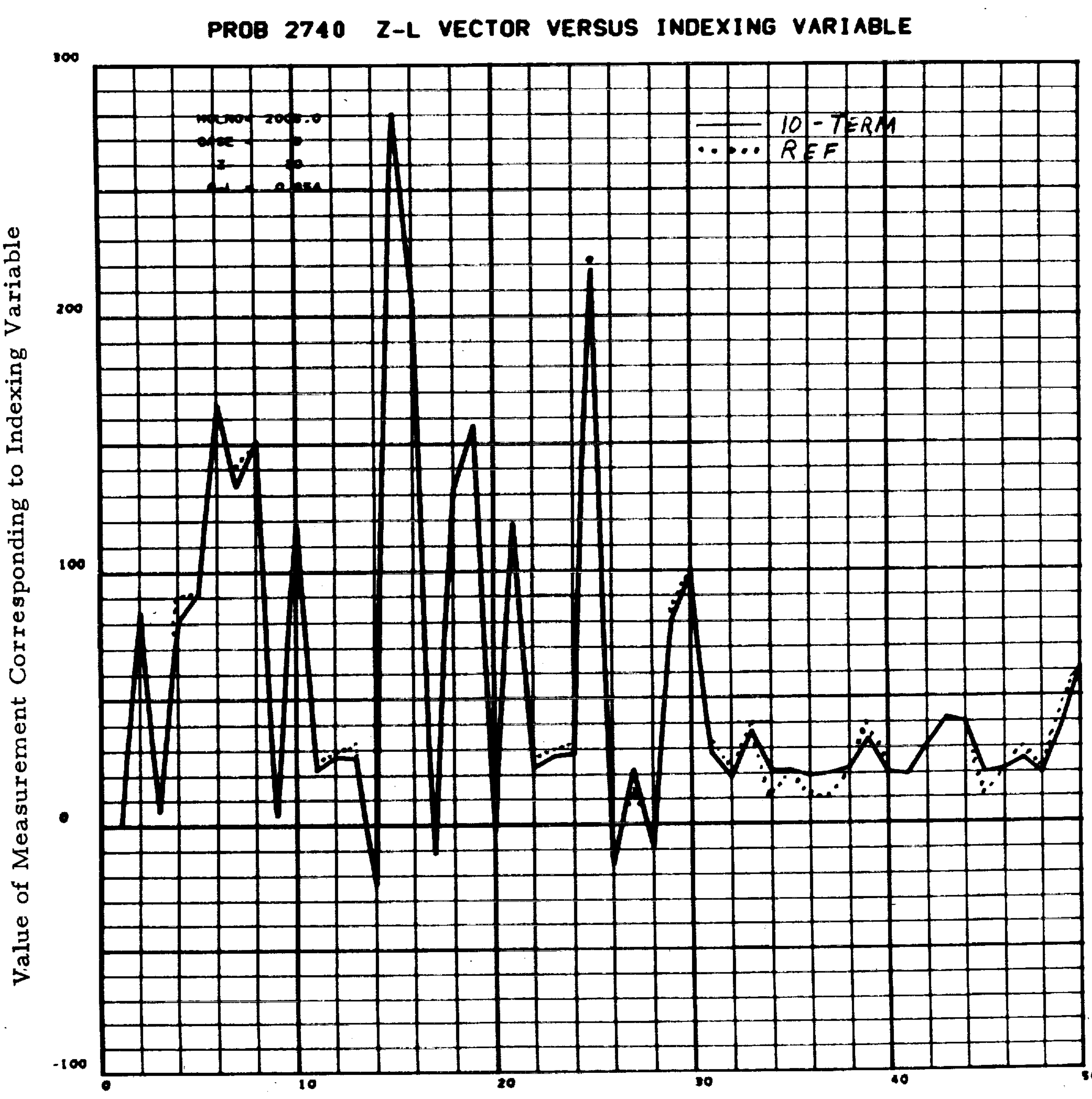


FIGURE 5.9 - SCATTER PLOT OF FIRST AND SECOND COEFFICIENTS OF GENERALIZED FOURIER SERIES REPRESENTATION OF CASES USED TO DERIVE ORIGINAL 50 MEASUREMENT BASE

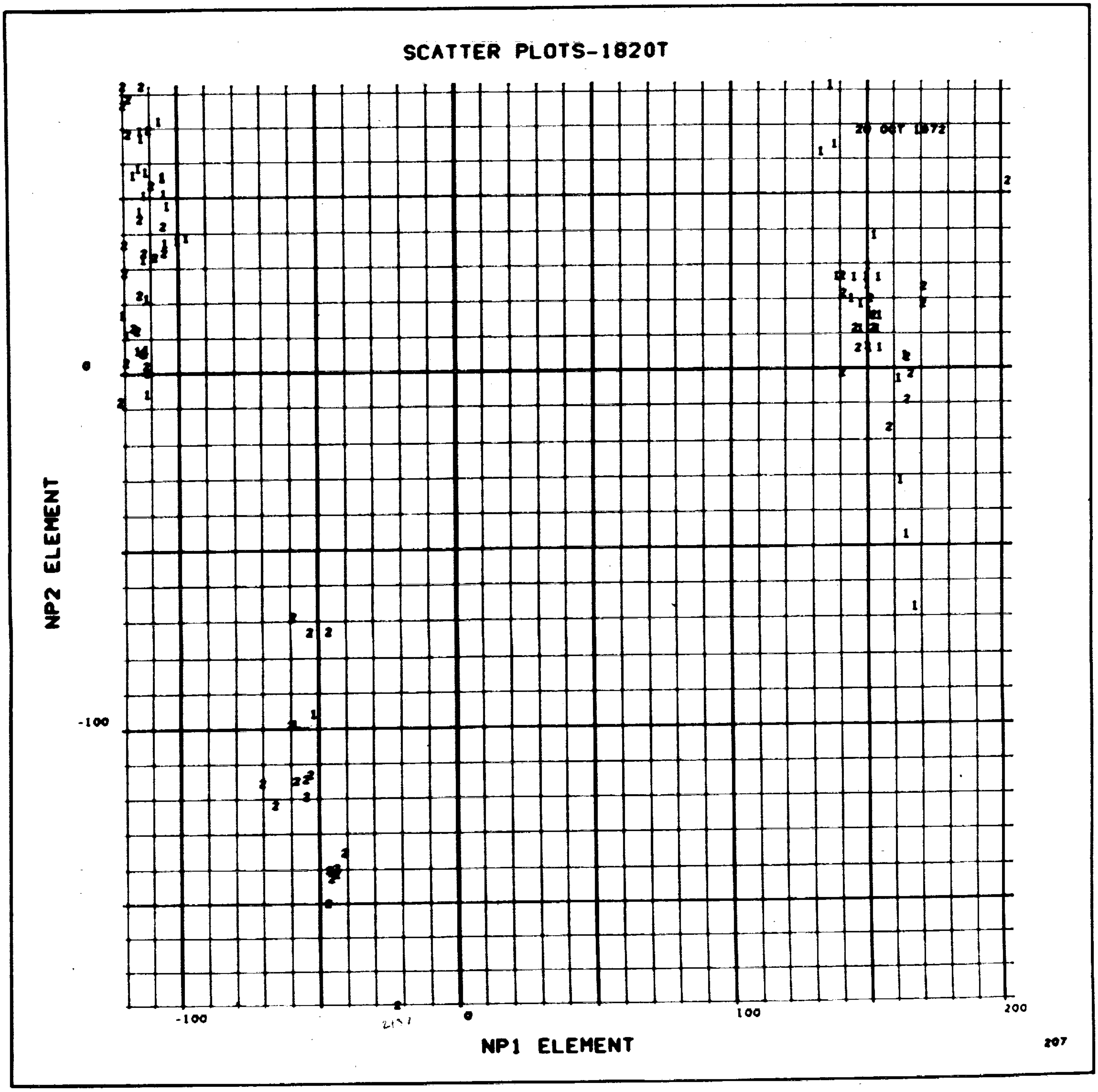


FIGURE 5.10 - SCATTER PLOT OF FOURTH AND FIFTH COEFFICIENTS OF GENERALIZED FOURIER SERIES REPRESENTATION FOR LEARNING DATA USED IN THE ORIGINAL 50 MEASUREMENT BASE

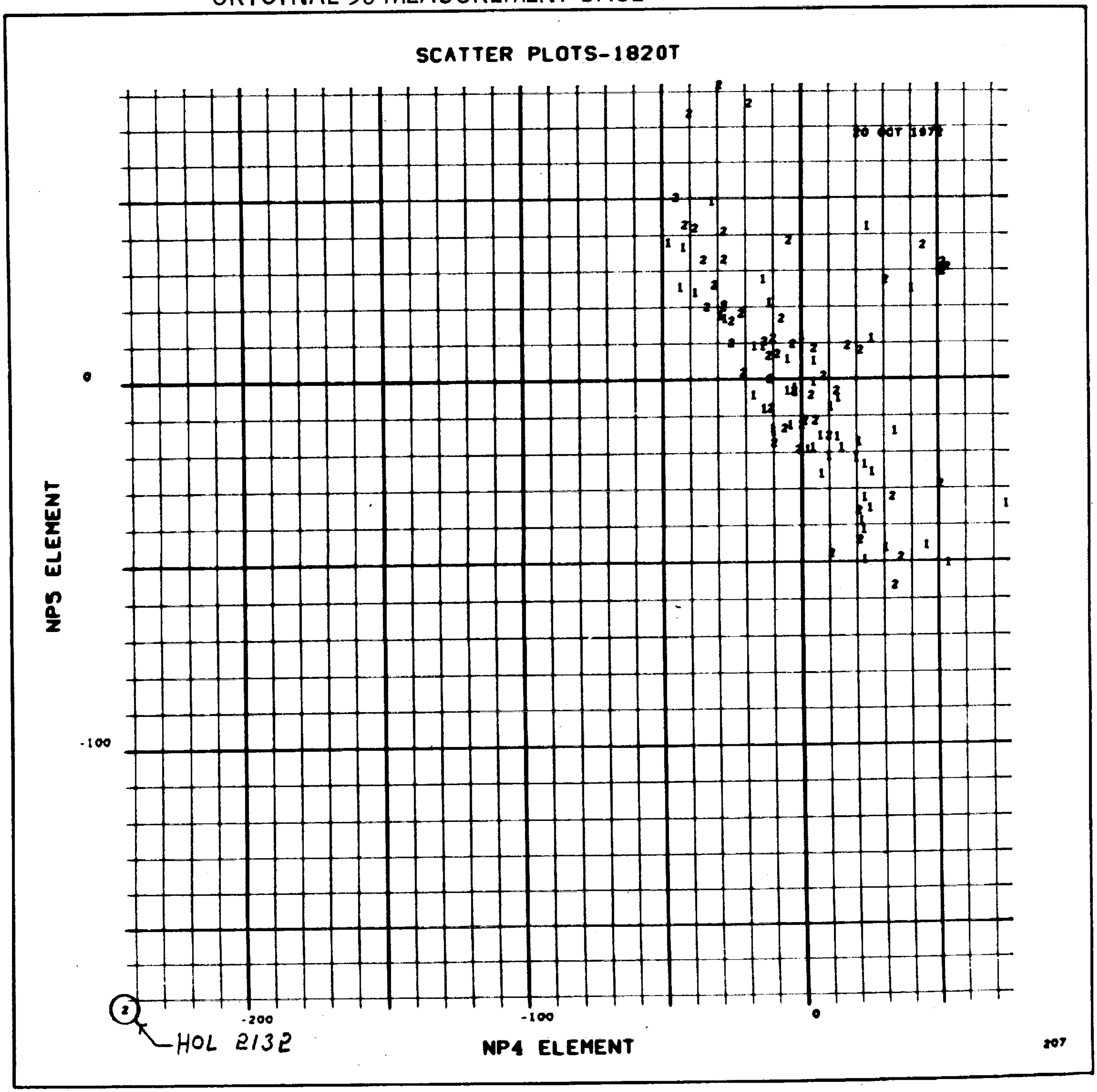


FIGURE 5.11

FOURTH OPTIMUM FUNCTION ORIGINAL 50 MEASUREMENT BASE

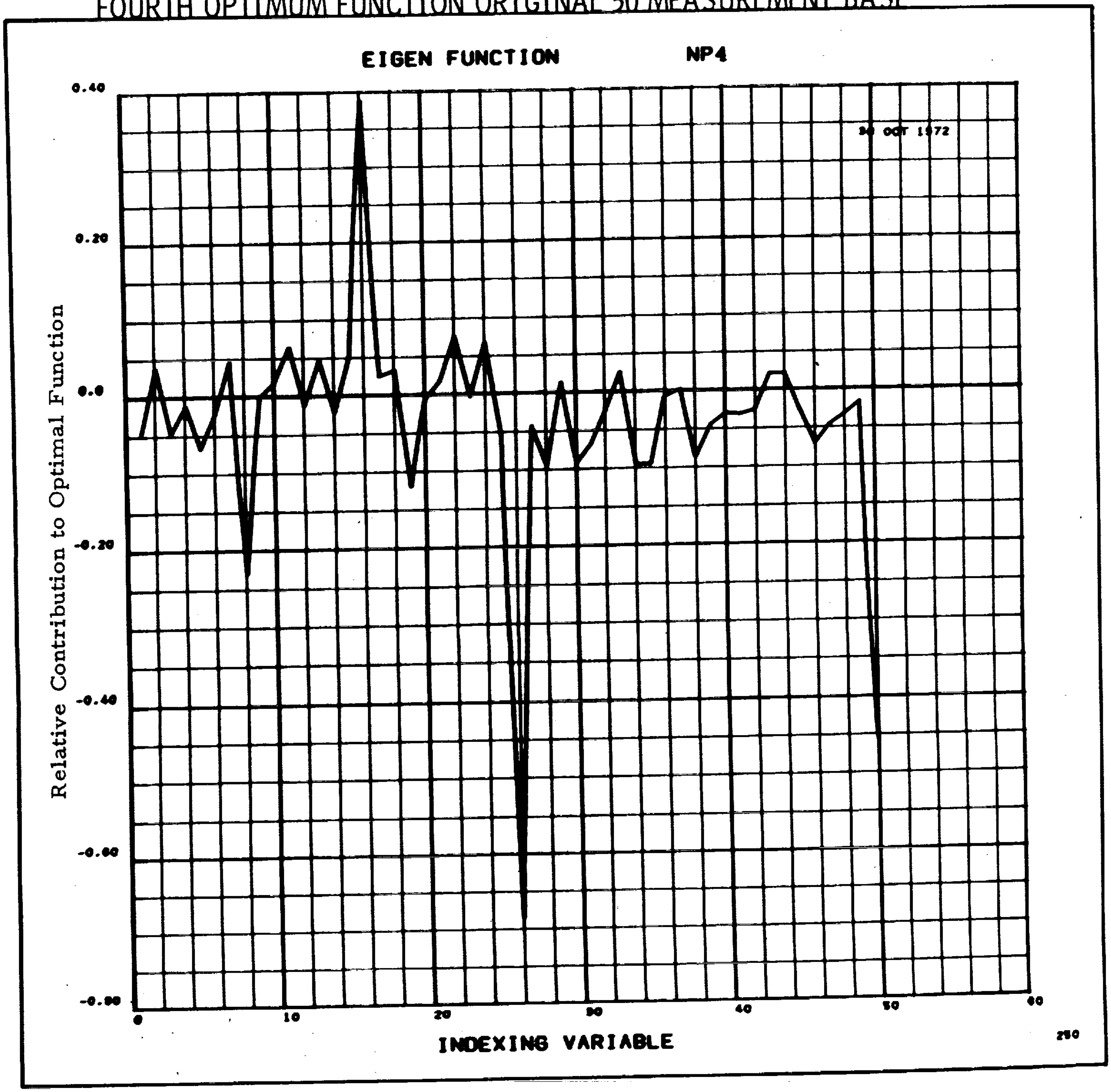


FIGURE 5.12

FIFTH OPTIMUM FUNCTION ORIGINAL 50 MEASUREMENT BASE

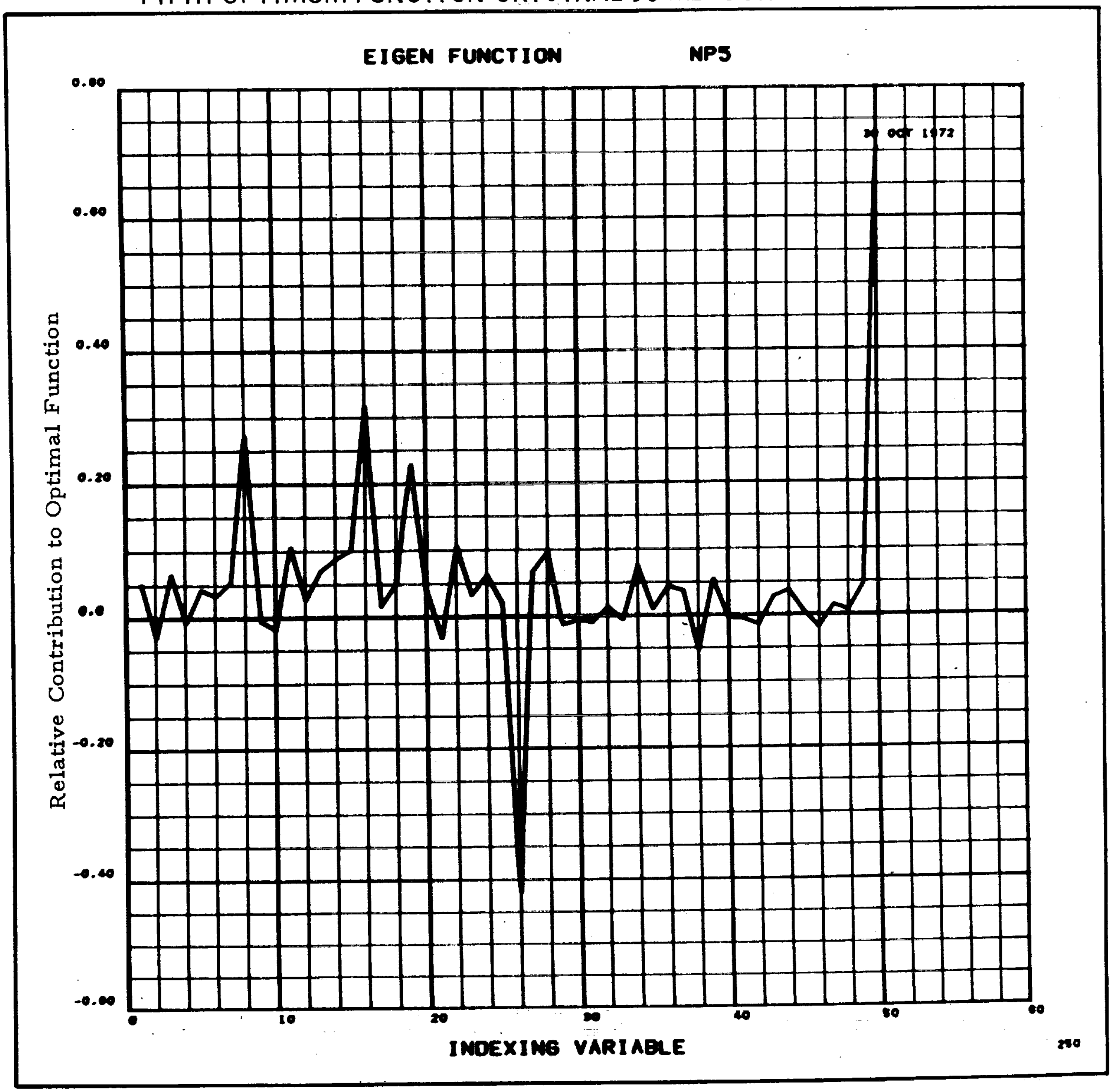


FIGURE 5.13 - VARIATION OF INFORMATION RETAINED WITH DIMENSIONALITY FOR REFERENCE 50 DIMENSIONAL BASE

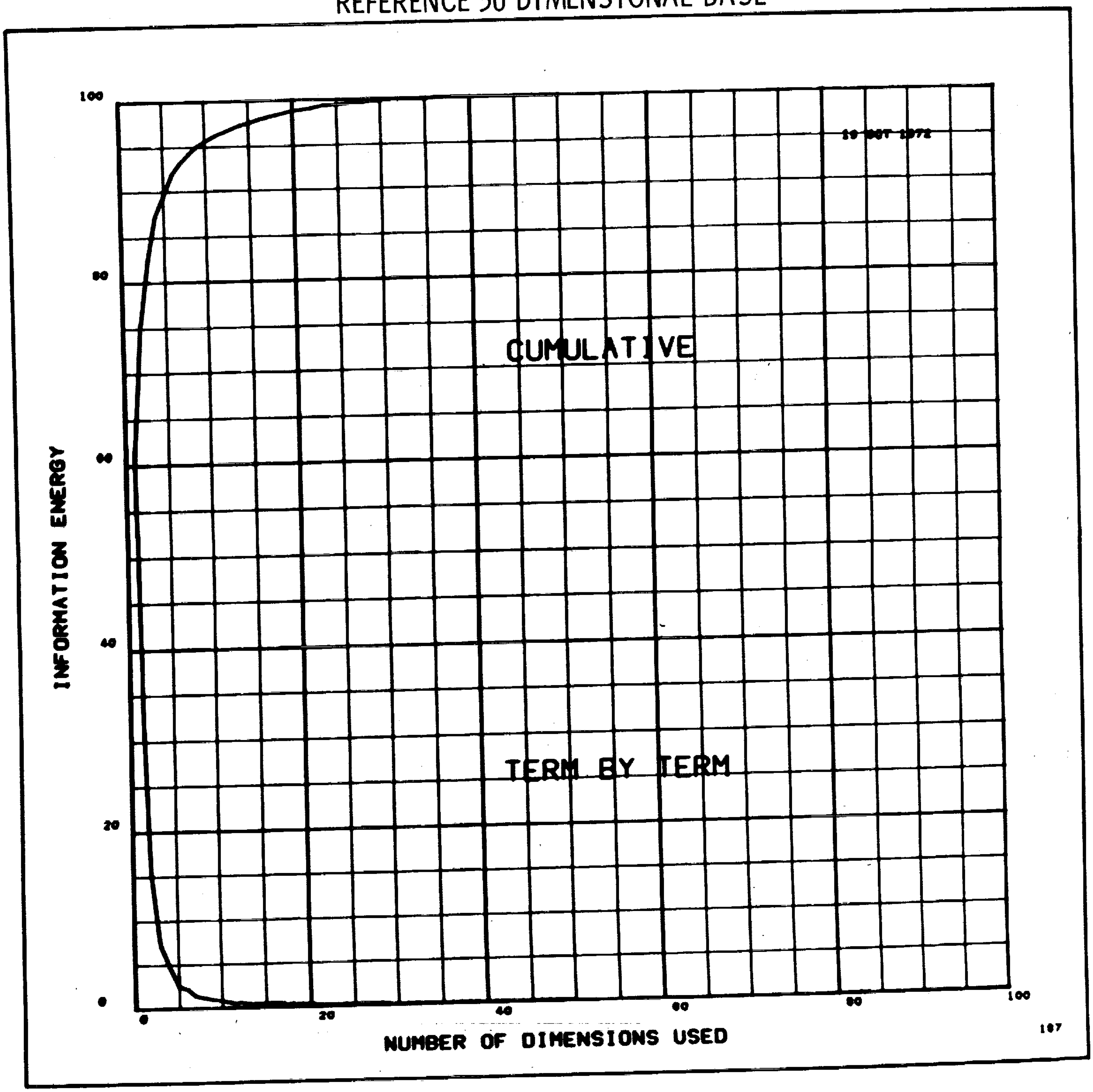


FIGURE 5.14 - FIRST OPTIMUM FUNCTION FOR REFERENCE 50 DIEMNSIONAL BASE

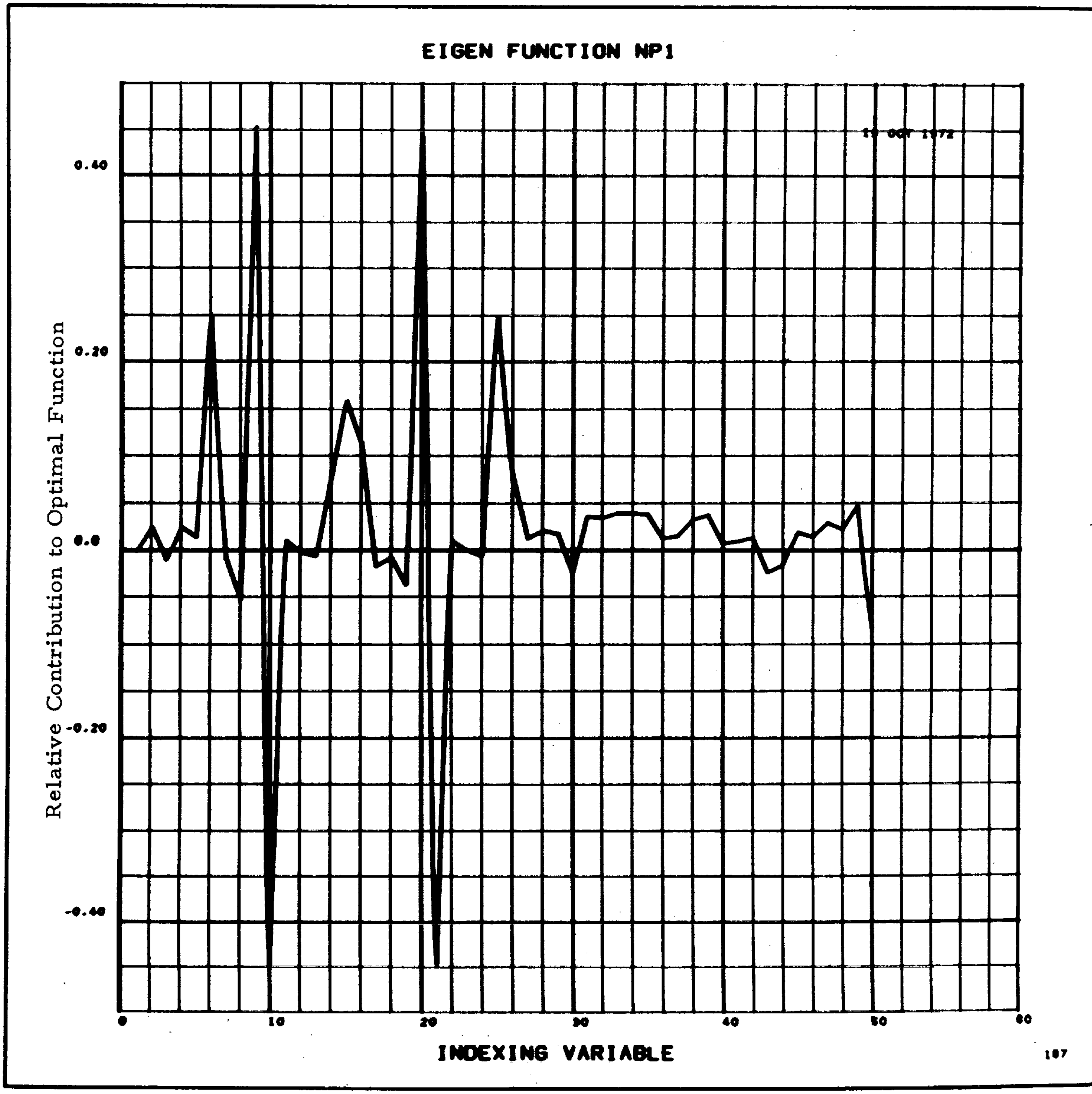


FIGURE 5.15 - SECOND OPTIMUM FUNCTION FOR REFERENCE 50 DIMENSIONAL BASE

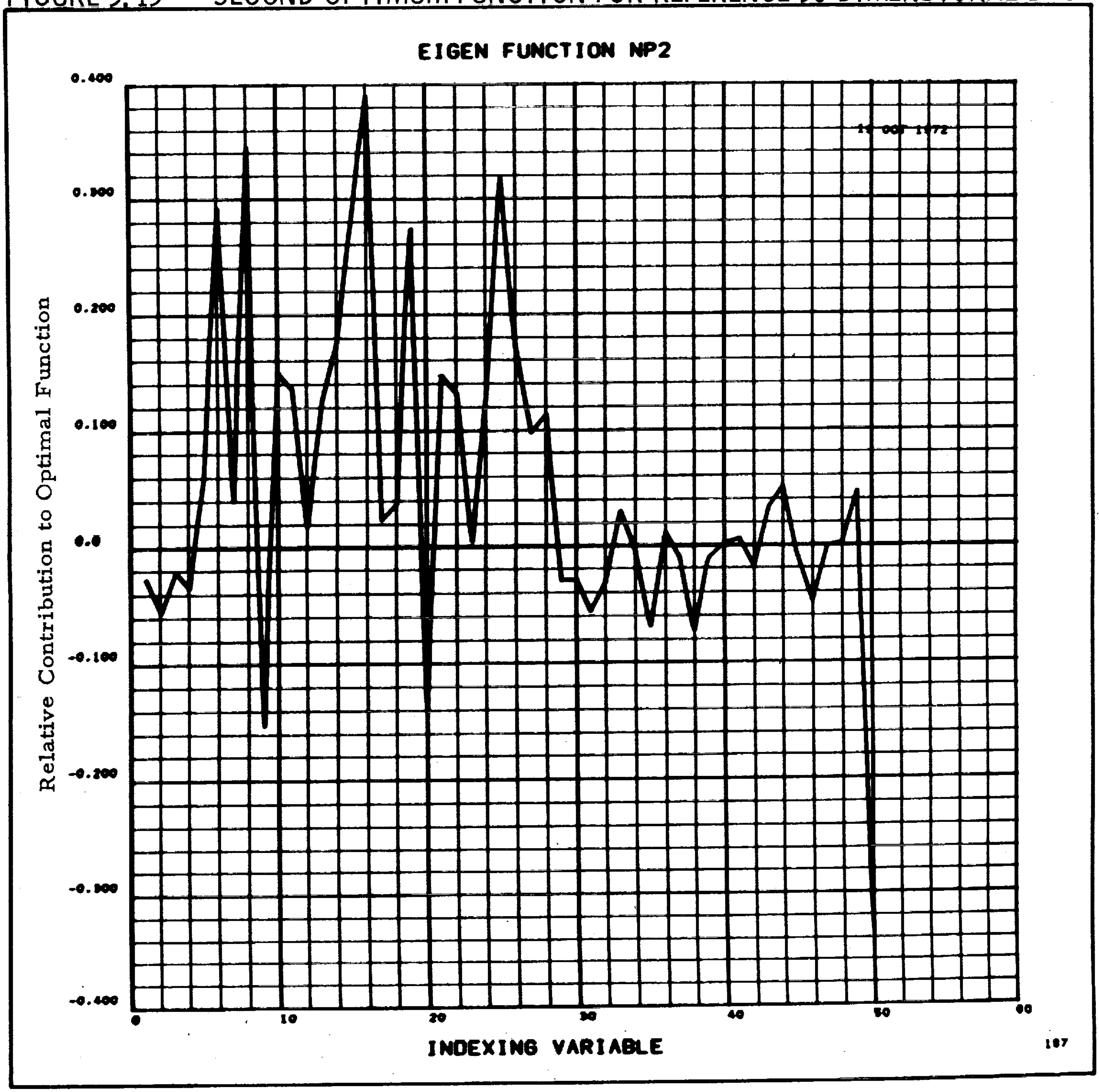


FIGURE 5.16

SCATTER PLOT OF FIRST AND SECOND COEFFICIENTS OF GENERALIZED FOURIER SERIES FOR CASES USED TO DEVELOP REFERENCE 50 MEASUREMENT BASE

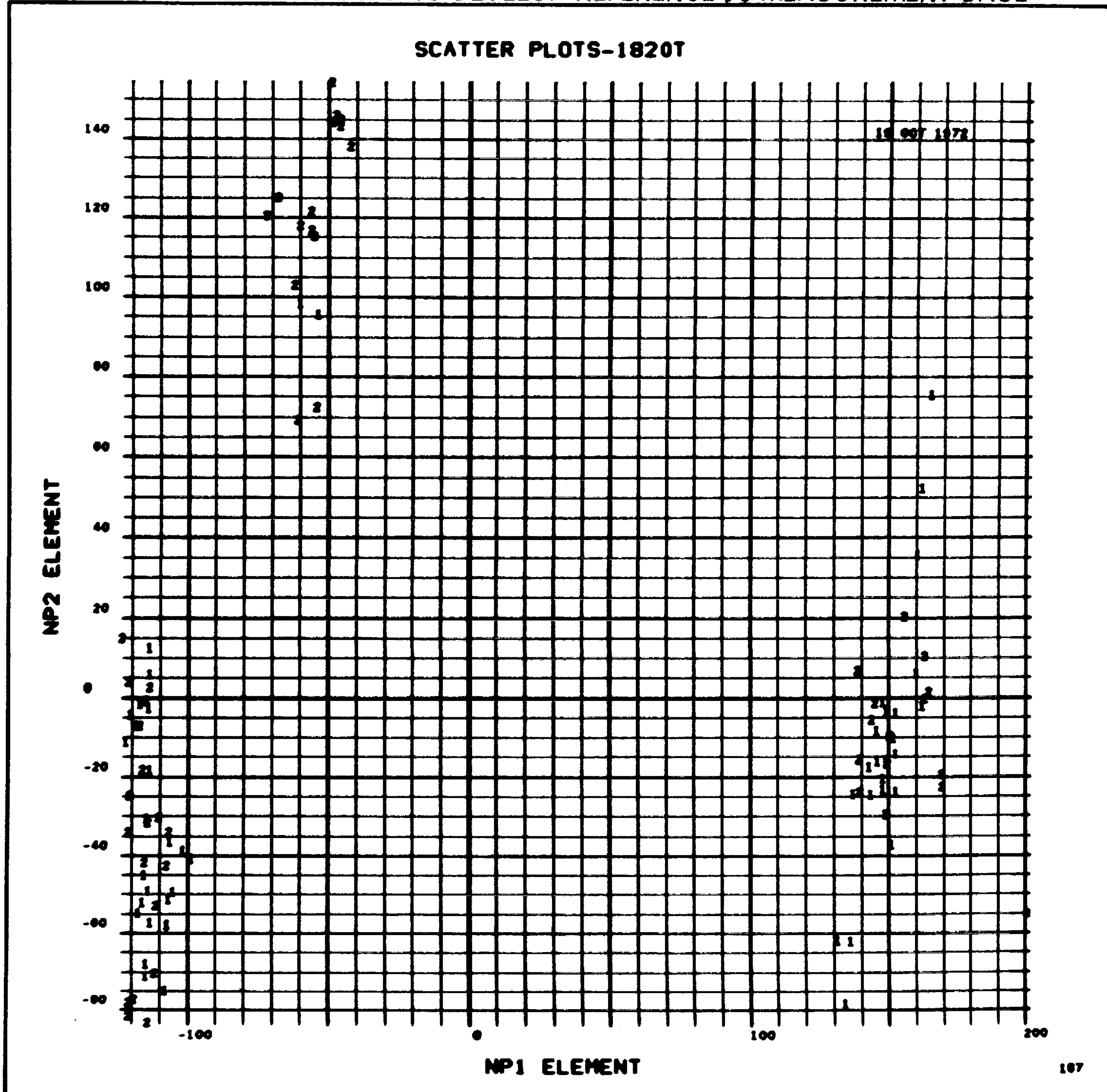


FIGURE 5.17 - SEVENTH OPTIMUM FUNCTION FOR ORIGINAL 50 MEASUREMENT BASE

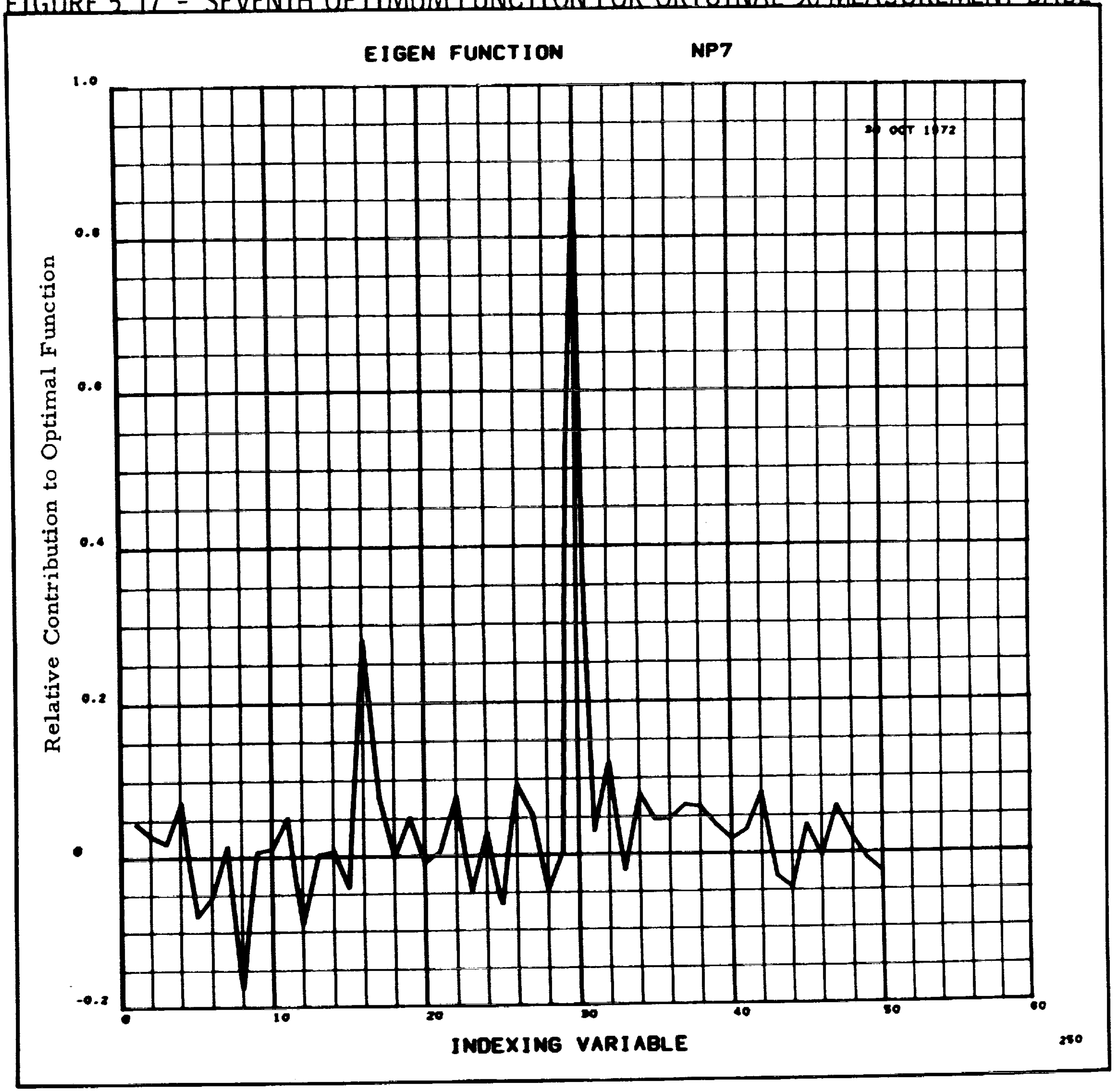


FIGURE 5.18 - SIXTH OPTIMUM FUNCTION FOR REFERENCE 50 MEASUREMENT BASE

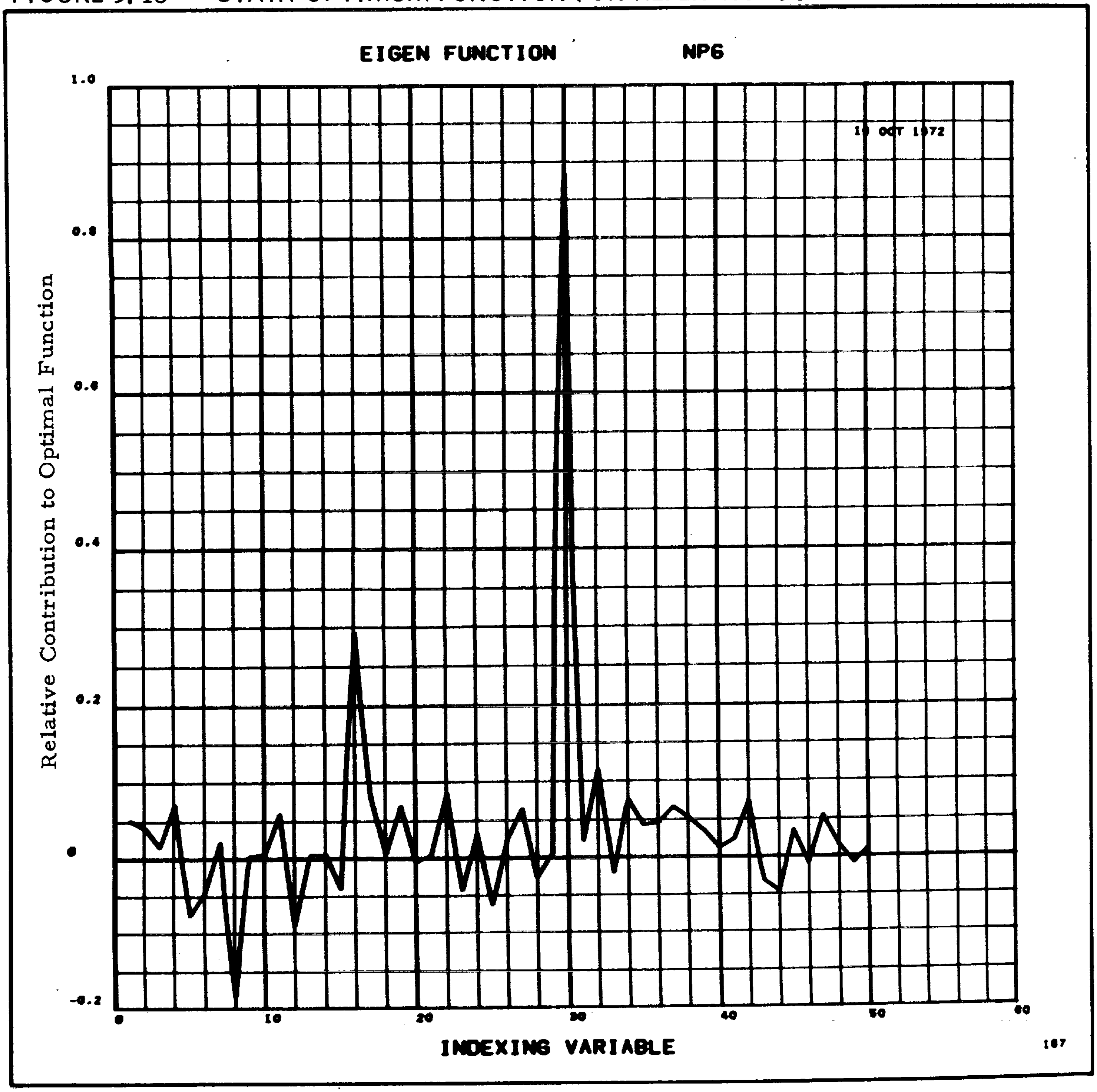


FIGURE 5.19 - THIRD OPTIMUM FUNCTION FOR REFERENCE 50 MEASUREMENT BASE

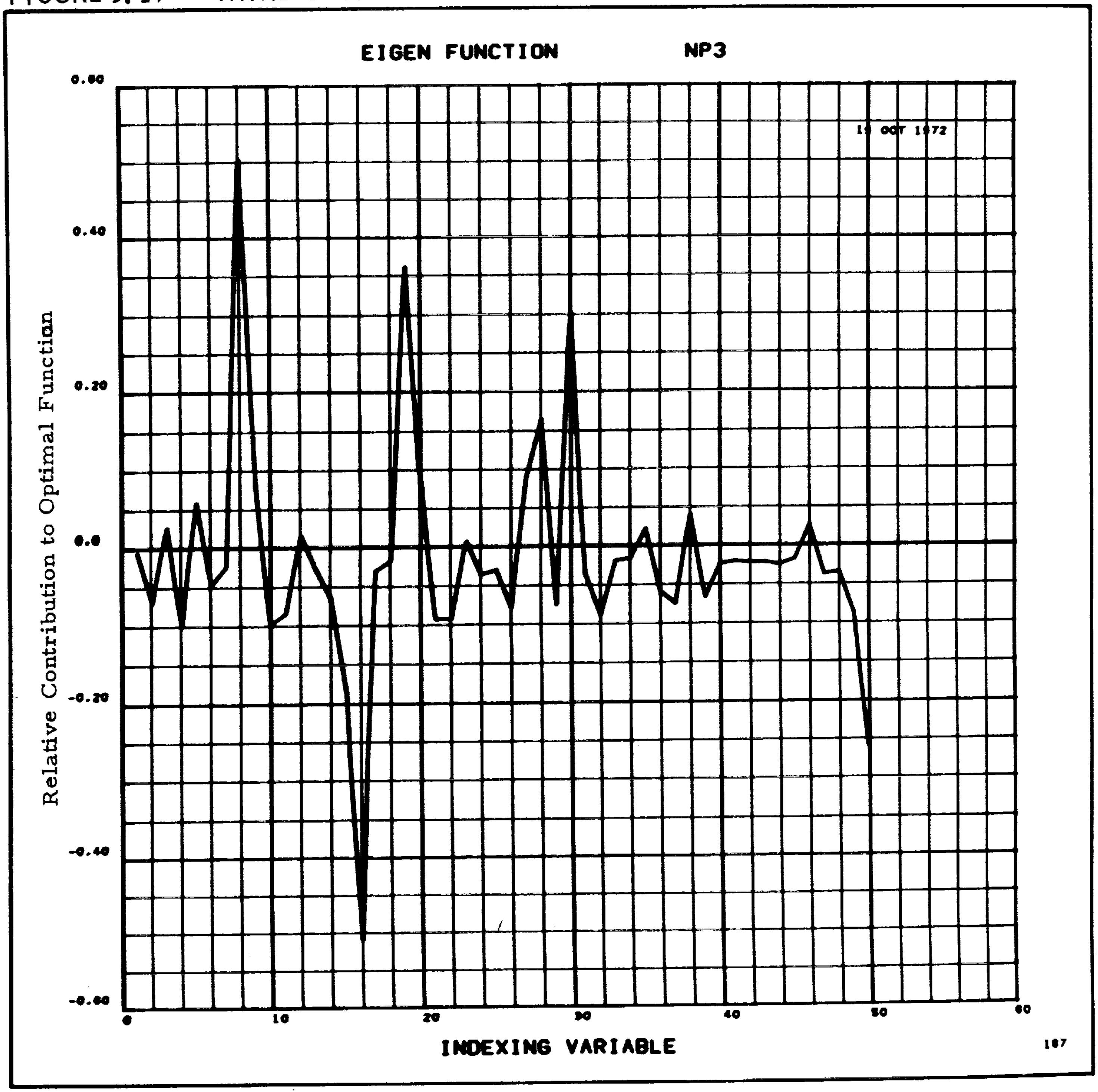


FIGURE 5.20 - FOURTH OPTIMUM FUNCTION FOR REFERENCE 50 MEASUREMENT BASE

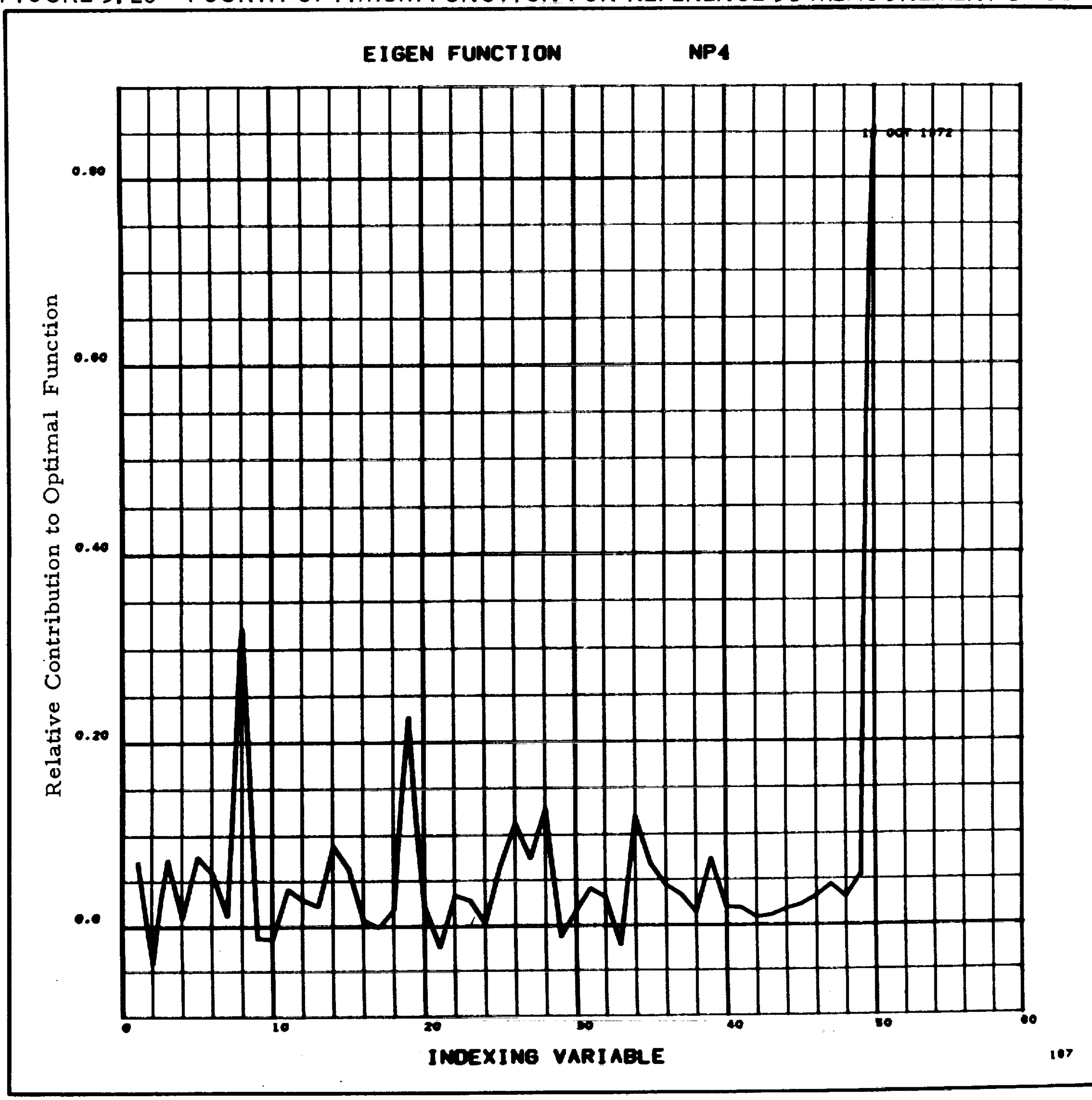


FIGURE 5.21 - EIGHTH OPTIMUM FUNCTION FOR REFERENCE 50 MEASUREMENT BASE

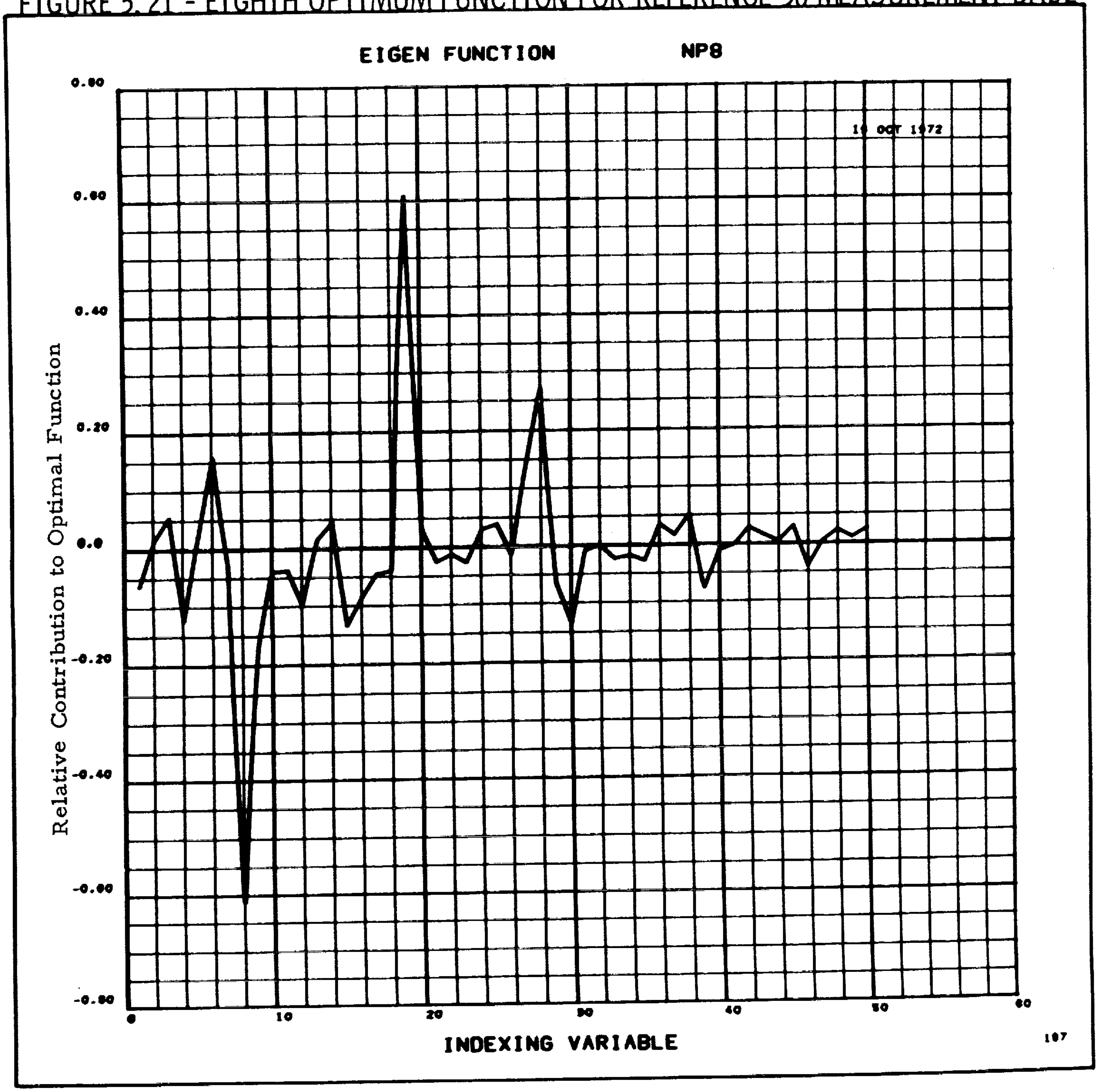


FIGURE 5.22 - NINETEENTH OPTIMUM FUNCTION FOR REFERENCE 50 MEASUREMENT BASE

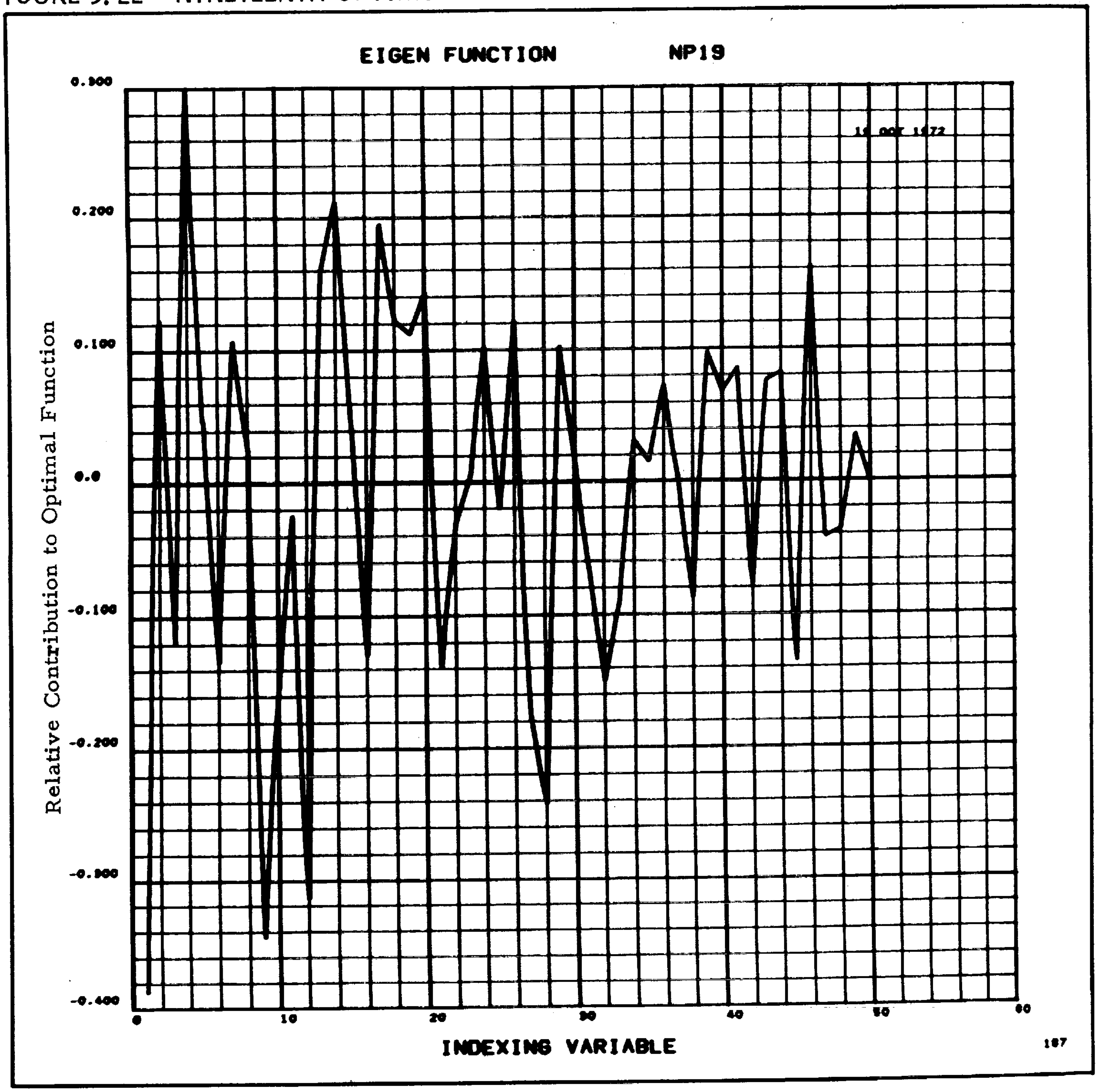


FIGURE 5.23 - PERFORMANCE MAP COMPARISON 29-CASE EXPLORATORY ALGORITHMS

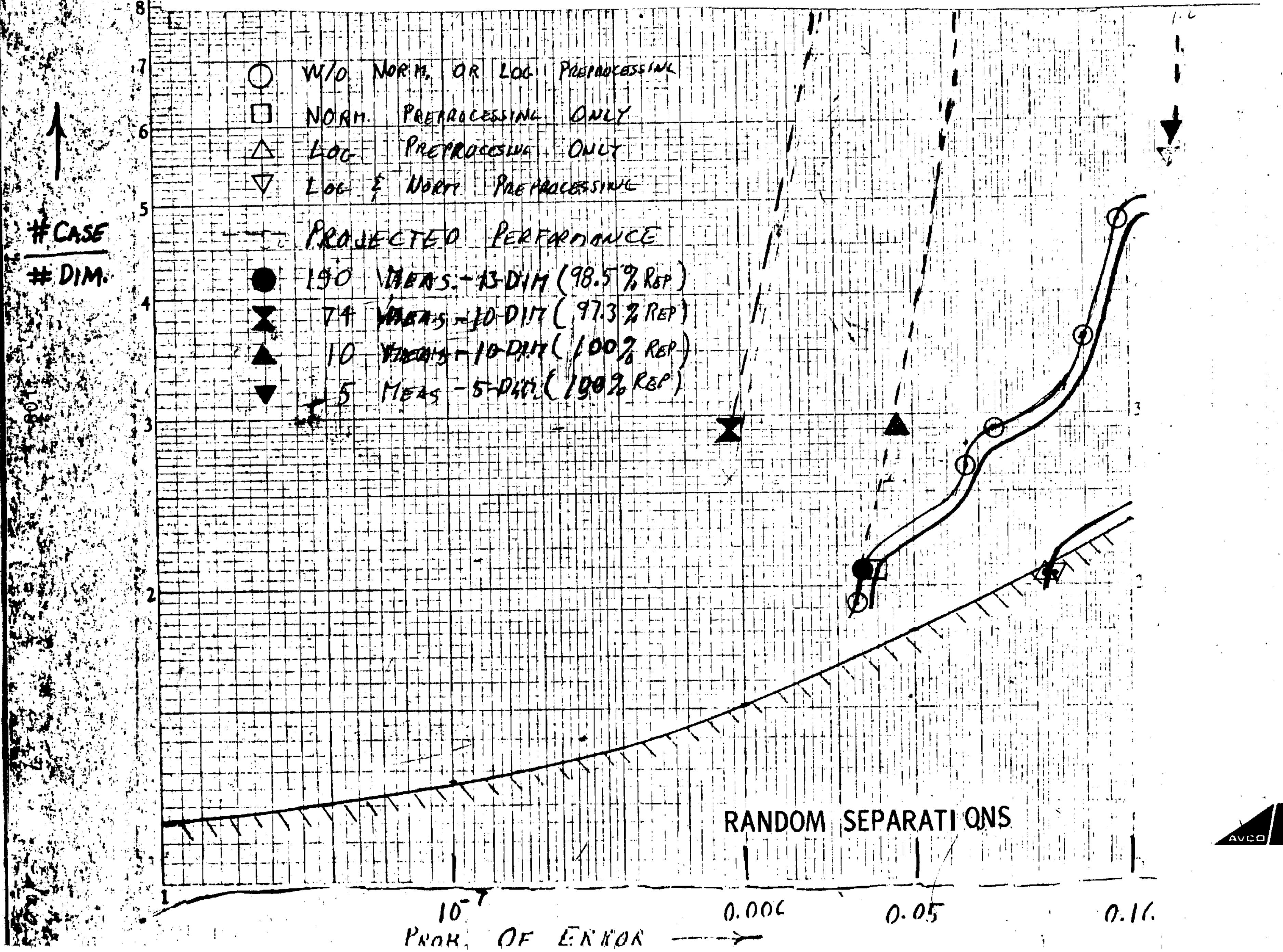
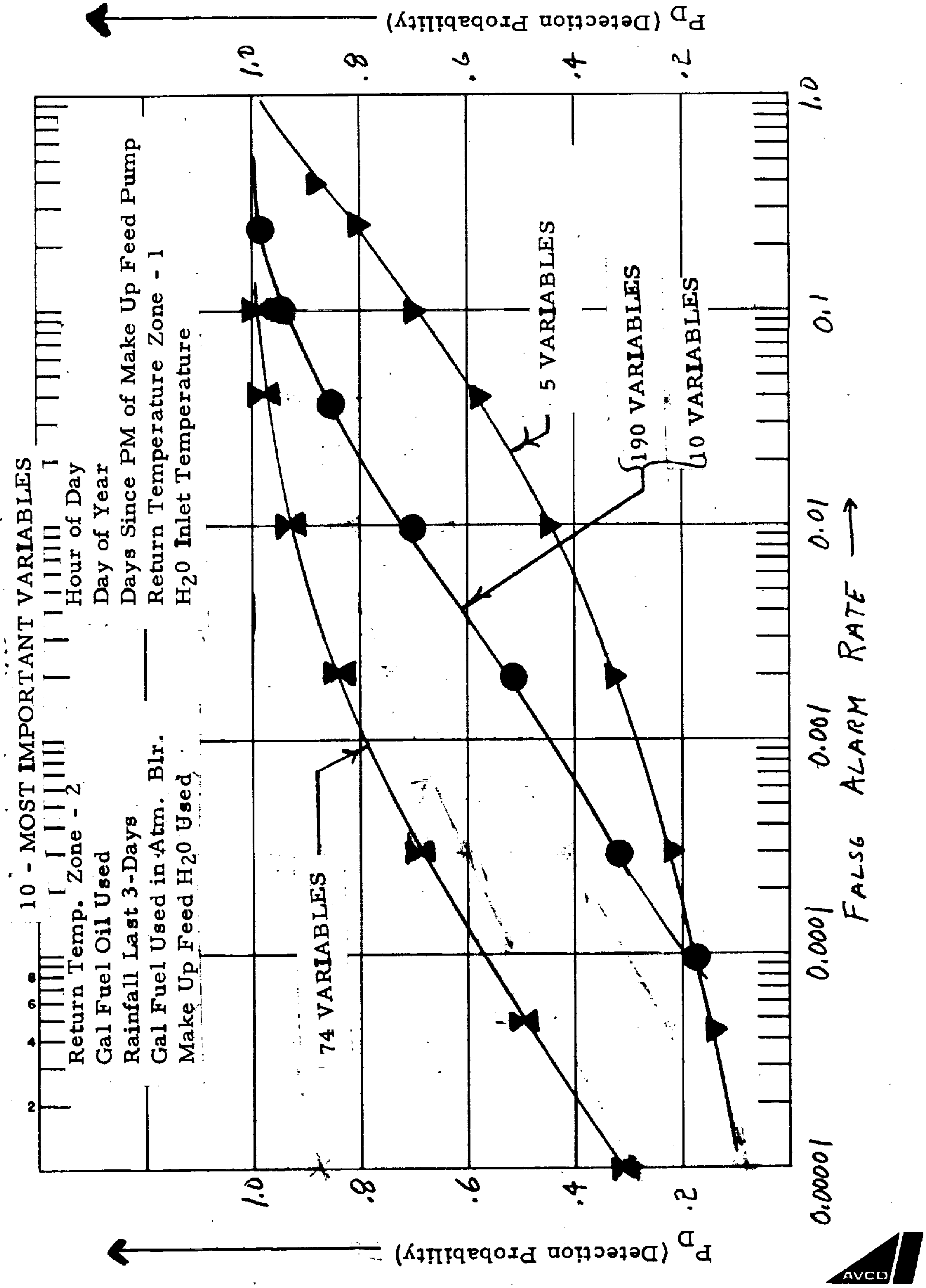


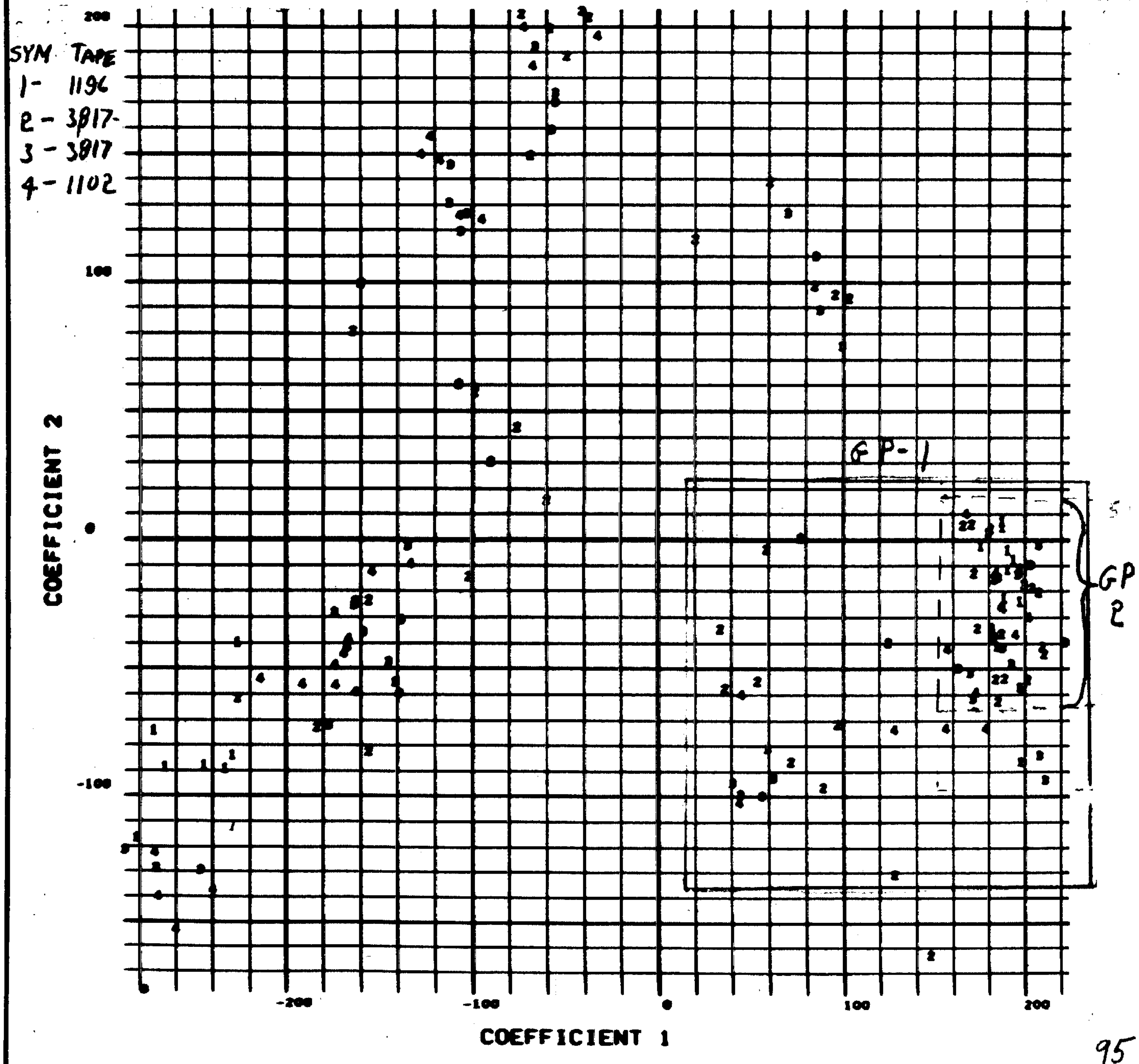
FIGURE 5.24
PROJECTED CLASSIFICATION PERFORMANCE TRADE-OFF CURVES
AS OF FUNCTION OF NUMBER OF VARIABLES USED



NOTE: USE TYPE B PENCIL FOR VUGRAPHS AND REPORT DATA.

FIGURE 5.25

SCATTER PLOT OF LEARNING DATA USED TO DERIVE 81 MEASUREMENT BASE



95

FIGURE 5.26

VARIATION OF INFORMATION RETAINED WITH DIMENSIONALITY FOR THE 81 MEASUREMENT BASE CONSTRUCTED FROM THE CASES BELONGING TO SCATTER PLOT GROUP 2

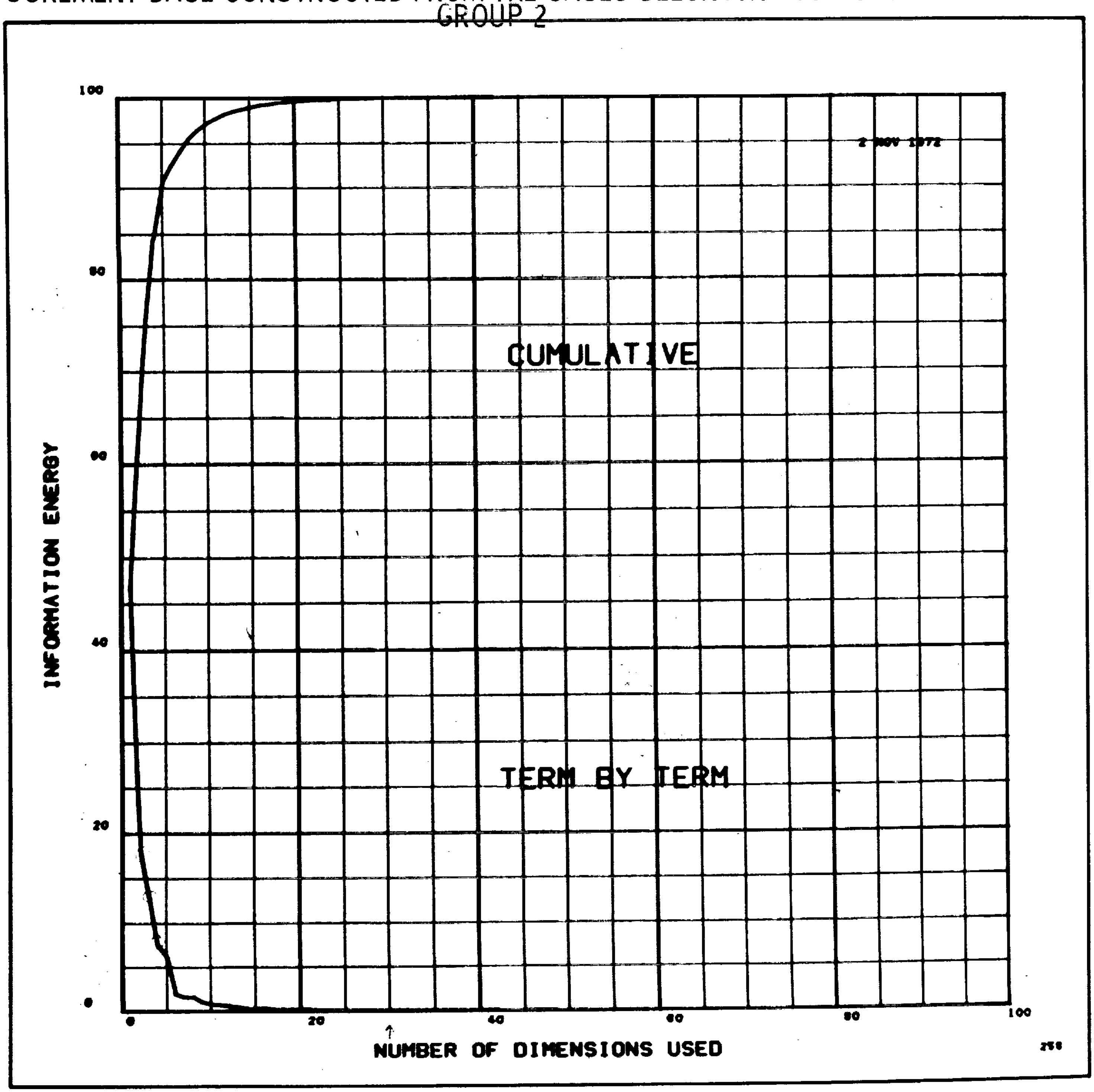


FIGURE 5.27 - VARIATION OF INFORMATION RETAINED WITH DIMENSIONALITY FOR REFERENCE 81 MEASUREMENT BASE

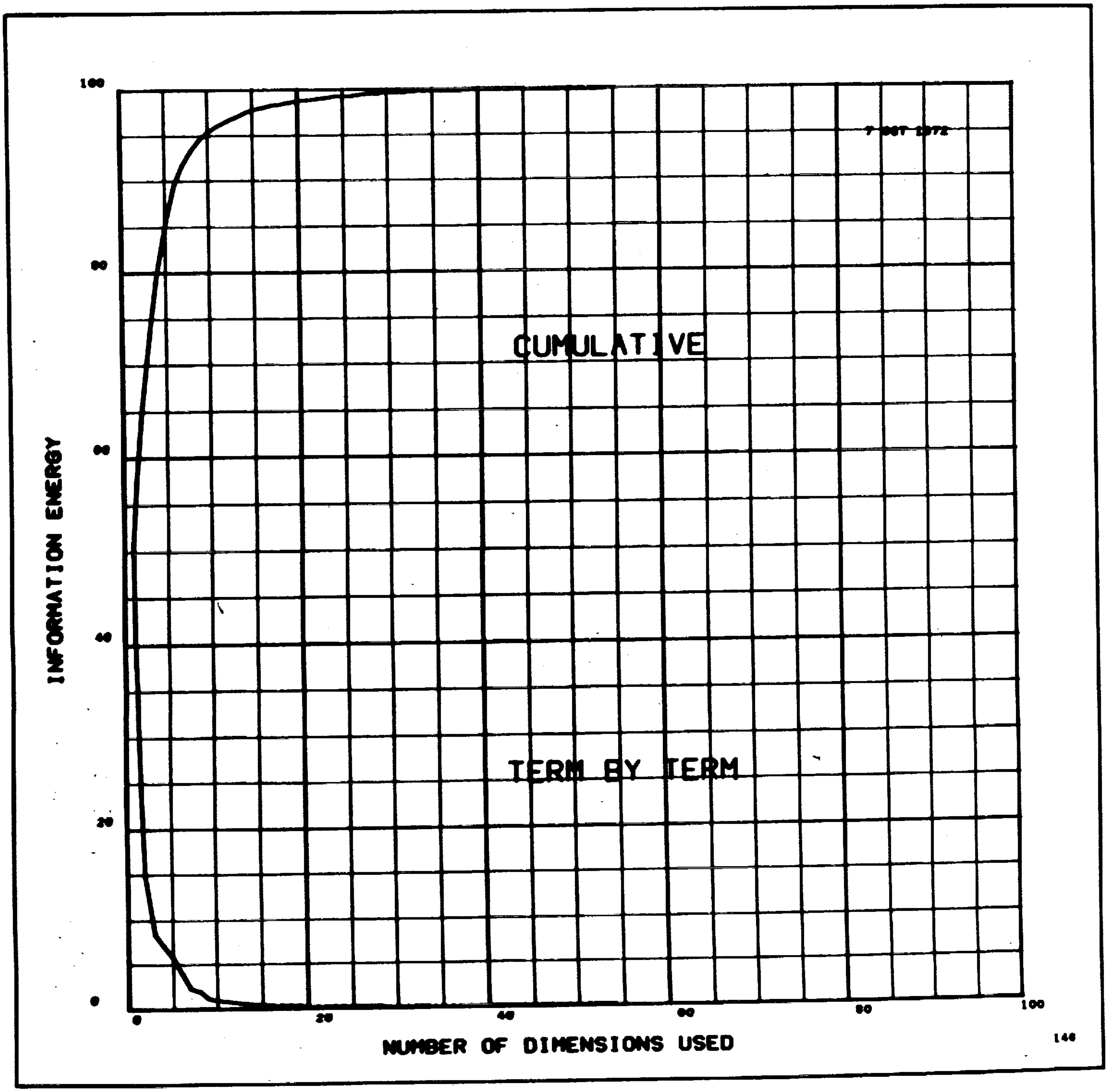


FIGURE 5.28 - PERFORMANCE MAP COMPARISON OF 81 MEASUREMENT DETECTION ALGORITHMS

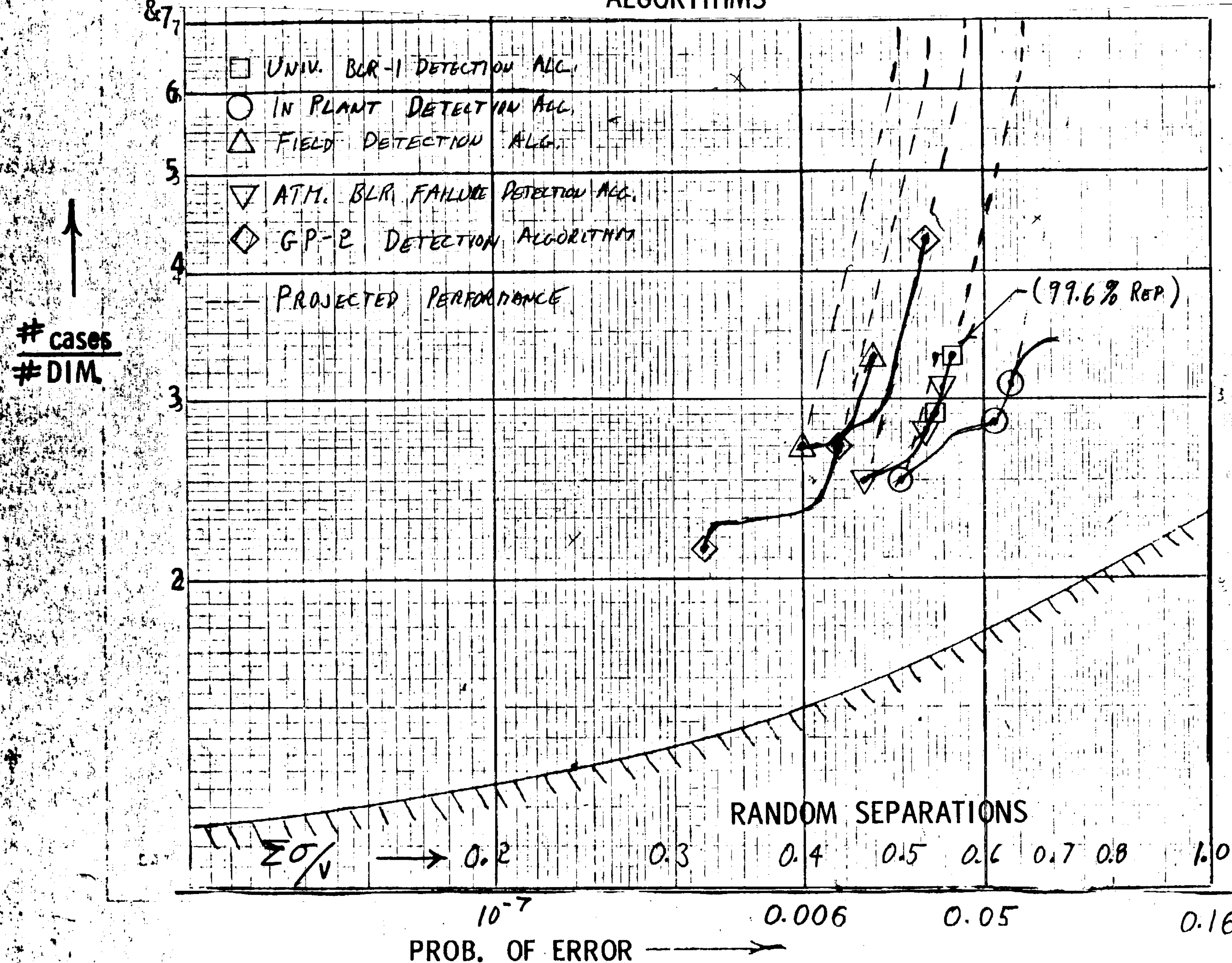
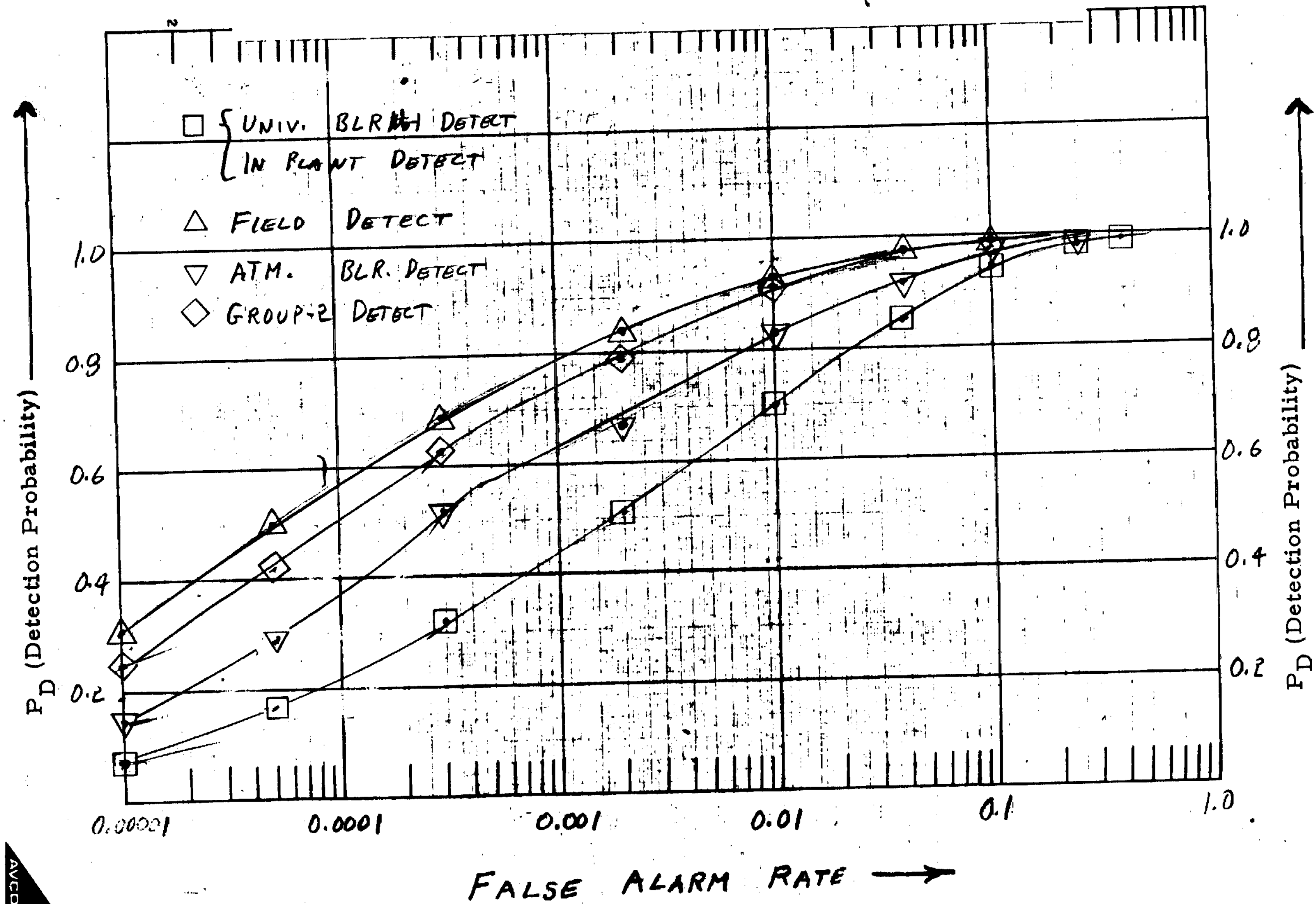


FIGURE 5.29 - CLASSIFICATION PERFORMANCE TRADE-OFF CURVES FOR COMPARING PROJECTED PERFORMANCE OF EXPLORATORY DETECTION ALGORITHMS



NOTE: USE TYPE B PENCIL FOR VUGRAPHS AND REPORT DATA.

FIGURE 5.30 - RELATIVE IMPORTANCE OF MEASUREMENTS FOR DETECTING INCIPIENT FAILURES USING 30 DIMENSIONS

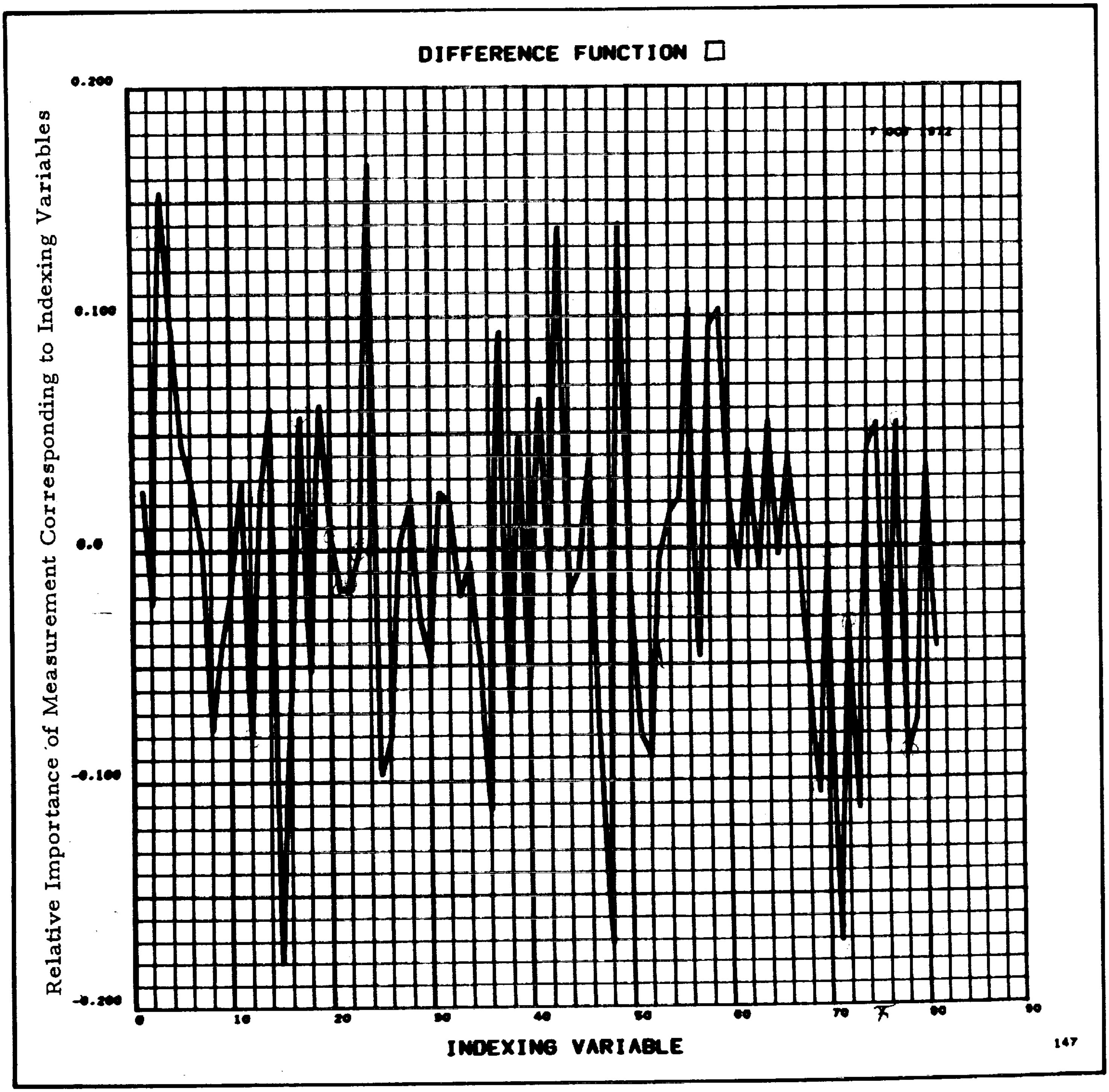


FIGURE 5.31 - RELATIVE IMPORTANCE OF MEASUREMENTS FOR DETECTING INCIPIENT IN-PLANT FAILURES USING 40 DIMENSIONS

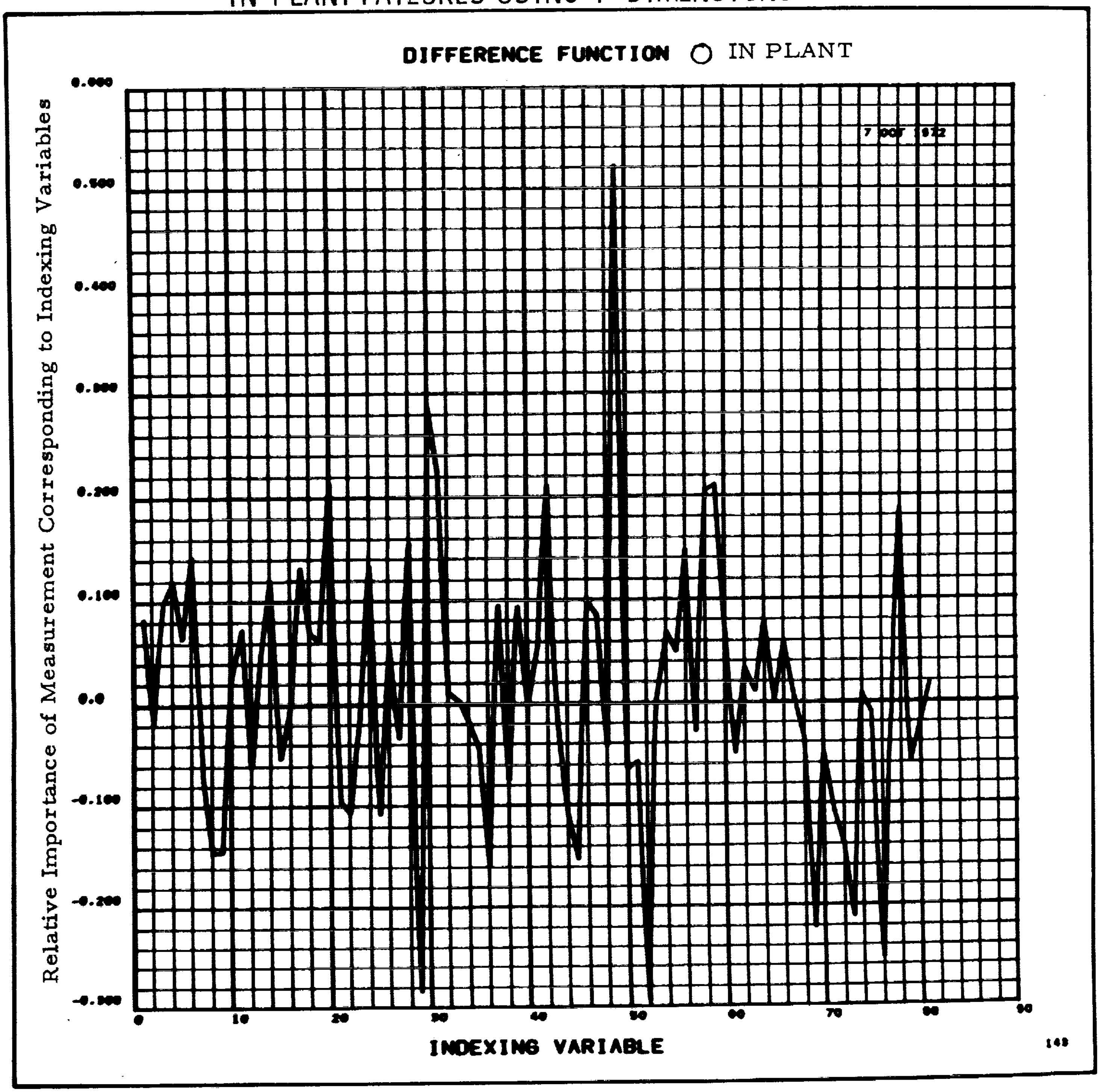


FIGURE 5.32 - RELATIVE IMPORTANCE OF MEASUREMENTS FOR DETECTING FIELD PROBLEMS USING 24 MEASUREMENTS

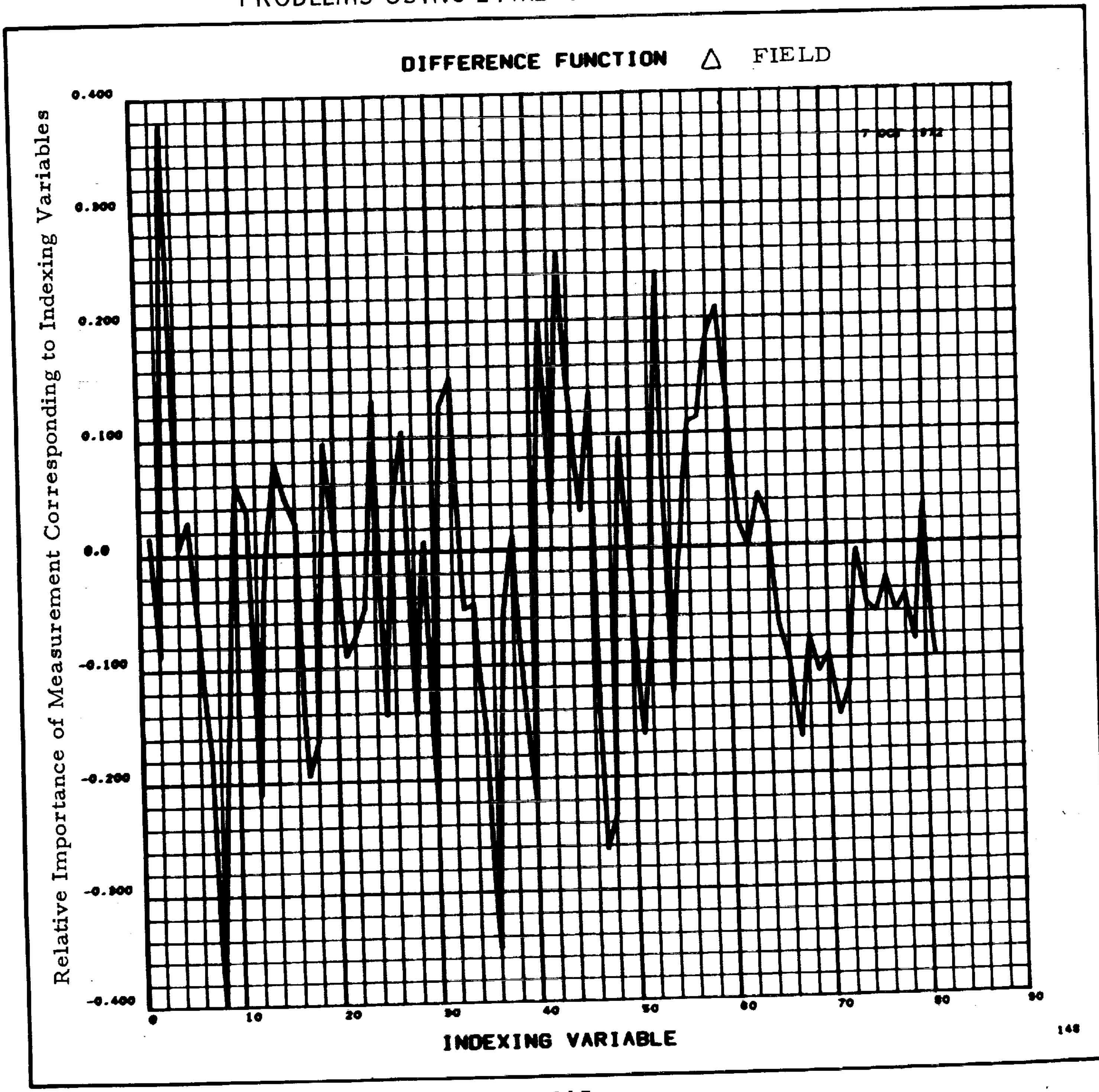


FIGURE 5.33 - RELATIVE IMPORTANCE OF MEASUREMENTS FOR DETECTING FAILURES OF THE ATOMIZING BOILER USING 38 DIMENSIONS

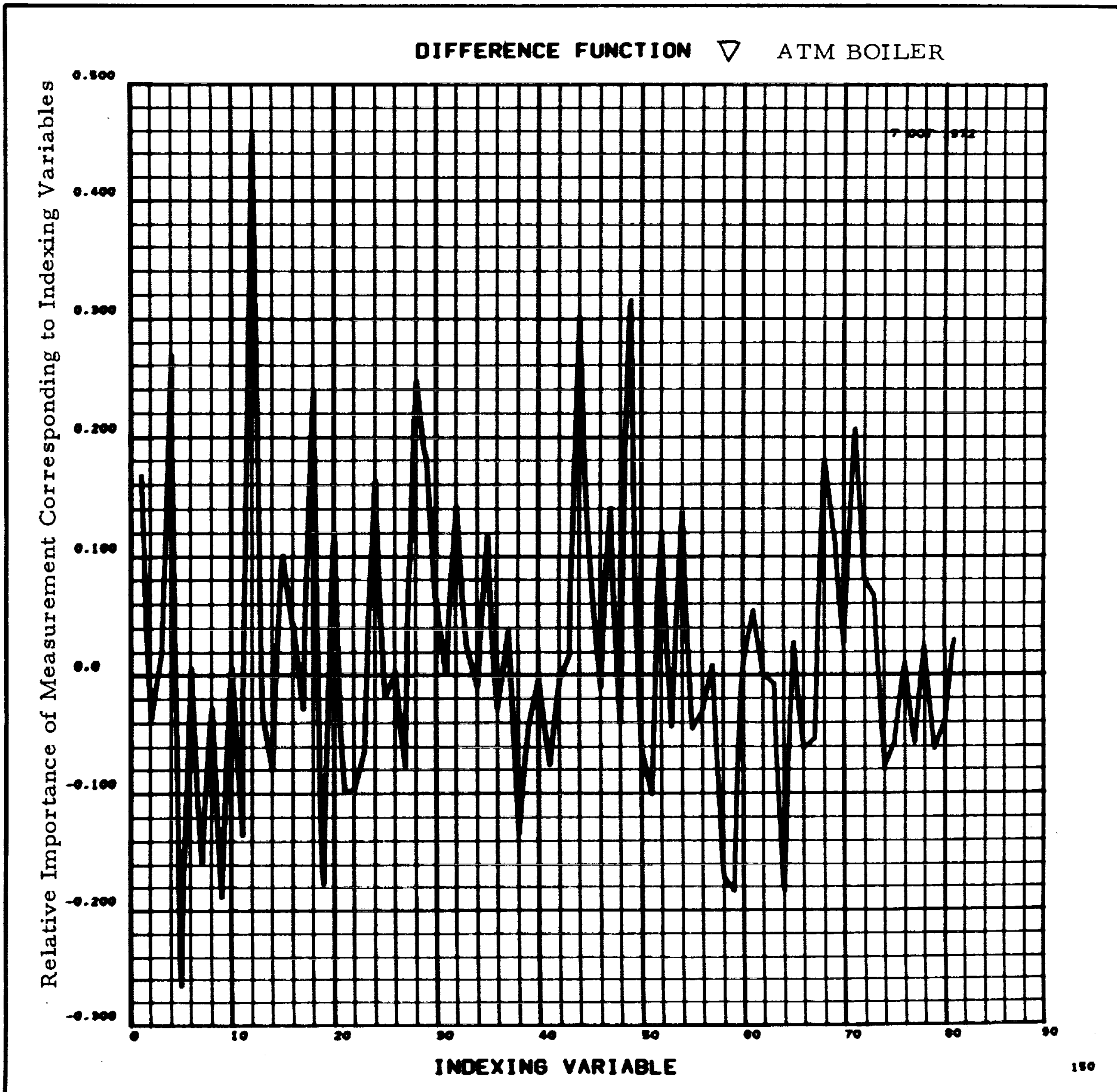


FIGURE 5.34 - RELATIVE IMPORTANCE OF MEASUREMENTS FOR DETECTING FAILURES IN SCATTER PLOT GROUP 2 USING 20 DIMENSIONS

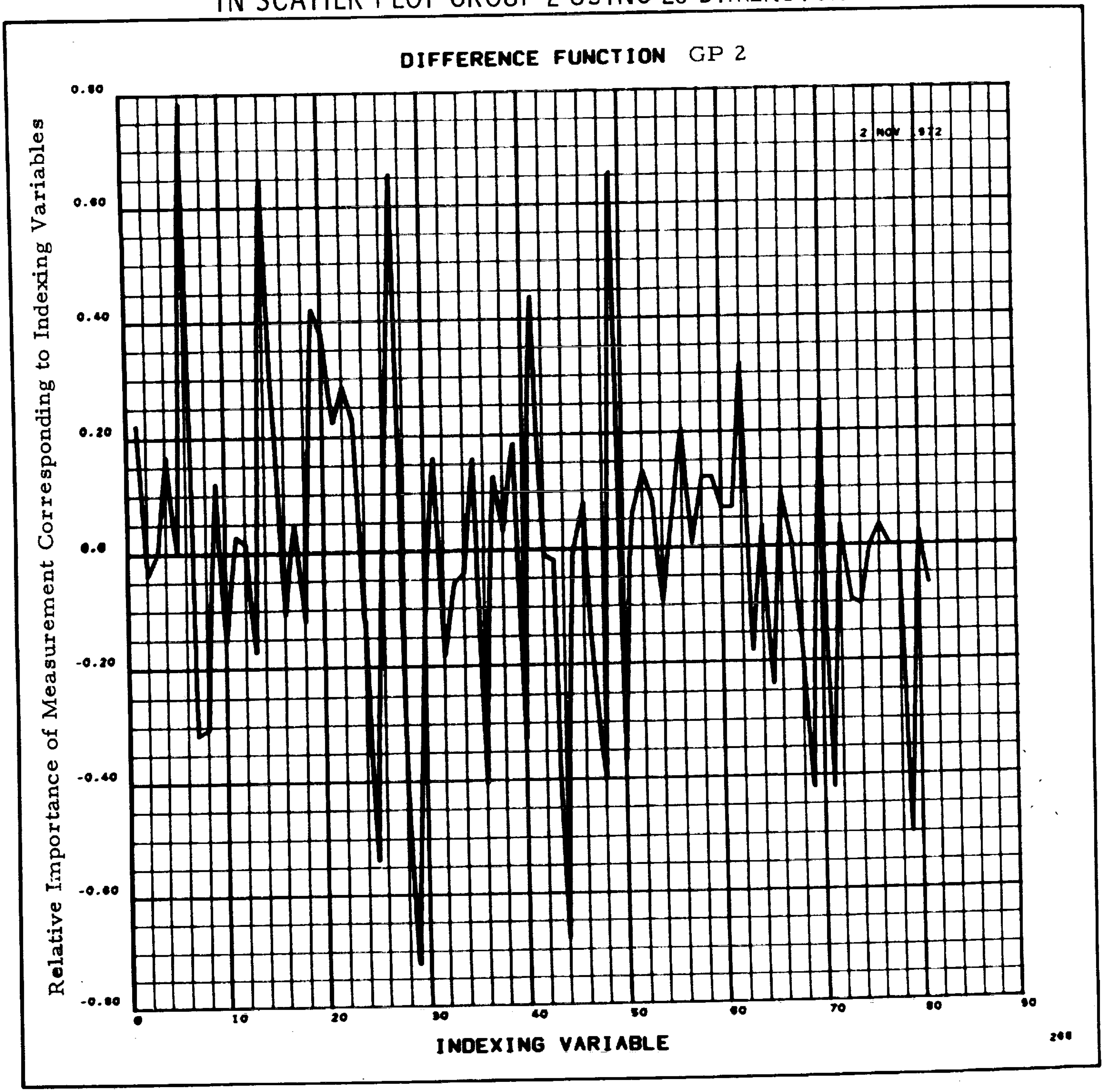


FIGURE 5.35 - VARIATION OF INFORMATION RETAINED WITH DIMENSIONALITY FOR REFERENCE 192 MEASUREMENT BASE

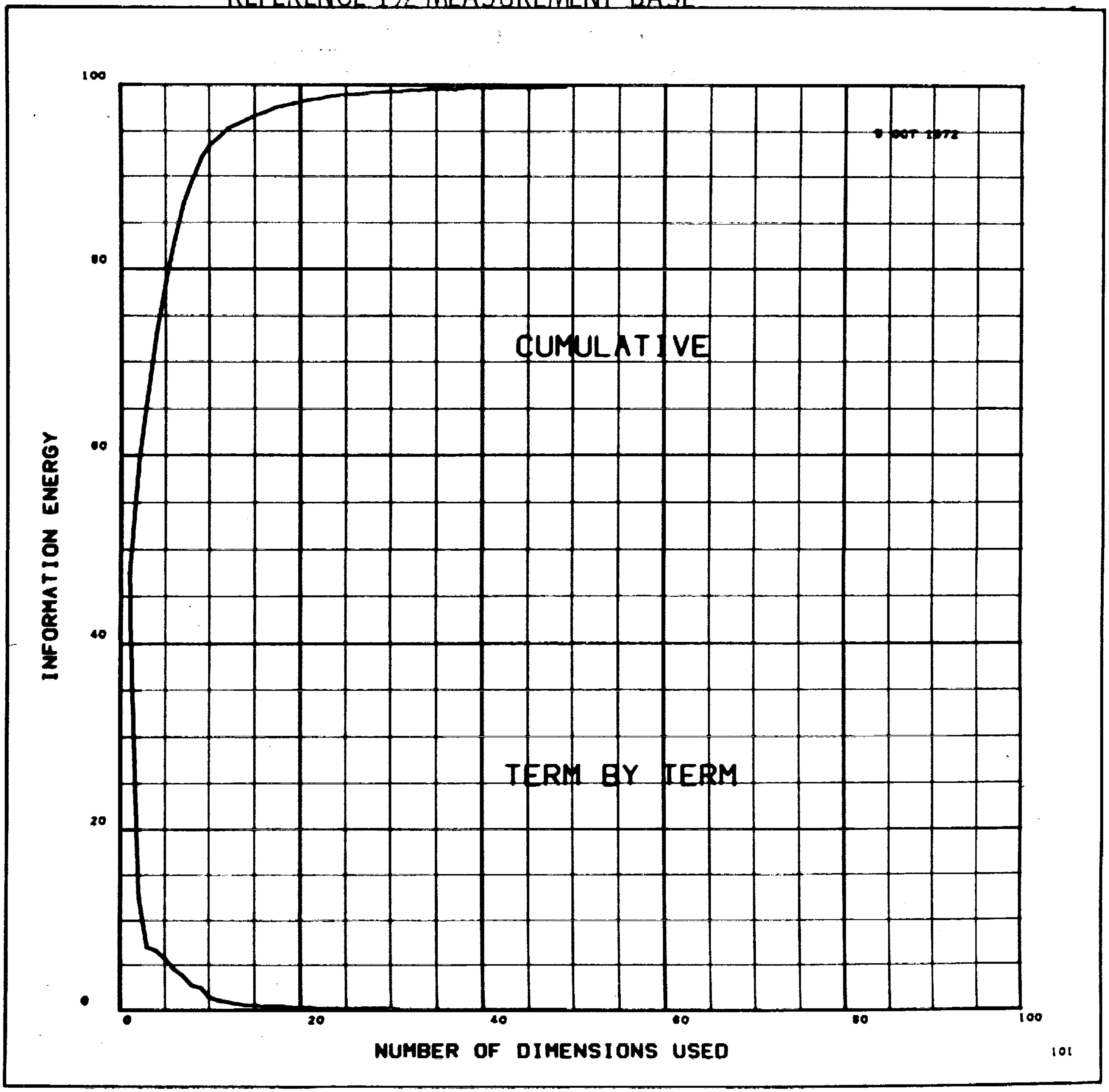


FIG. 5.36 - SCATTER PLOT OF THE FIRST AND SECOND COEFFICIENT OF THE GENERALIZED FOURIER SERIES REPRESENTATION OF THE CASES USED TO CONSTRUCT 1[illegible]2 MEASUREMENT REFERENCE BASE

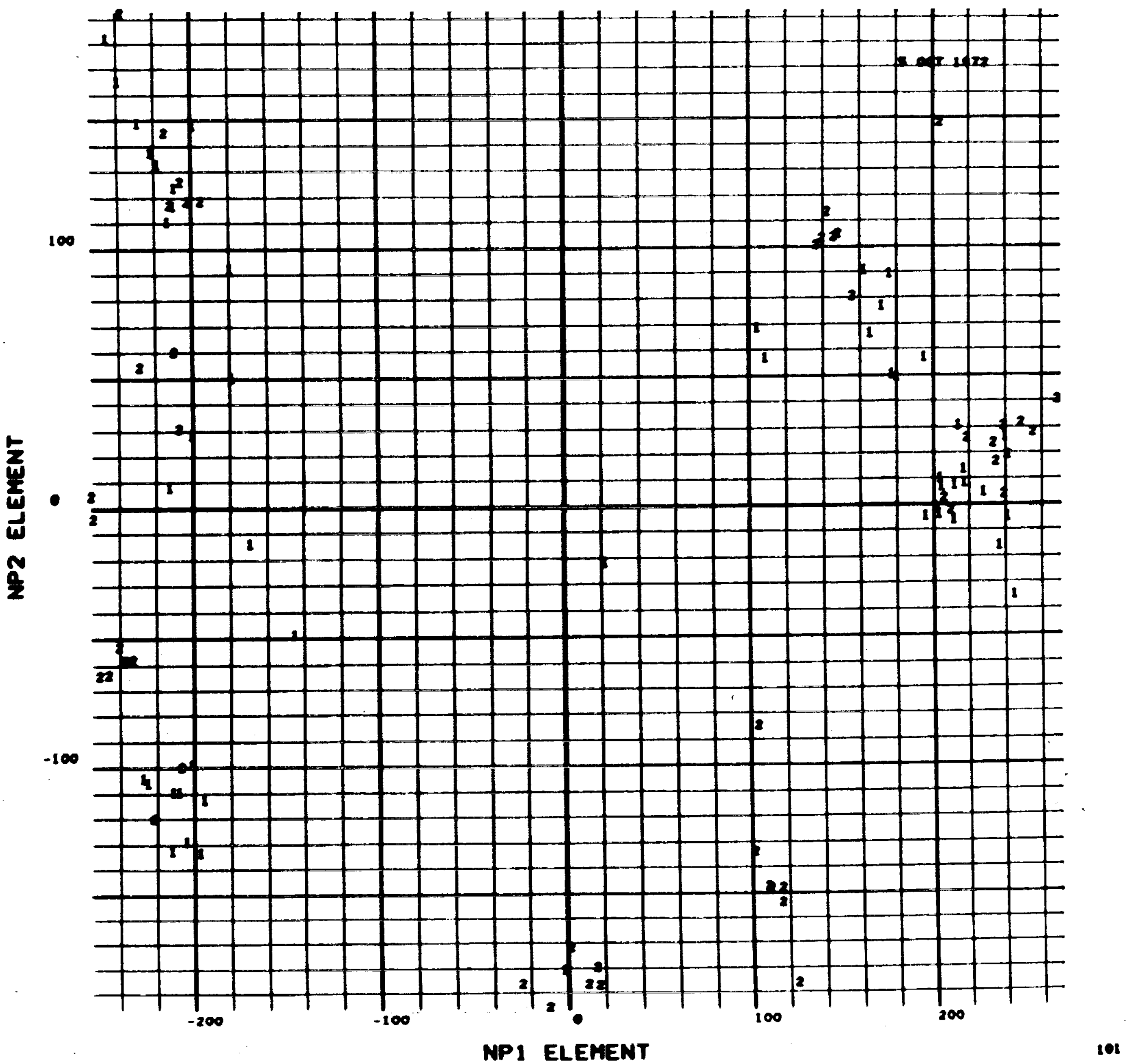


FIGURE 5.37 - FIRST OPTIMUM FUNCTION 192 MEASUREMENT REFERENCE BASE

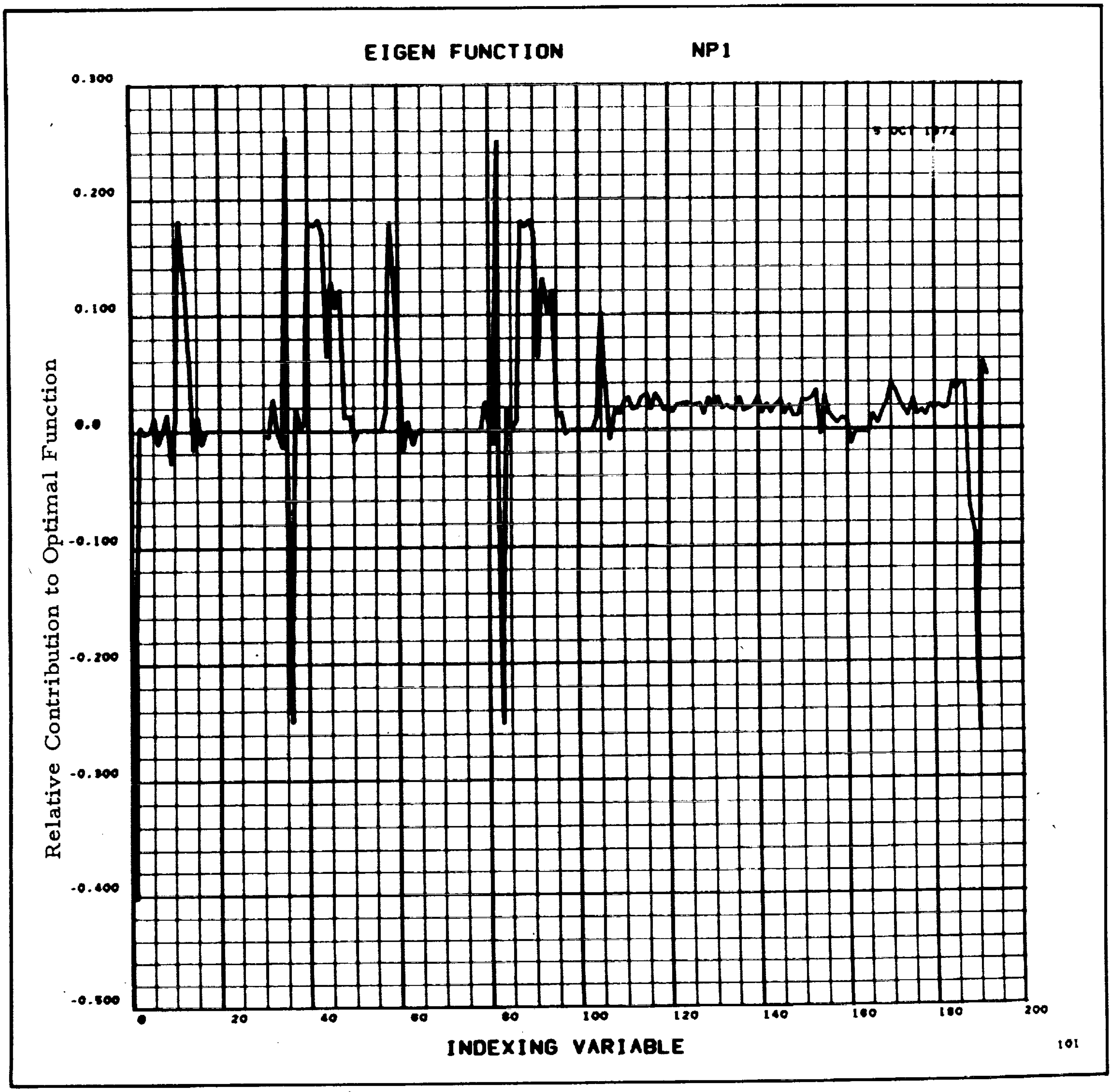


FIGURE 5.38 - SECOND OPTIMUM FUNCTION 192 MEASUREMENT REFERENCE BASE

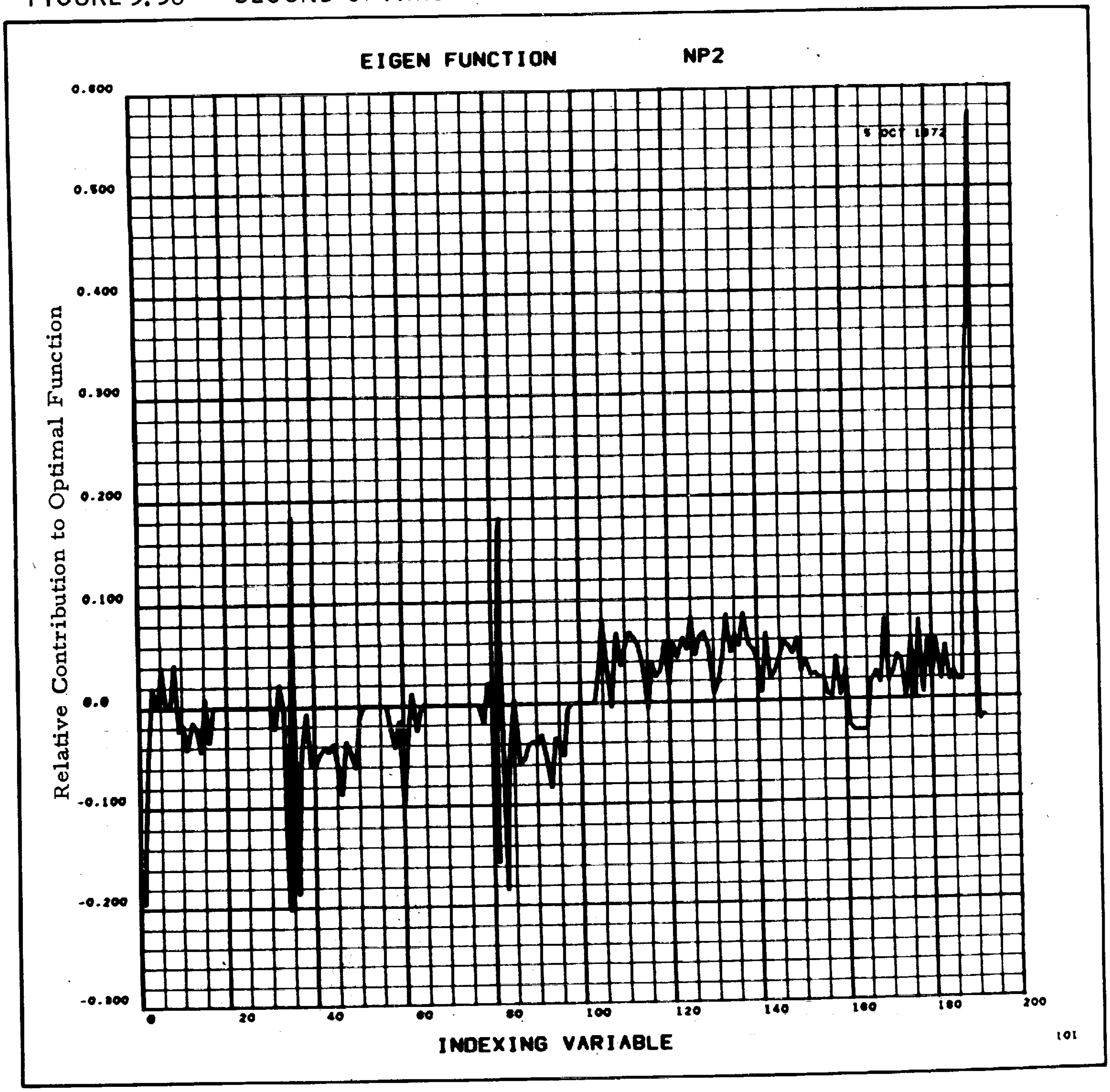


FIGURE 5.39

RELATIVE IMPORTANCE OF OPTIMAL DIRECTIONS FOR DETECTING INCIPIENT FAILURES

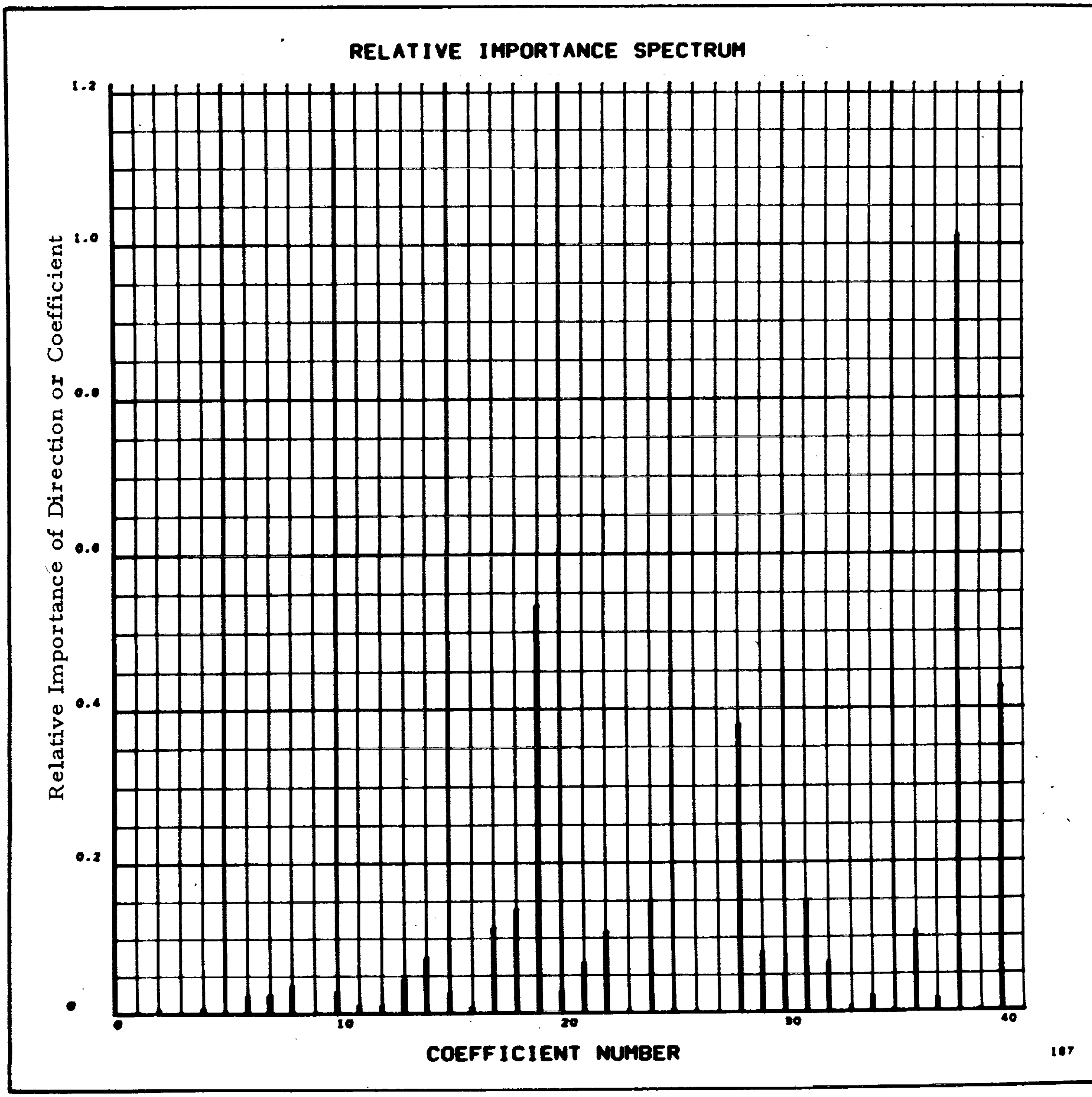


FIGURE 5.40 - PERFORMANCE MAP-UNIV BOILER #1 DETECTION ALGORITHM USING 50 MEASUREMENTS

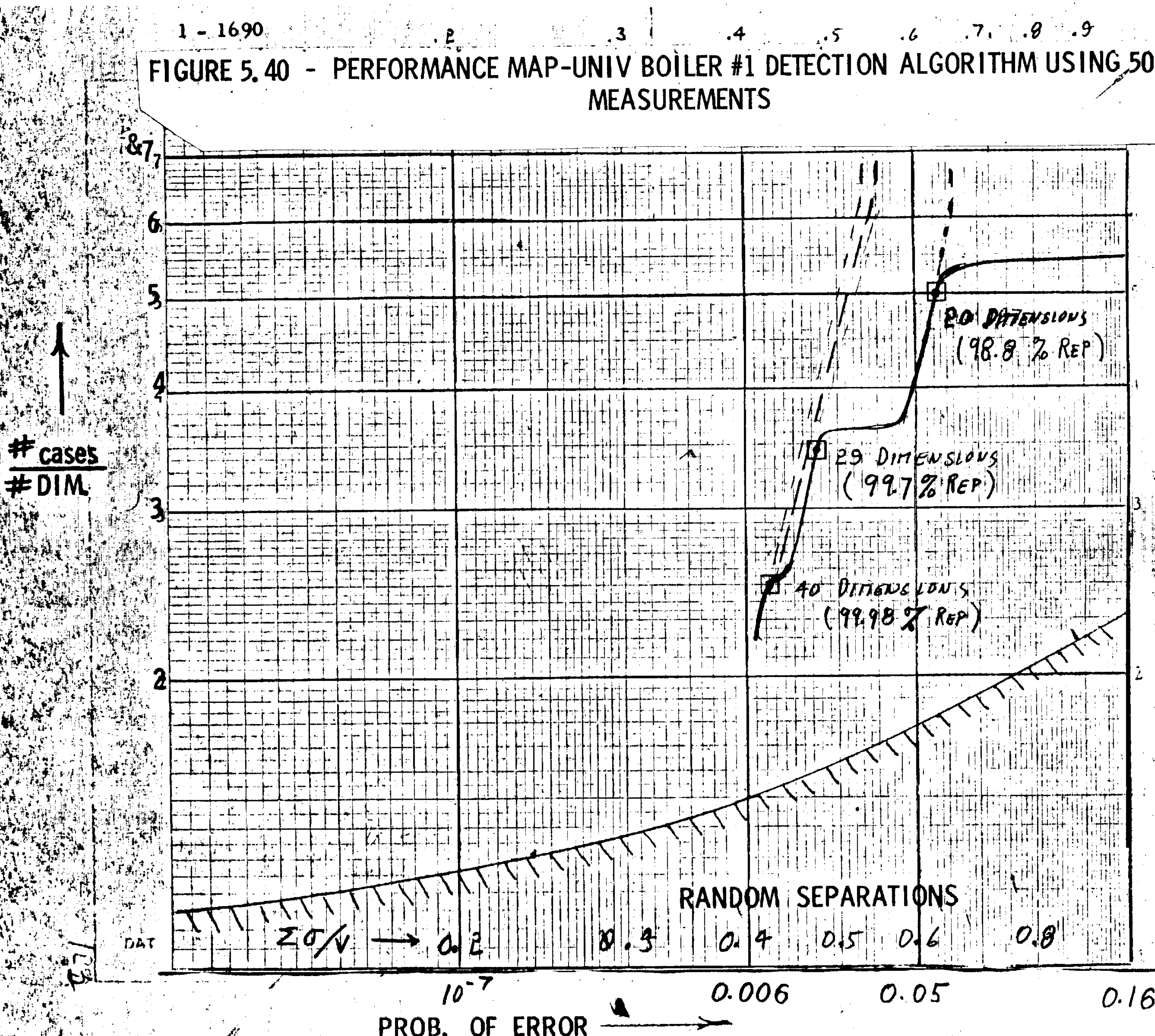
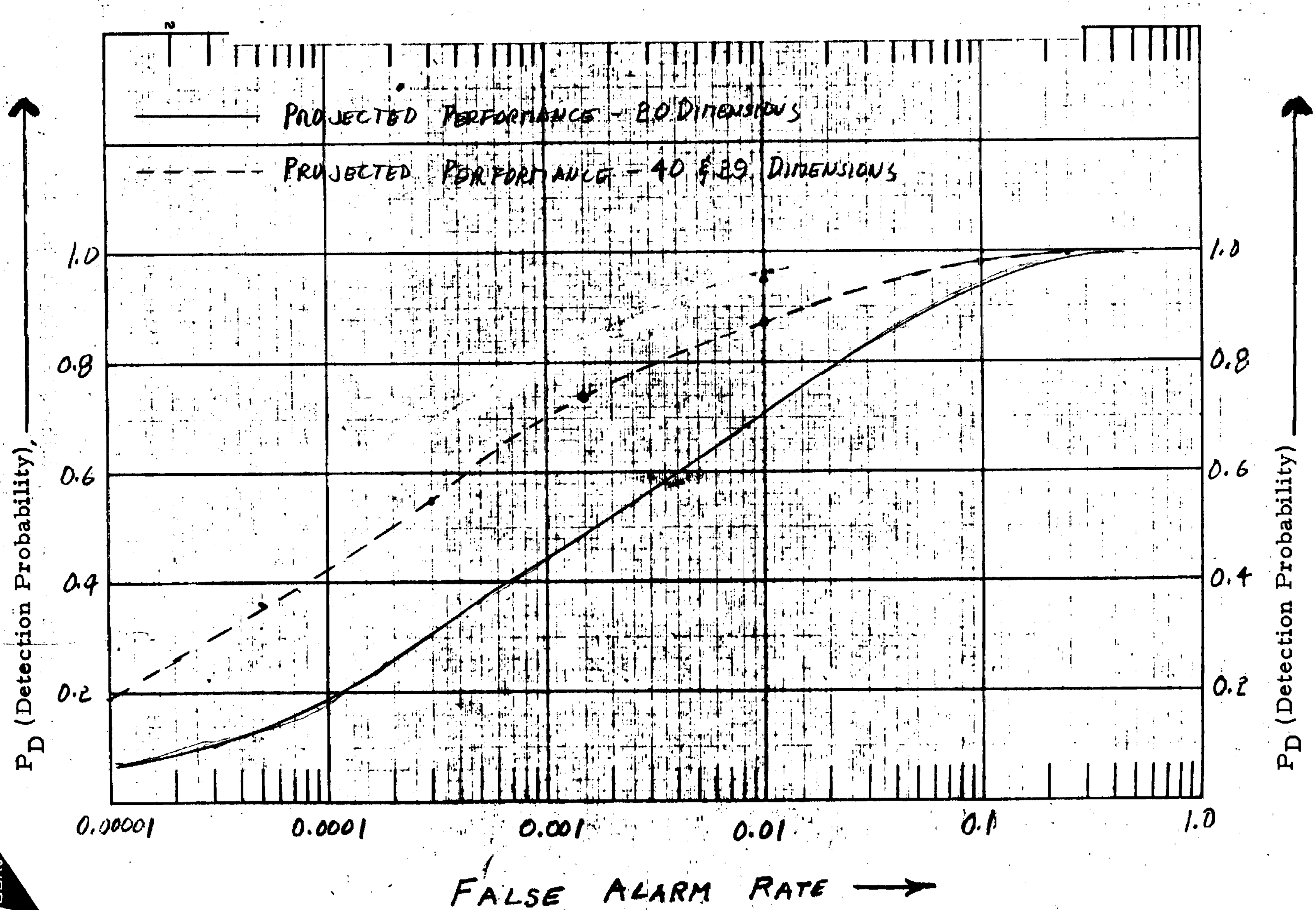


FIGURE 5.41 - CLASSIFICATION PERFORMANCE TRADE-OFF CURVE FOR COMPARING EXPECTED PERFORMANCE OF 50 MEASUREMENT ALGORITHMS



NOTE: USE TYPE B PENCIL FOR VUGRAPHS AND REPORT DATA.

FIGURE 5.42 RELATIVE IMPORTANCE OF MEASUREMENTS FOR 40 DIMENSIONAL UNIVERSAL DETECTION ALGORITHM

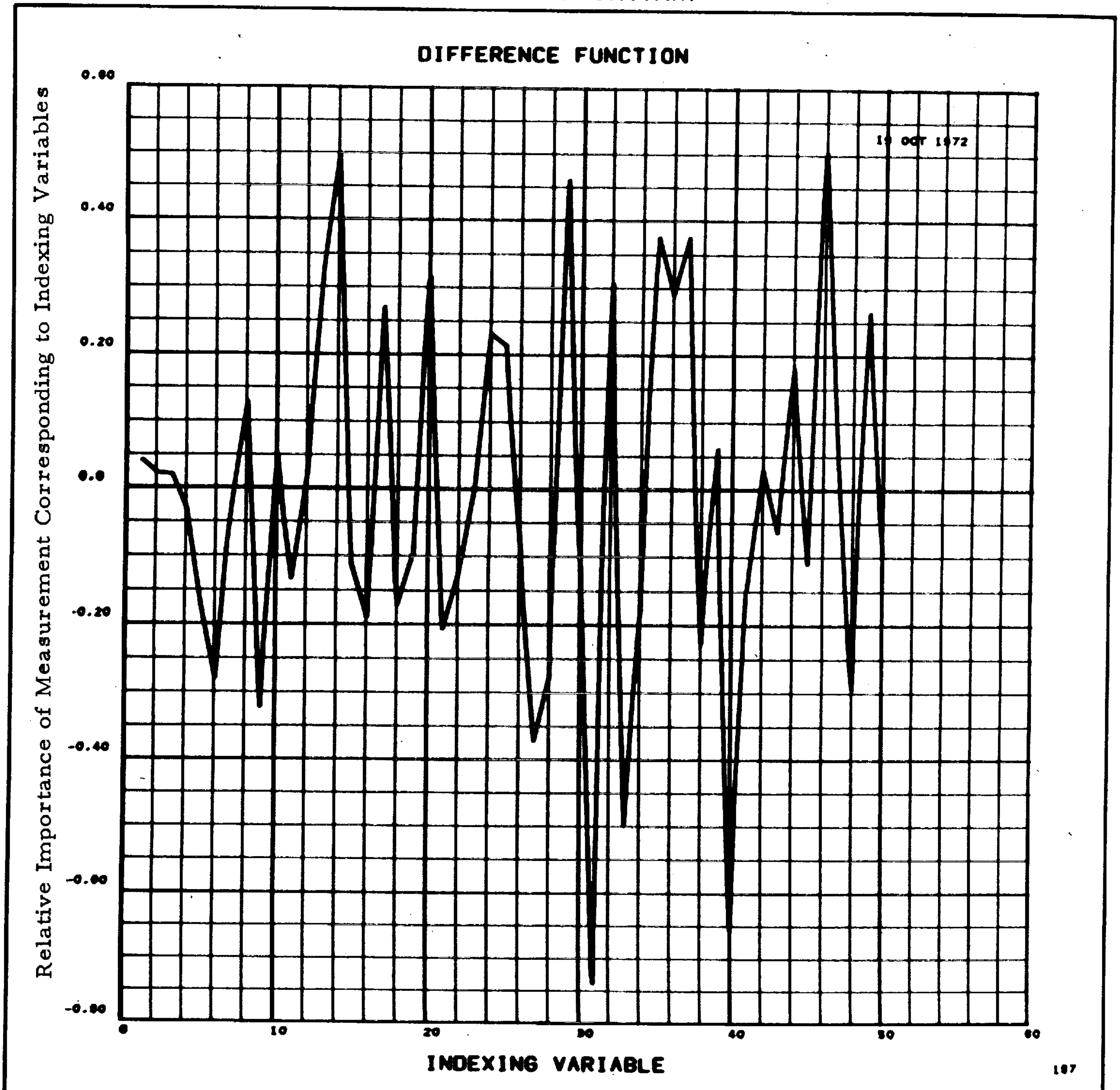


FIGURE 5.43 - RELATIVE IMPORTANCE OF MEASUREMENTS FOR 29 DIMENSIONAL UNIVERSAL DETECTION ALGORITHM

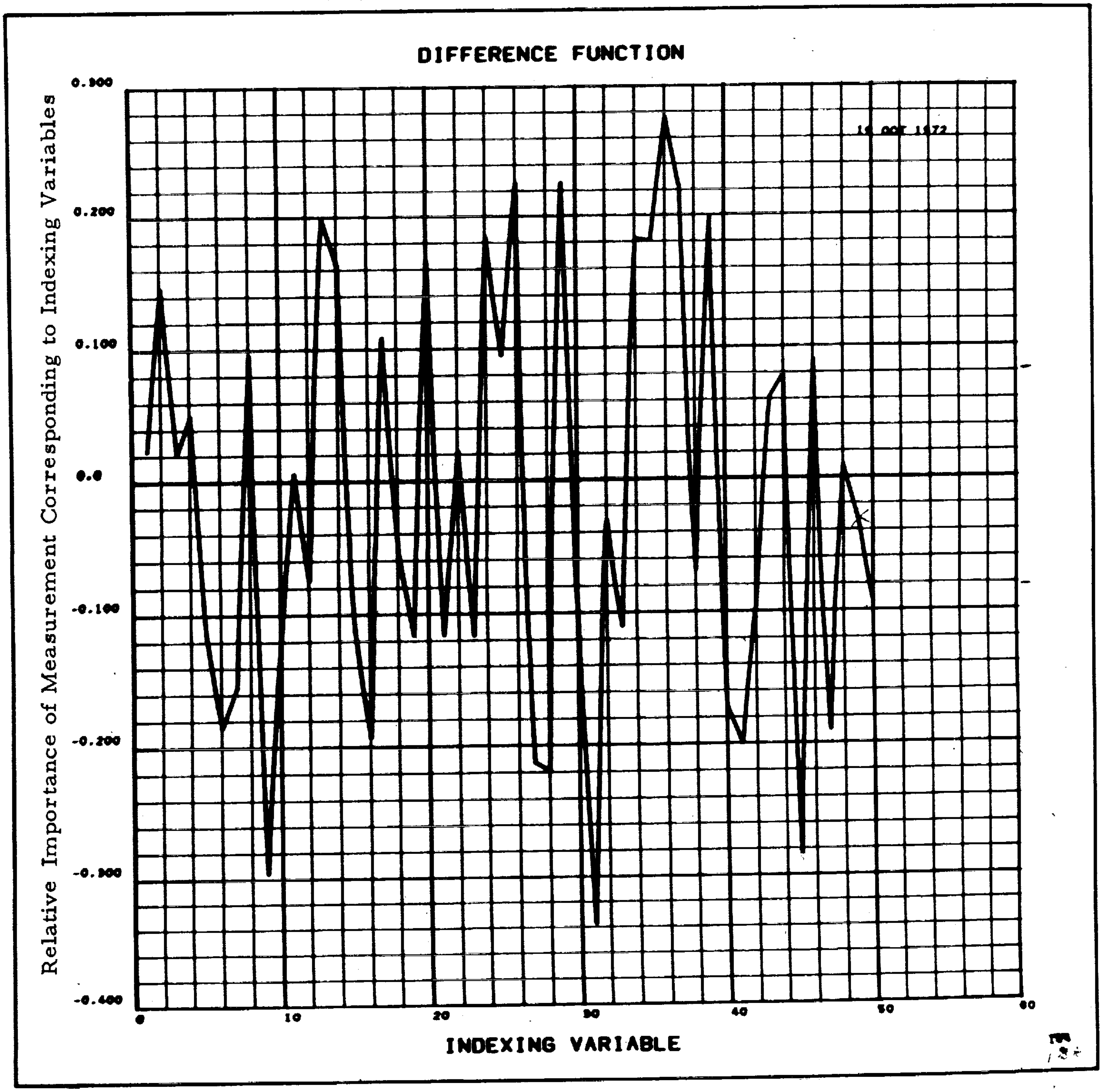
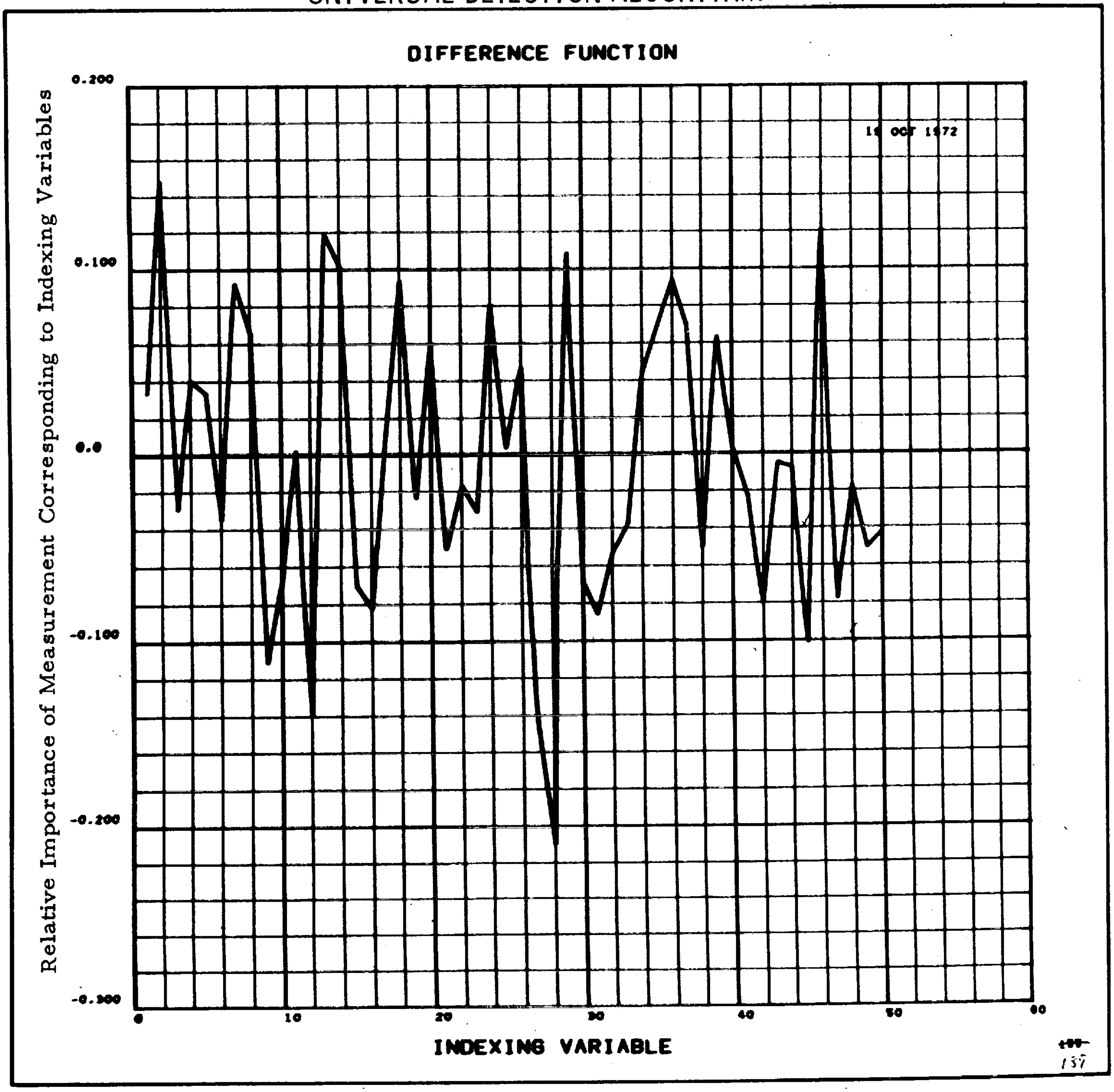


FIGURE 5.44 - RELATIVE IMPORTANCE OF MEASUREMENTS FOR 20 DIMENSIONAL UNIVERSAL DETECTION ALGORITHM



137

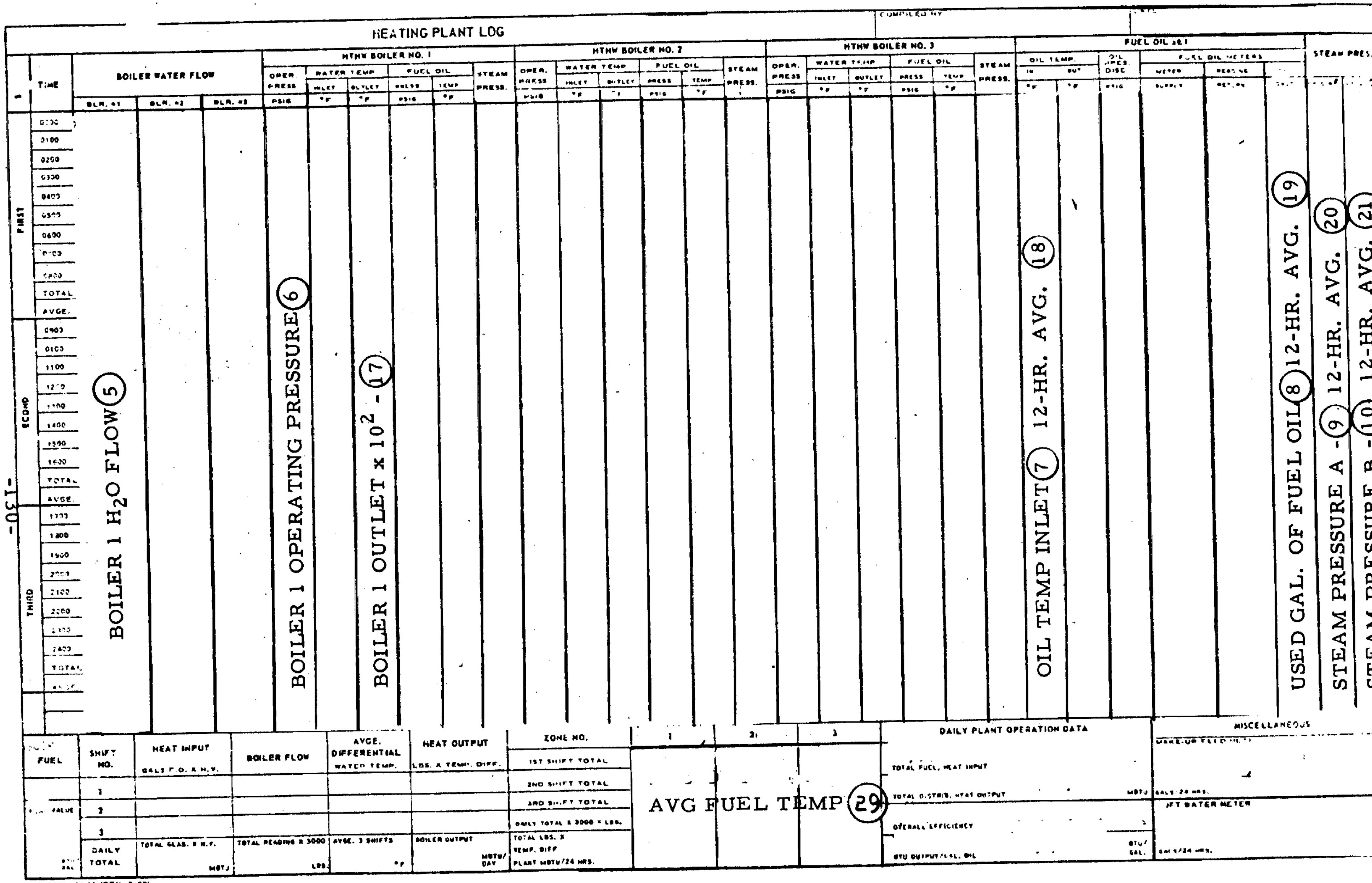


FIGURE 5.45

NOMENCLATURE FOR VARIABLES OBTAINED FROM THE HEATING PLANT LOG

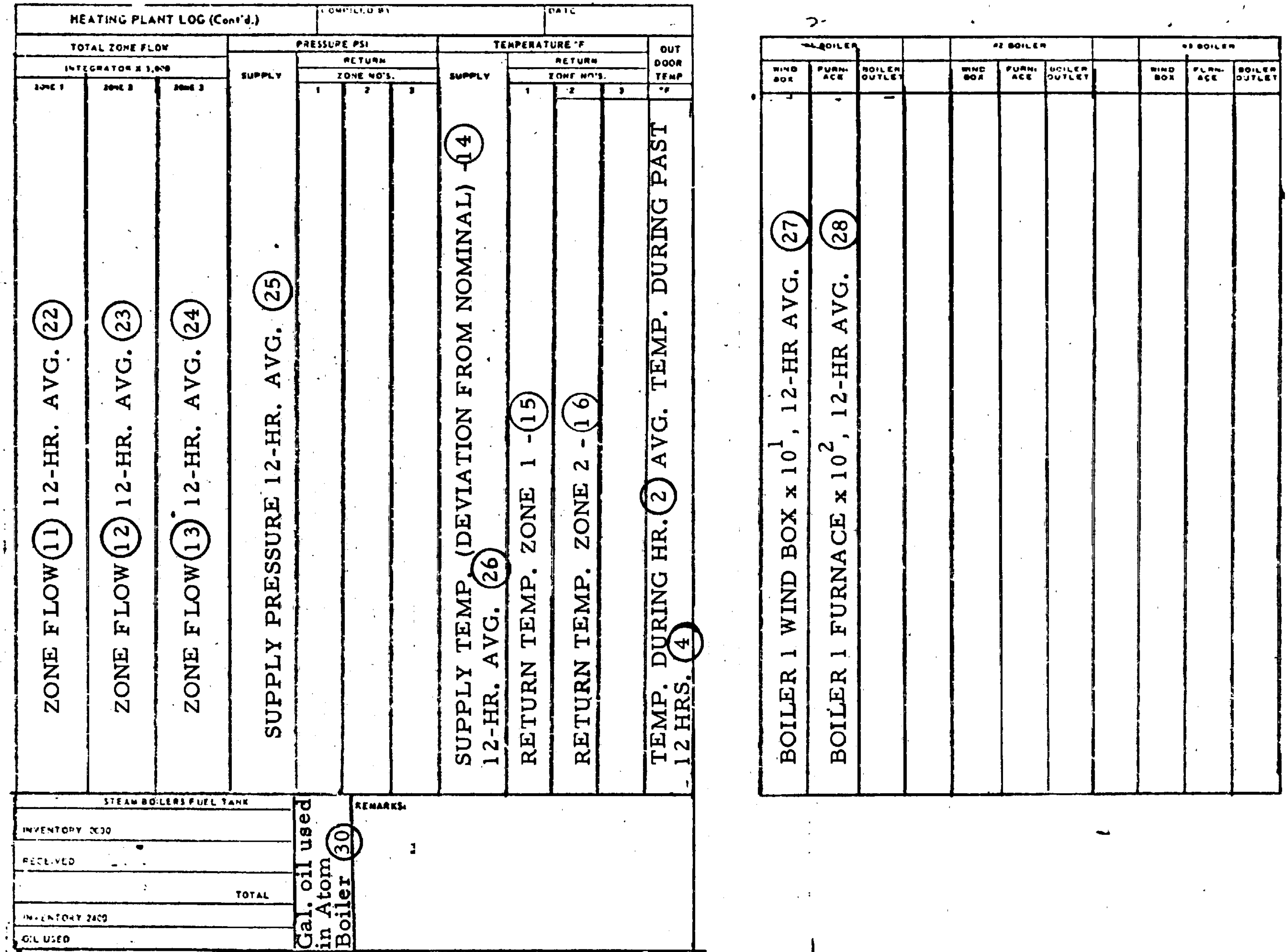


FIGURE 5.45 (continued)
NOMENCLATURE FOR VARIABLES OBTAINED FROM HEATING PLANT LOG

FIGURE 5.46 - SCATTER PLOT OF FIRST AND SECOND COEFFICIENTS OF GENERALIZED FOURIER SERIES REPRESENTATION OF NON-FAILING TEST CASES

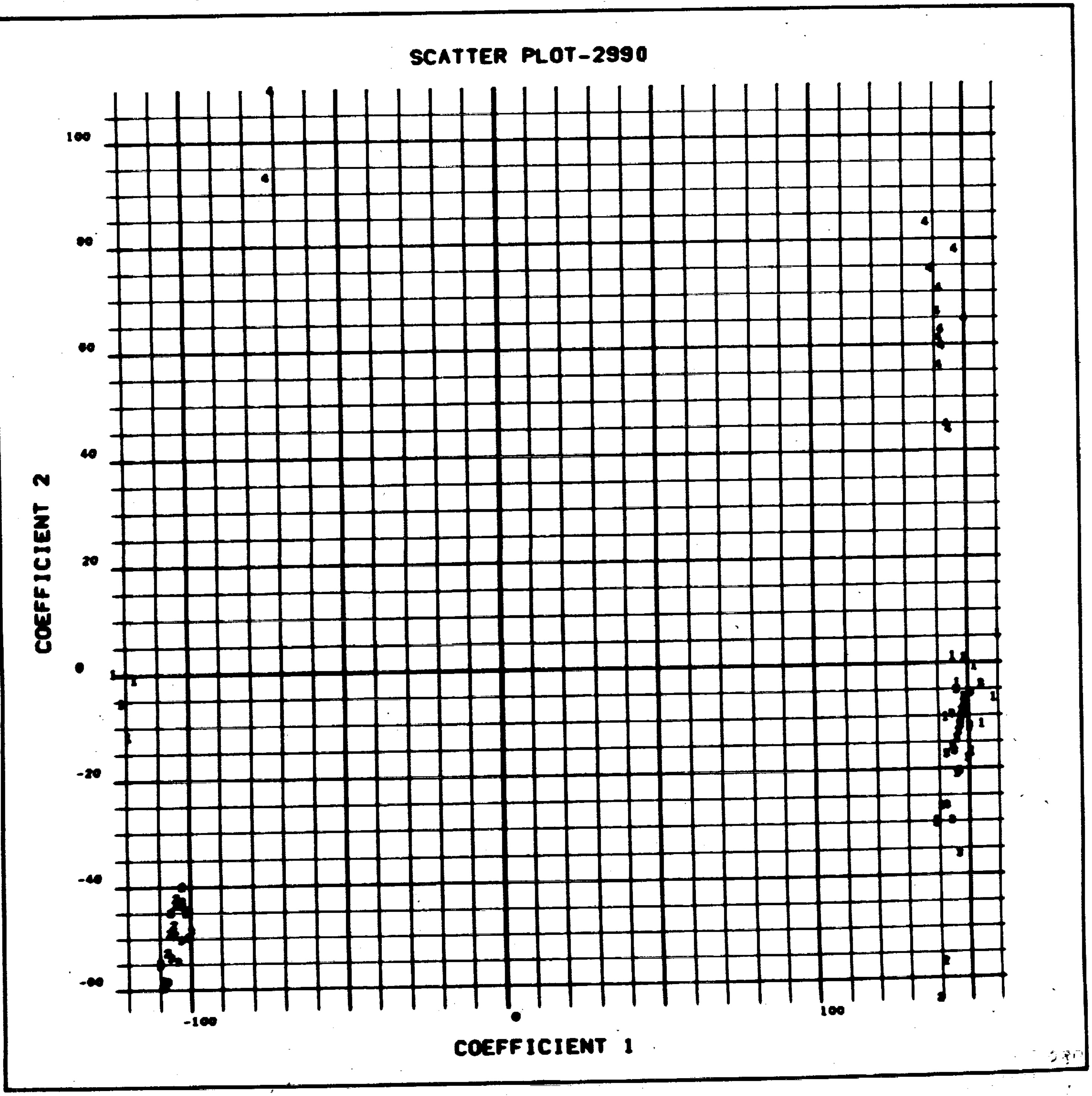
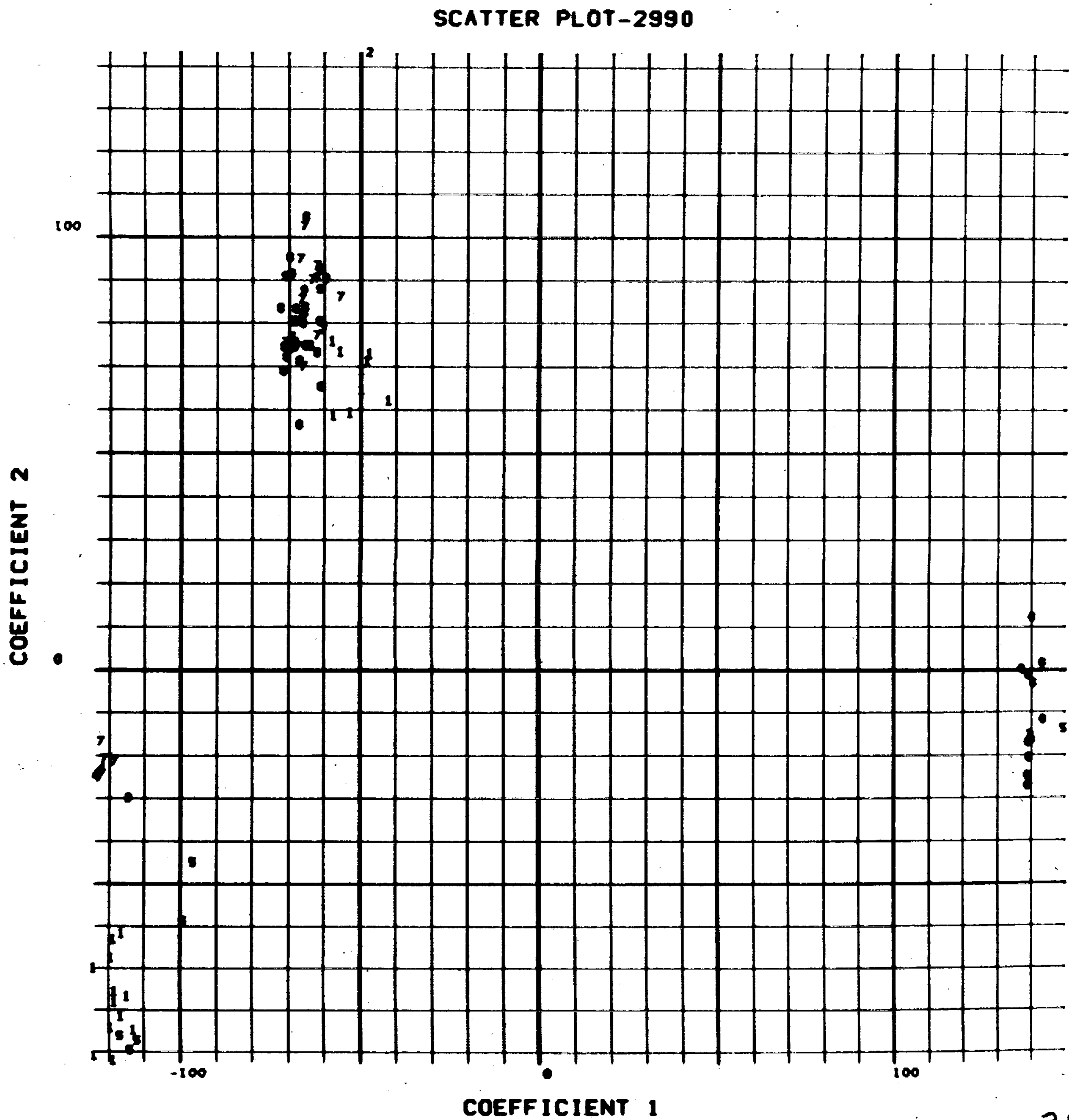


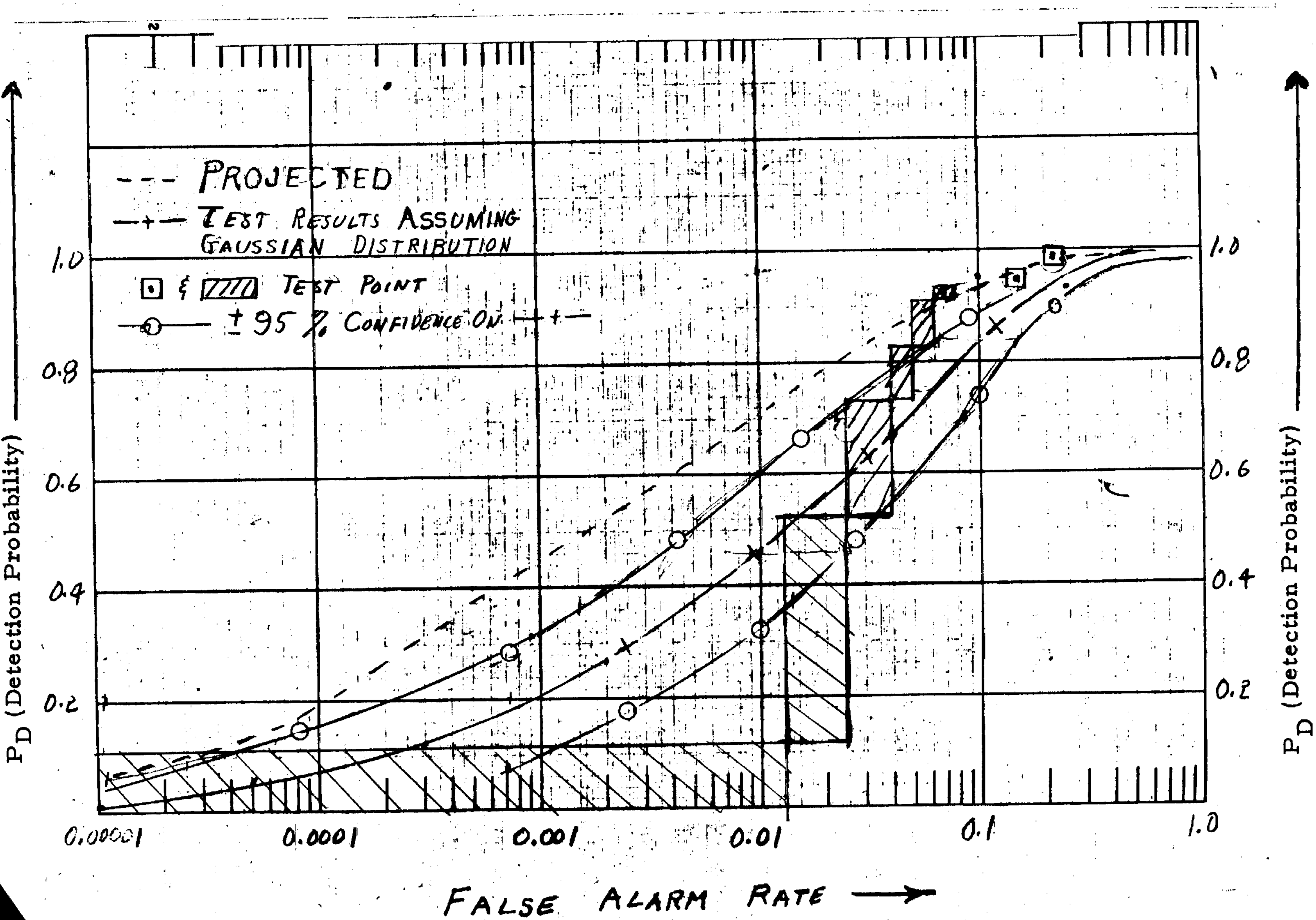
FIGURE 5.47

SCATTER PLOT OF FIRST AND SECOND COEFFICIENTS OF GENERALIZED FOURIER SERIES REPRESENTATION FOR TEST CASES IN WHICH HEATING PLANT FAILURES OCCURRED



295

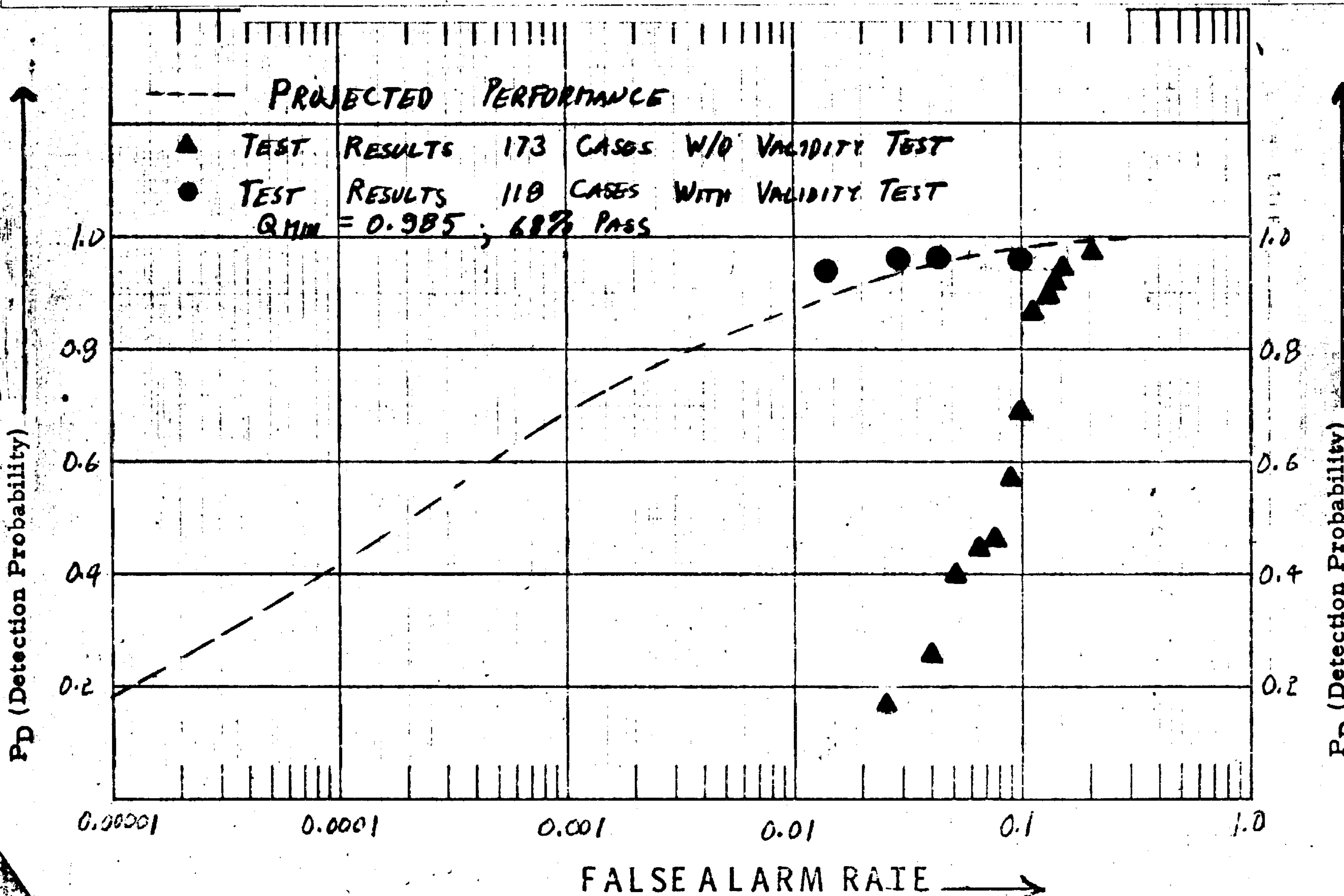
FIGURE 5.48 - CLASSIFICATION PERFORMANCE TRADE-OFF CURVES FOR COMPARING PROJECTED WITH TEST PERFORMANCE FOR 20 DIMENSIONAL 50 MEASUREMENT ALGORITHM



NOTE: USE TYPE B PENCIL FOR VUGRAPHS AND REPORT DATA.

FIGURE 5.49

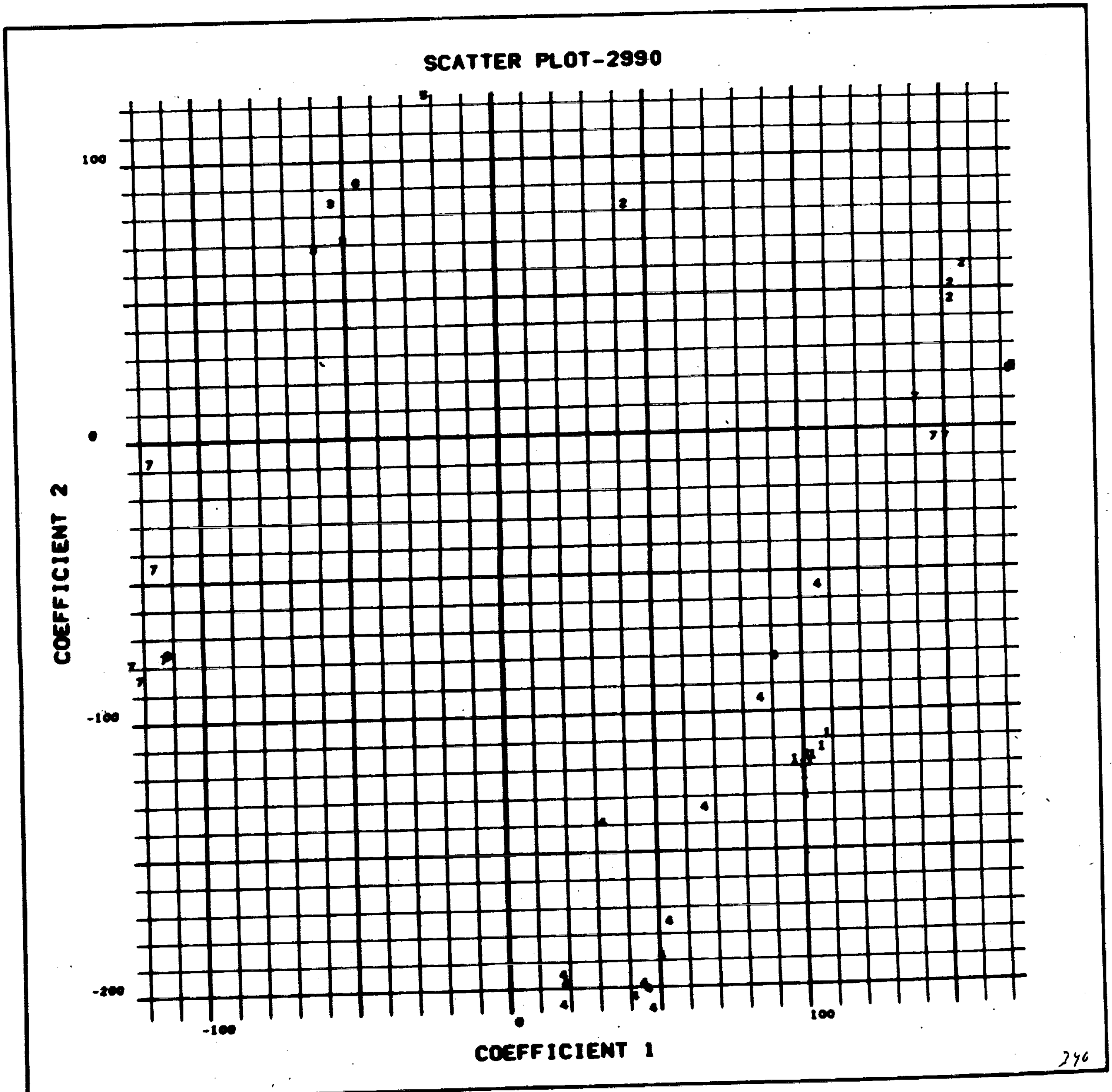
COMPARISON OF CLASSIFICATION PERFORMANCE TRADE-OFF CURVES FOR PROJECTED AND TEST PERFORMANCE OF 29 DIMENSIONAL UNIVERSAL BOILER NO. 1 DETECTION ALGORITHM



NOTE: USE TYPE B PENCIL FOR VUGRAPHS AND REPORT DATA.

FIGURE 5.50

SCATTER PLOT OF FIRST AND SECOND COEFFICIENTS OF GENERALIZED FOURIER SERIES REPRESENTATION OF EVALUATION TEST CASES



6.0 DIAGNOSTIC ALGORITHMS

The previous section showed that it would be feasible to develop algorithms to detect incipient failures for the KSC central heat plant. Thus, it becomes meaningful to determine if this same data can be used to diagnose what component or region of the central heat plant is about to fail. Although success of a diagnostic algorithm is not absolutely essential to establish the usefulness of ADAPT algorithms for demand maintenance of a system such as this, the availability of diagnostic algorithms will considerably improve this capability. Diagnosis can in principal be performed in two ways: 1) develop detection algorithms for detecting specific faults versus all other cases, and 2) if one has already performed the detection, diagnostic algorithm is a classification algorithm separating the particular failure of interest from all other failures. Since it is quite likely that many failures will have much in common, the prognosis of successfully developing diagnostic algorithms is better for the second approach, that is separating a particular failure from all other failures than for the first approach. This section will show the feasibility of developing diagnostic algorithms of this second kind based on the central heat plant data similar to that used to develop a diagnostic algorithm in the preceding section.

Since a separate diagnostic algorithm must be developed for each failure mode, the cost of developing the family of diagnostic algorithms is considerably greater than the corresponding cost of developing the detection algorithm. Furthermore, the feasibility of the predictive preventive maintenance system is not as critically dependent on the ability to develop a particular failure diagnostic algorithm as it was on the ability to develop algorithms to detect incipient failures. Thus, the feasibility demonstration will not include the derivation of a complete set of operational diagnostic algorithms but will be limited to demonstrating the feasibility on two typical failures. The feasibility will be demonstrated based on the projected performance of the initial exploratory algorithm development. The ability to achieve and improve on this projected performance was demonstrated on the detection algorithms as described in the preceding section. Although a similar demonstration of the optimization and proof testing of these algorithms can be achieved from a technical standpoint, the task is more appropriate to the development of the predictive preventive maintenance system than to the demonstration of the feasibility of this system.

In principal, it would not be necessary to develop a new base to develop the diagnostic algorithms. However, one might expect that in general the performance will be slightly better if the base used to develop a given diagnostic algorithm has been derived for that specific task and that the results achieved will more typical of those that would be expected for the development of diagnostic algorithms for different types of failures if a new base were developed for each of the two failure modes being investigated.

The first base which was developed was that to be used for diagnosing the failure of Boiler No. 1. This base was derived using a total of 50 cases, 13 failures of Boiler No. 1, and 37 of other types of failures. All of the analysis for the diagnostic algorithms were accomplished using all 192 measurements. Figures 6.1 thru 6.4 present principal characteristics of the representation for the base used for diagnosing the Boiler No. 1 failure. These figures may be compared with Figures 5.35 thru 5.38 which presented similar information for 192 measurement base derived for separating good cases for incipient failures. Comparison of Figures 6.1 and 5.35 shows that the representation for the diagnostic cases is easier than for the detection cases. This would be expected from the fact that the diagnostic base only includes failure type cases and need not account for the variation associated with good cases. Some of this difference is also probably due to the fact that the diagnostic base was constructed using half as many cases as the detection base.

Figure 6.2 presents the scatter plot for the detection of Boiler No. 1 failures. Note that on this scatter plot the numeral's 1 indicate those cases for which Boiler No. 1 failure occurred and the 2's indicate other types of failures. Thus, both the 1's and 2's would appear as 2's on Figure 5.36. In comparing Figures 6.2 and 5.36 the reader is cautioned that a double mirror image has occurred, that is both the first and second optimum functions have selected opposite signs for the base derived for detecting Boiler No. 1 failures. When this is accounted for the scatter plots are indeed quite similar. This is in agreement with the results that are obtained by examining Figures 6.3 and 6.4 and comparing them with Figures 5.37 and 5.38. In general, we see that the first optimum functions are very similar except that the sign is reversed. There are no qualitative difference between the first optimum functions of the boiler diagnostic base and the universal detection base and no significant quantitative differences. Again, when the sign is accounted for, careful examination of the second optimum function presented in Figures 6.4 and 5.38 shows only four significant qualitative differences between these two optimum functions. These differences are the appearance of spikes associated with variables 46, 60, 93, and 106 which appeared in the base for diagnosing Boiler No. 1 failures. Referring to Table 2.5 we see that these variables correspond to the return temperature associated with zone 2, the 12-hr. average of Boiler No. 1 outlet temperature, the 12-hr. average of the zone 2 return temperature, and the soft water meter. Recalling that these variables are important to failure diagnosis, it is not surprising that a base consistent only of failure cases would be more likely to include these variables earlier in the representation. This also suggests that further improvement in the diagnostic algorithms developed in Section 5 might be achieved by using a failure base rather than a combined base for deriving the universal detection algorithm. This is another possibility which might be investigated as part of the development program but which was not necessary to establish feasibility.

Figures 6.5 thru 6.8 present the corresponding information for the base used to derive the algorithm for diagnosing failures in the atomizing steam boiler. This base was derived using 75 cases, 35 of these cases represented failures of the atomizing steam boiler and 40 of the cases represented other failure types.

Comparison of the information energy curves presented in Figure 5.35, 6.1, and 6.5 shows that the difficulty of representation is approximately equal for both of the bases utilizing only the failure data. But both of these (i.e. Figures 6.1 and 6.5) are considerable easier to represent than the combined base of good cases and failure cases. Figure 6.6 presents the scatter plot presentation of the cases used to develop the algorithm for diagnosing atomizing steam boiler failures. On this figure the numerals 1 represent those cases for which atomizing steam boiler failures occur and numerals 2 are other failure cases. Comparison of Figures 6.2 and 6.6 shows that the scatter plots for these two bases are quite similar.

Comparison of Figures 5.37, 6.3, and 6.7 shows that there is still no significant variation both quantitative or qualitative between the first optimum functions of all three of these bases. However, comparison of the second optimum function shown in Figures 5.38, 6.4 and 6.8 shows that there is variation in a few variables between the base used to detect the atomizing steam boiler failure and the base for detecting the Boiler No. 1 failures. In particular, the soft water make-up feed meter measurement is the only measurement of the four measurements which differed between the Boiler No. 1 diagnostic base and the universal detection base which is still important in the atomizing steam boiler base. With the exception of the remaining three measurements and noting the mirror image effect, one sees that the second optimum function of the two diagnostic bases is still very similar and in fact in a qualitative sense much more similar than the universal detection base relative to either or both of the diagnostic bases.

The relative importance vectors for the two diagnostic algorithms derived on these bases, that is the algorithms for diagnosing Boiler No. 1 failures versus all other failures, and the algorithm for diagnosing atomizing steam boiler failures versus all other failures are presented in Figures 6.9 and 6.10. These figures provide a basis for reducing the number of measurements to be used for the diagnosis and thereby beginning the optimization process of the diagnostic algorithms. They also provide a basis for understanding the failure mechanisms and thus improving both the system and its maintenance. The reduction of the number of measurements would be a part of the development program to implement the predictive preventive maintenance approach. The analysis of the relative importance vectors to understand the failure mechanisms of the system although extremely useful are beyond the scope of the present study. However, these plots in conjunction with the corresponding relative importance plots presented in Section 5 provide the reader with the information required to carry out this analysis. The load on the system is very important to both the universal detection algorithm, the atomizing steam boiler detection algorithm, the atomizing steam boiler diagnostic algorithm, and the Boiler No. 1 diagnostic algorithm. This is shown by the fact that the amount of rainfall, the temperature and the maintenance on the sump pumps tend to be important for all of these algorithms. Figure 6.9 shows that both the 3-day and the 10-day rainfall are extremely important to diagnosing Boiler No. 1 failures. However, we note that the 3-day rainfall appears as a negative parameter and the 10-day rainfall as a positive parameter. This implies that Boiler No. 1 failure tends to occur later after a rain storm than the other failures. The fact that each of these variables has approximately equal absolute magnitude also suggests that since it is the difference between these rainfalls which actually enters the diagnostics that the actual relative importance of these two variables taken together may be significantly less than suggested by the initial cursory examination of Figure 6.9. Considerable information should be available from detail analysis of these figures and this is recommended as a further study program.

The performance of these two algorithms is summarized on Figure 6.11. The Boiler No. 1 diagnostic algorithm indicated by the triangles was developed at 20 and 15 dimensions. Both of these algorithms projected to the same performance indicating that a minimum of significant information is lost in the reduction from 20 to 15 dimensions. Examination of the relative importance spectrum for these algorithms indicated that this performance should continue clear down to 13 dimensions. The projected performance is obtained by extending the algorithm track with a fixed slope until the ratio of number of cases to number of dimensions is approximately 6. This yields an expected performance for the final Boiler No. 1 diagnostic algorithm of approximately .49 which corresponds to a probability of error of approximately 1 in 100.

The diagnostic algorithm for the atomizing steam boiler which was developed at 35 and 20 dimensions is also shown by the squares on Figure 6.11. This algorithm projects to considerably different performances for the 35 and 20 dimensional algorithm. This indicates, as does the relative importance spectrum, that significant information is lost as one decreases from 35 to 20 dimensions. For this reason the performance estimate has been based on projecting the performance of the 35 dimensional algorithm. This position appears to be sufficiently far from the random separations that the performance projection should be satisfactory. Project to performance parameter has a value of approximately .37 which corresponds to a probability of error between .001 and .005.

The trade-off between detection probability and false alarm rate which can be expected for these two algorithms is presented in Figure 6.23. Examination of this figure shows clearly that both of these algorithms have a performance which is completely adequate for application to the predictive preventive maintenance schemes outlined. Even without introducing techniques of multiple applications one finds that if false alarm rates of 1 in 100, the detection probability are of the order of 98 to 100%. The performance shown on Figure 6.23 has been projected in exactly the same manner as performance for the universal detection algorithm was projected in Sections 5.2 and 5.3. These projections were verified by the tests presented in Section 5.4 and thus it is believed that the performance shown in Figure 6.23 can be taken as a demonstration of the feasibility of utilizing the ADAPT programs to derive the diagnostic algorithms required for the straight forward implementation of a predictive preventive maintenance system.

E

The Boiler No. 1 diagnostic algorithm provides an excellent example for illustrating the effect of the Fisher weighting parameter on the trade-off between the detection probability and false alarm rate. Experience has shown that the performance parameter $\Sigma\sigma/V$ is relatively insensitive to the change in the Fisher weighting parameter (See Ref. 9). Experience has also shown that the parameter $2/\sqrt{V}$ which for the special case of $\sigma_1 = \sigma_2$ is identical to the performance parameter is quite sensitive to the Fisher weighting parameter. Thus, when projecting the performance of an algorithm to the case of $\sigma_1 = \sigma_2$, the projection must always be performed utilizing the performance parameter and calculating the corresponding value of V. This is illustrated by Figure 6.12 where the upward facing triangles present the performance trade-off for the Boiler No. 1 detection algorithm for the case of $\sigma_1 = \sigma_2$. The downward facing triangles present the same curve for the case of $\sigma_2 = 3.3\ \sigma_1$ which can be obtained from the same set of data as the upward facing triangle curve by simply changing the value of the Fisher weighting parameter. Although the downward facing triangle curve at first appears to have poorer performance than the curve for equal standard deviations, this is a result of the linear scale for detection probability. In fact, the downward facing triangle curve has significantly better detection probabilities for false alarm rates greater than .01. For example the detection probability at a false alarm rate of .04 is .9998 for the downward facing triangles and only .997 for the upper triangles and at .01 the downward facing triangles have a detection probability of .9999993 as compared to .99994 for the upward facing triangles. Thus the Fisher weighting parameter for the downward facing triangles is better for those cases where false alarm rates are greater than .01. The selection of the proper Fisher weighting parameter as a function of false alarm rate is discussed in more detail in Appendix C.

FIGURE 6.1 - VARIATION OF INFORMATION RETAINED WITH DIMENSIONALITY FOR FOR BOILER NO. 1 DIAGNOSTICS BASE

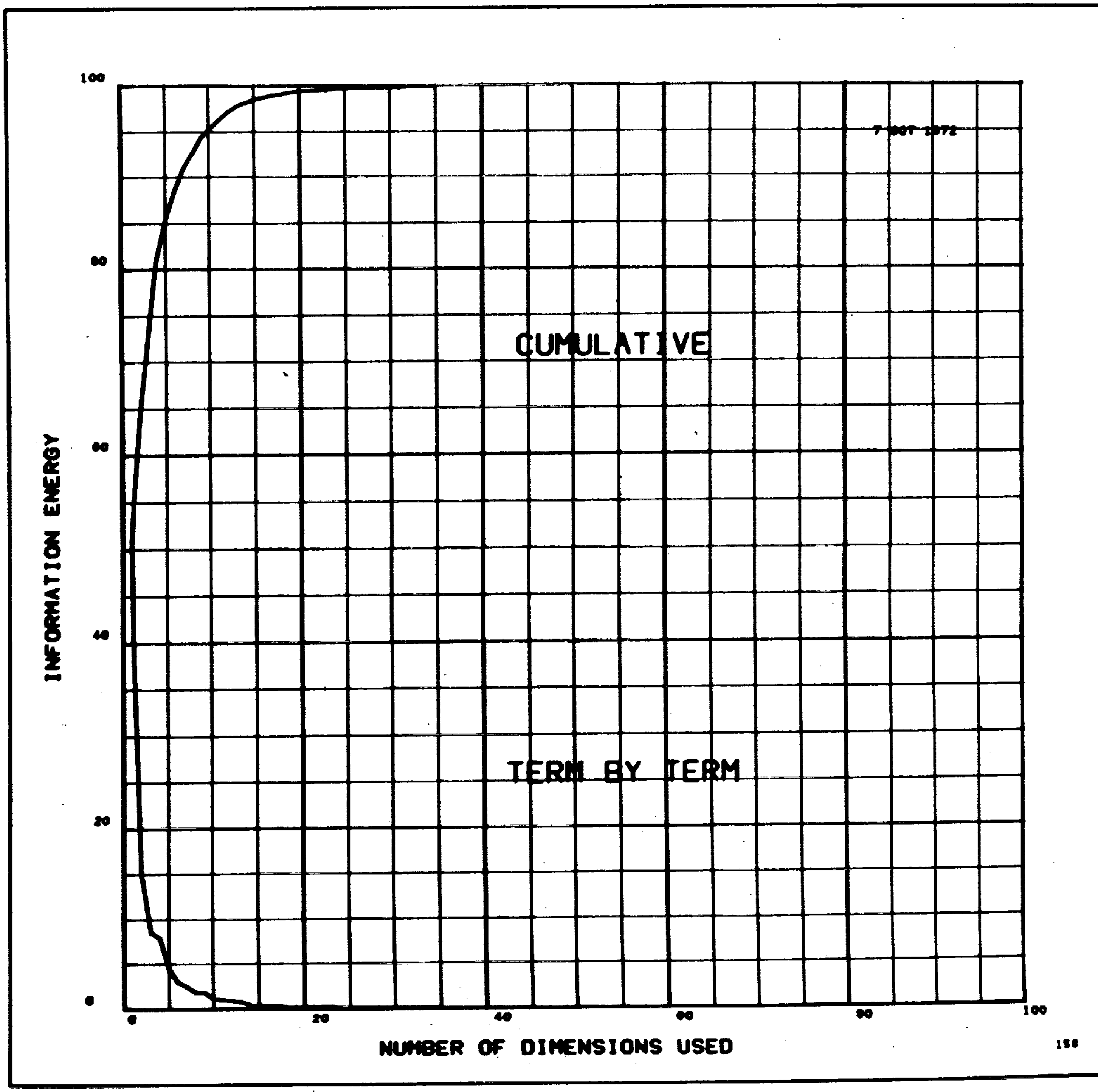


FIGURE 6.2 - SCATTER PLOT OF FIRST AND SECOND COEFFICIENTS OF GENERALIZED FOURIER SERIES REPRESENTATION OF DATA USED TO DEVELOP BOILER NO. 1 DIAGNOSTIC ALGORITHMS

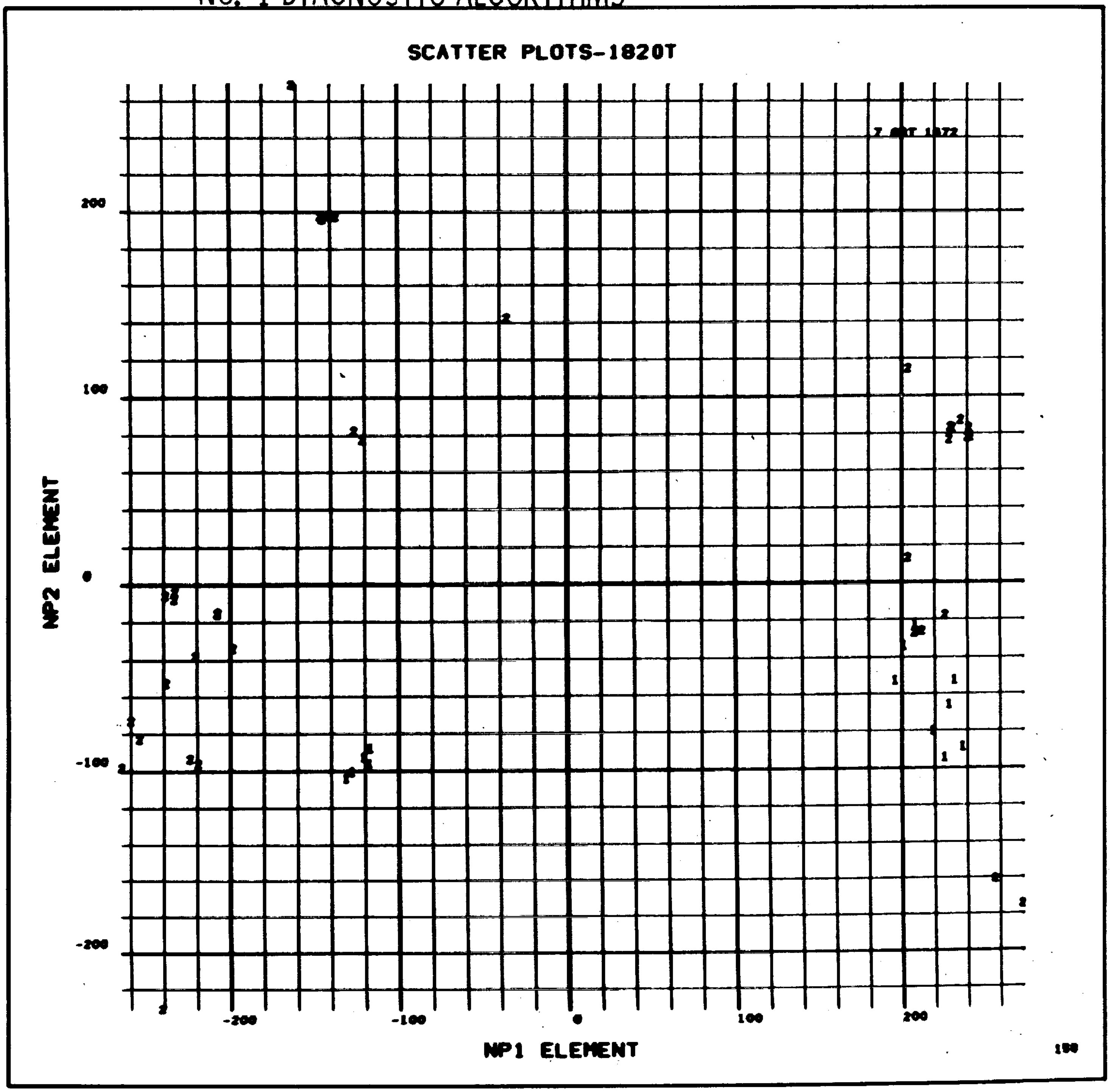


FIGURE 6.3
FIRST OPTIMUM FUNCTION FOR BOILER NO. 1 DIAGNOSTIC ALGORITHM BASE

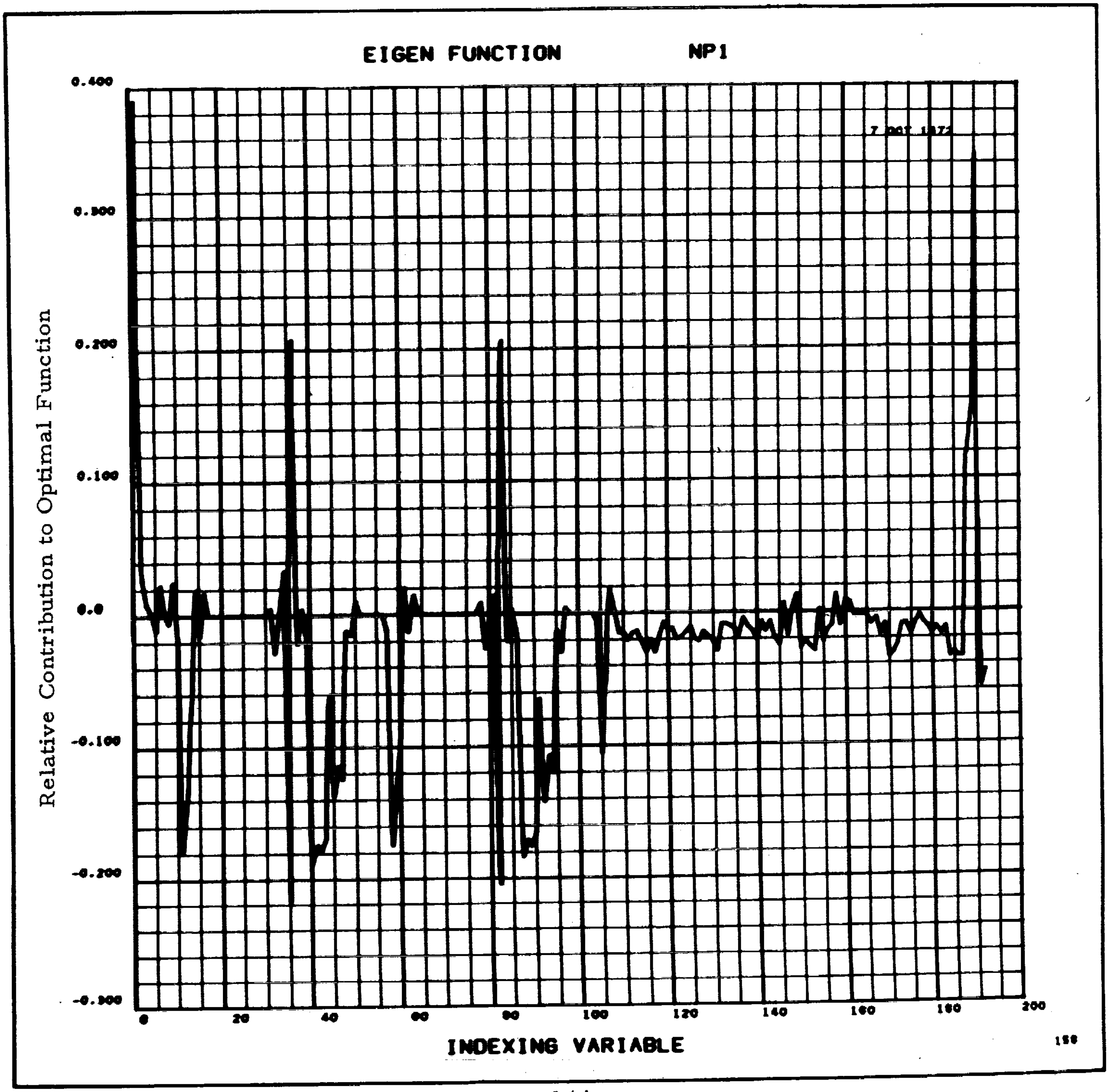


FIGURE 6.4
SECOND OPTIMUM FUNCTION FOR BOILER NO. 1 DIAGNOSTIC ALGORITHM BASE

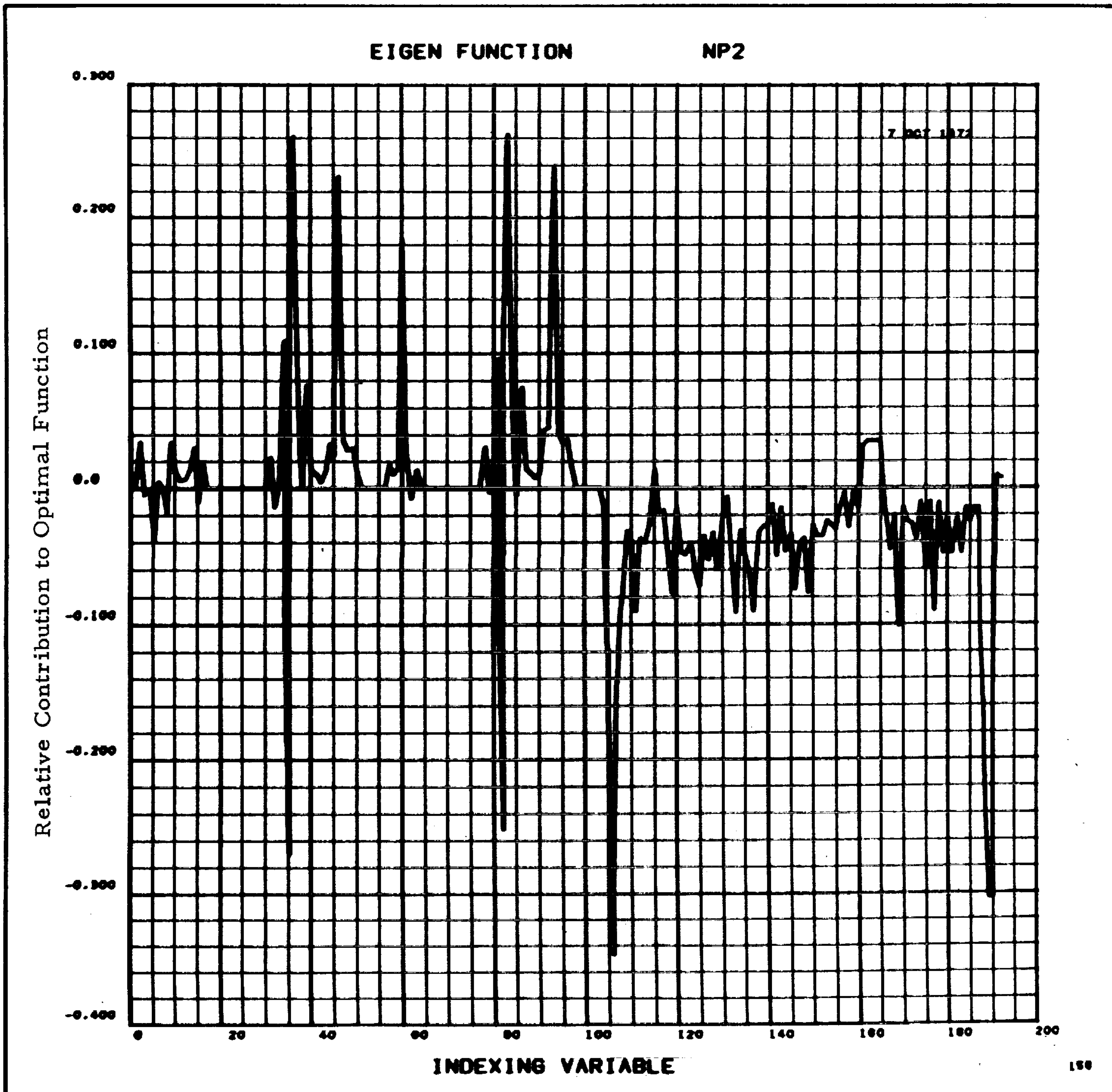


FIGURE 6.5

VARIATION OF INFORMATION ENERGY RETAINED WITH DIMENSIONALITY FOR ATOMIZING BOILER DIAGNOSTIC BASE

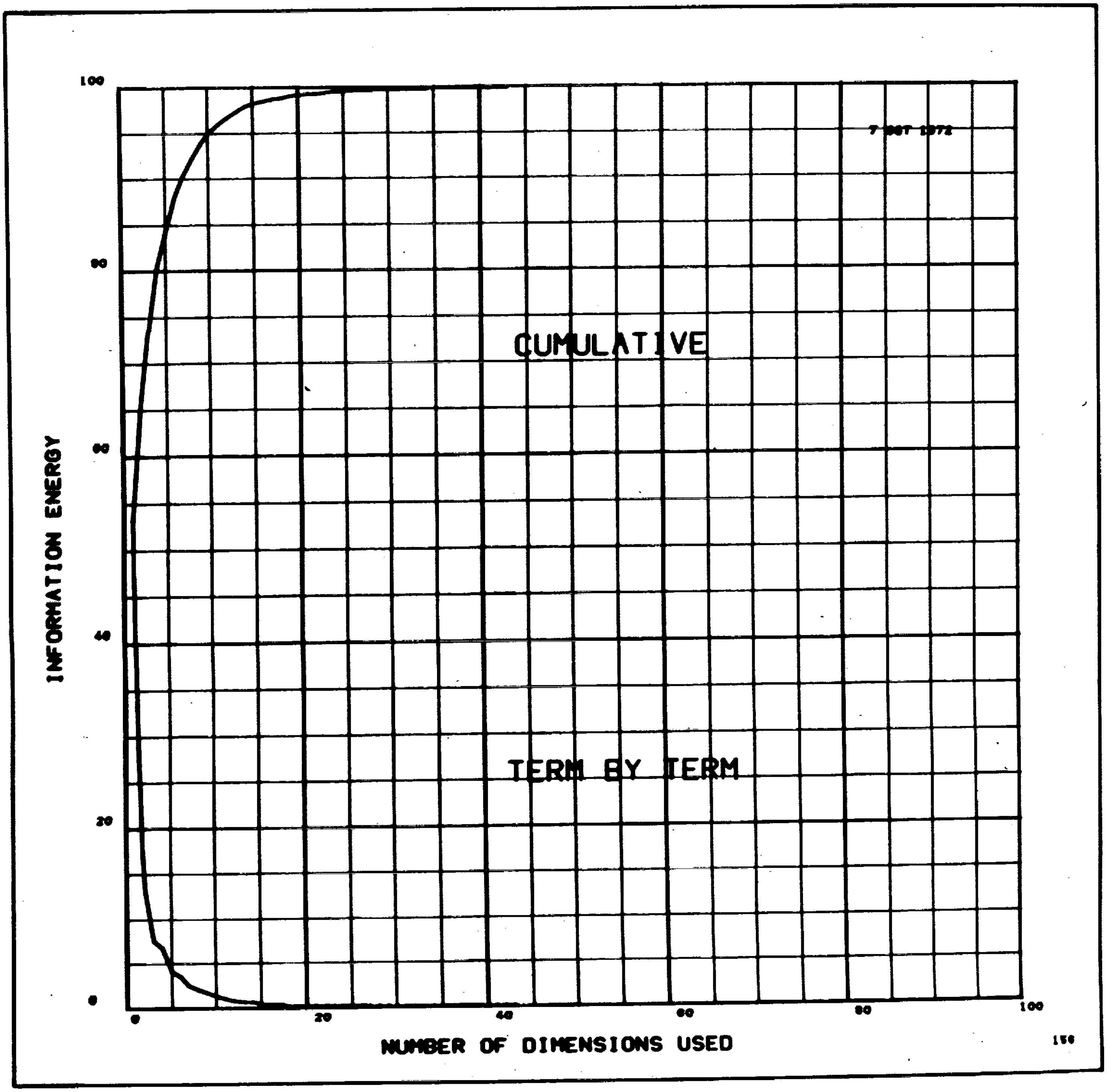


FIGURE 6.6

SCATTER PLOT OF FIRST AND SECOND COEFFICIENTS OF GENERALIZED FOURIER SERIES REPRESENTATION FOR LEARNING CASES USED IN DEVELOPING ALGORITHM FOR DIAGNOSING ATOMIZING STEAM BOILER FAILURES

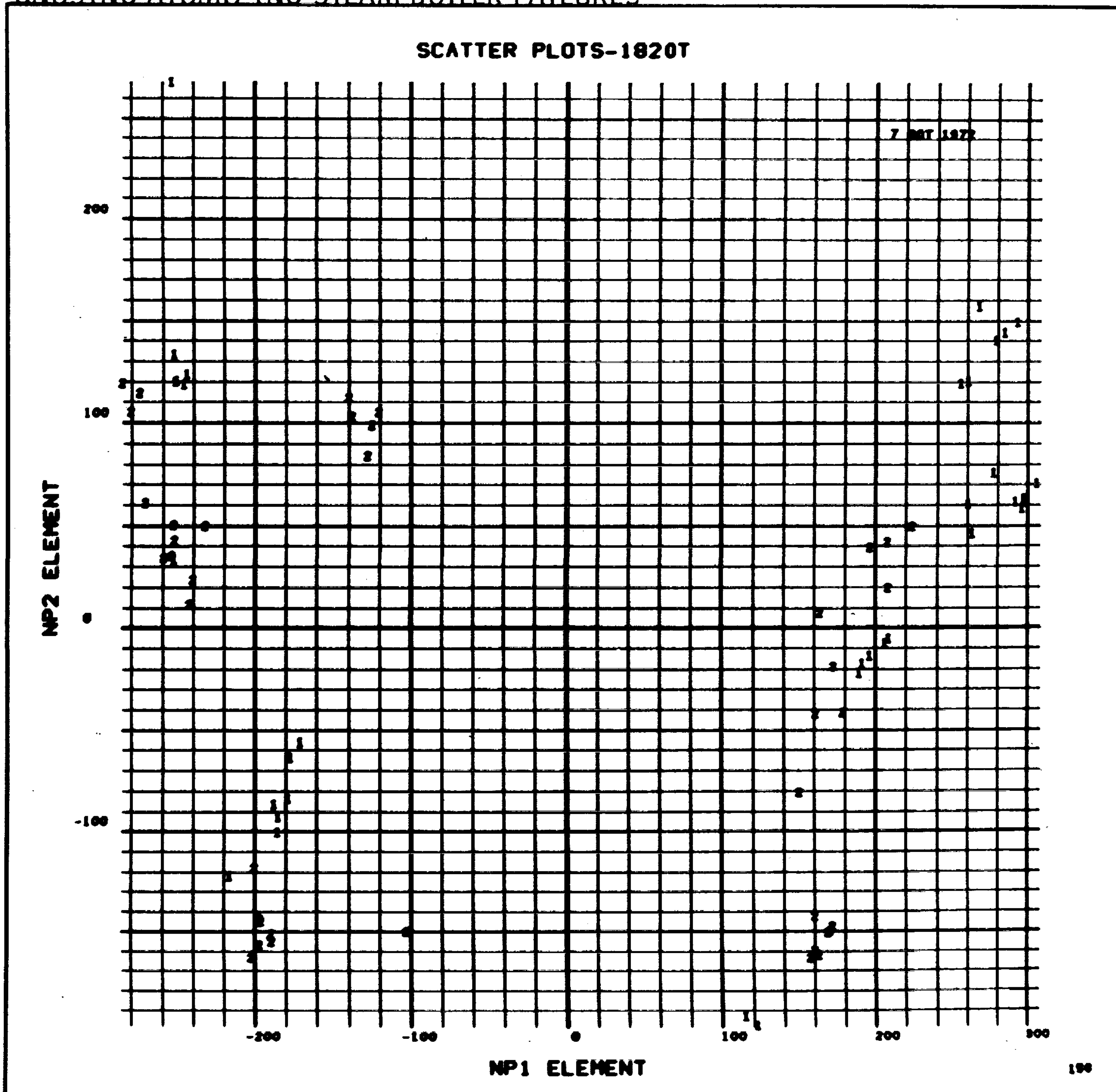


FIGURE 6.7
FIRST OPTIMUM FUNCTION FOR ATOMIZING STEAM BOILER DIAGNOSTIC BASE

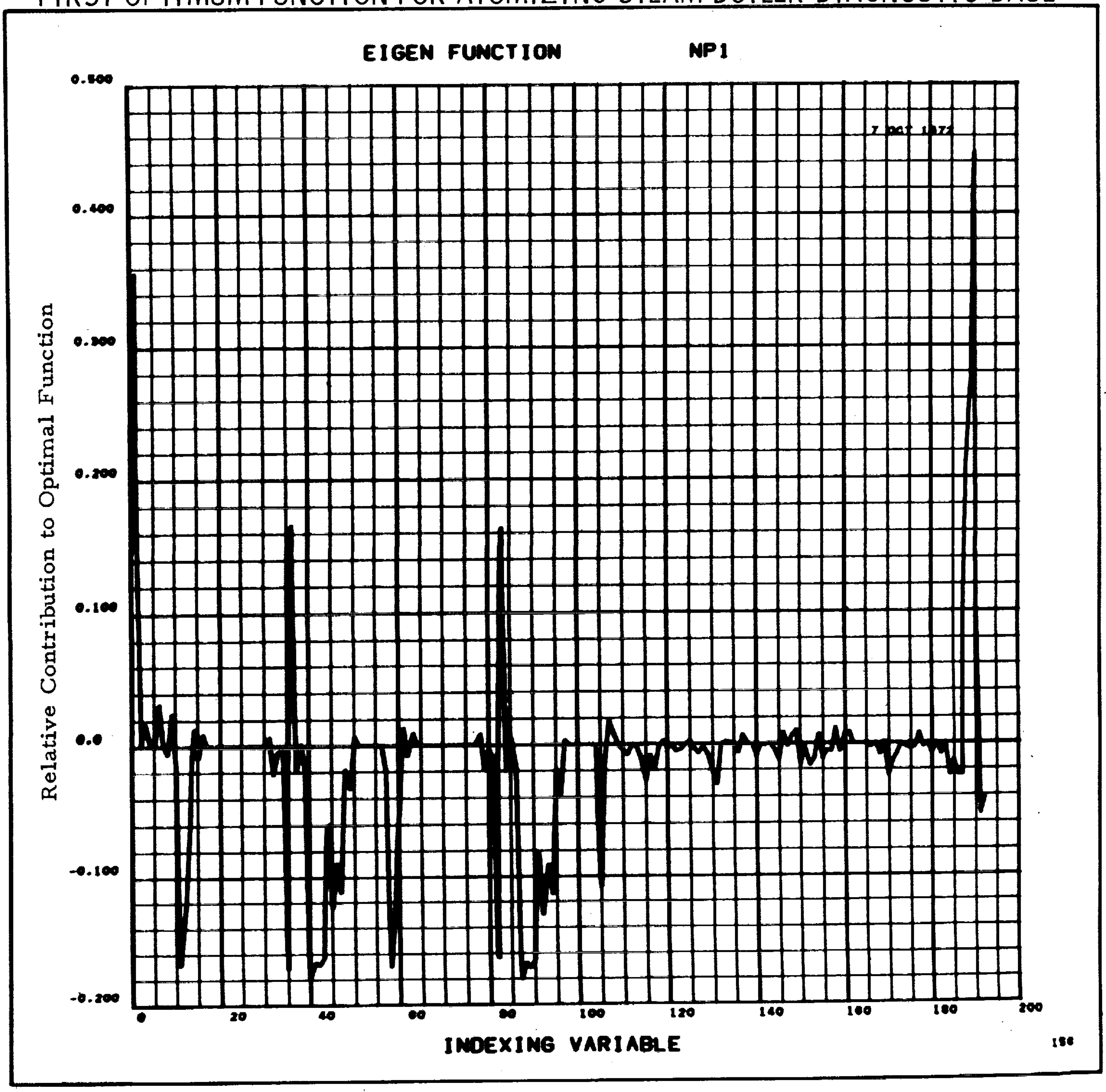


FIGURE 6.8

SECOND OPTIMUM FUNCTION FOR ATOMIZING STEAM BOILER DIAGNOSTIC BASE

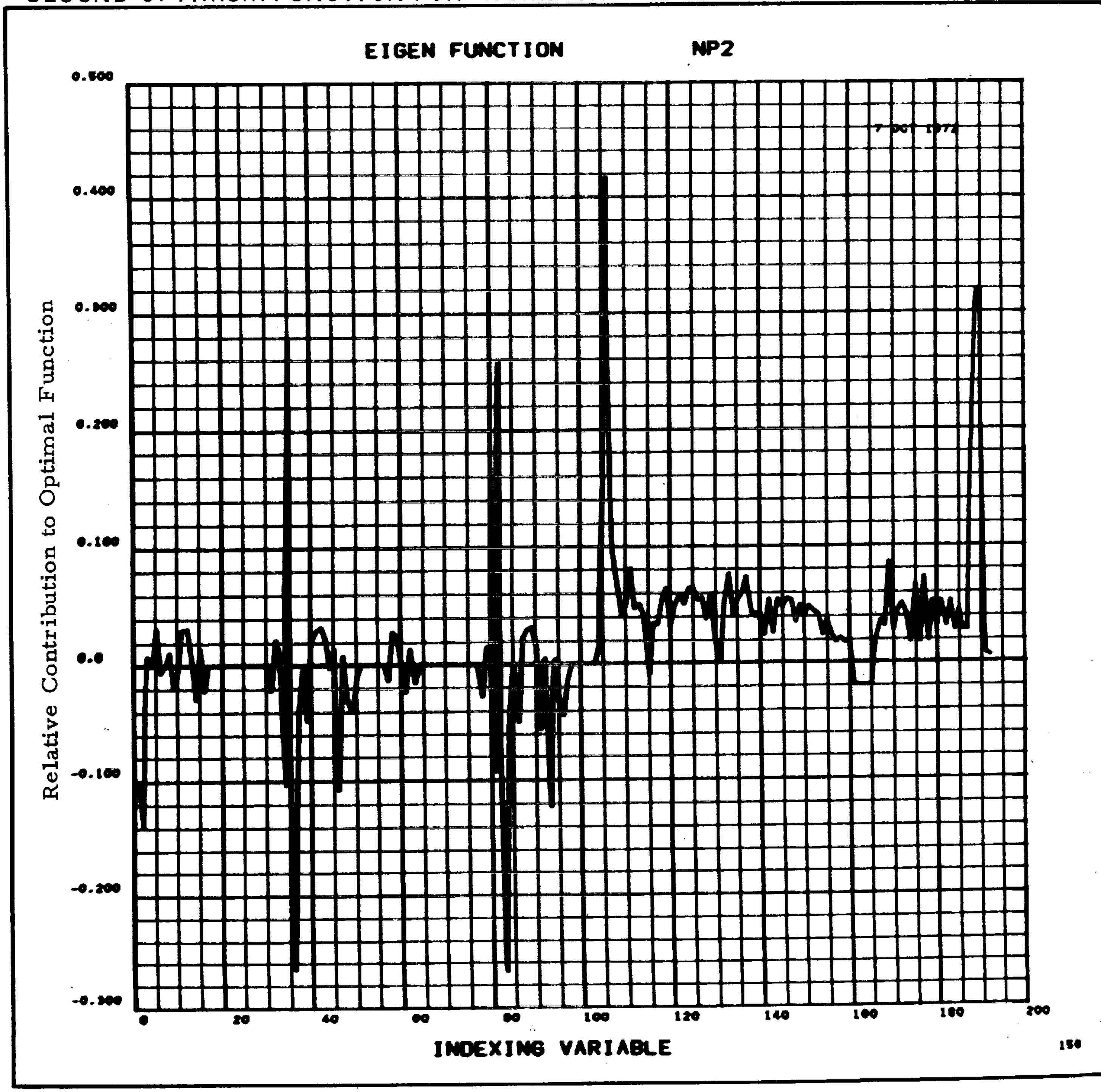


FIGURE 6.9 - RELATIVE IMPORTANCE OF MEASUREMENTS FOR DIAGNOSING FAILURES IN BOILER NO. 1 USING 20 DIMENSIONS

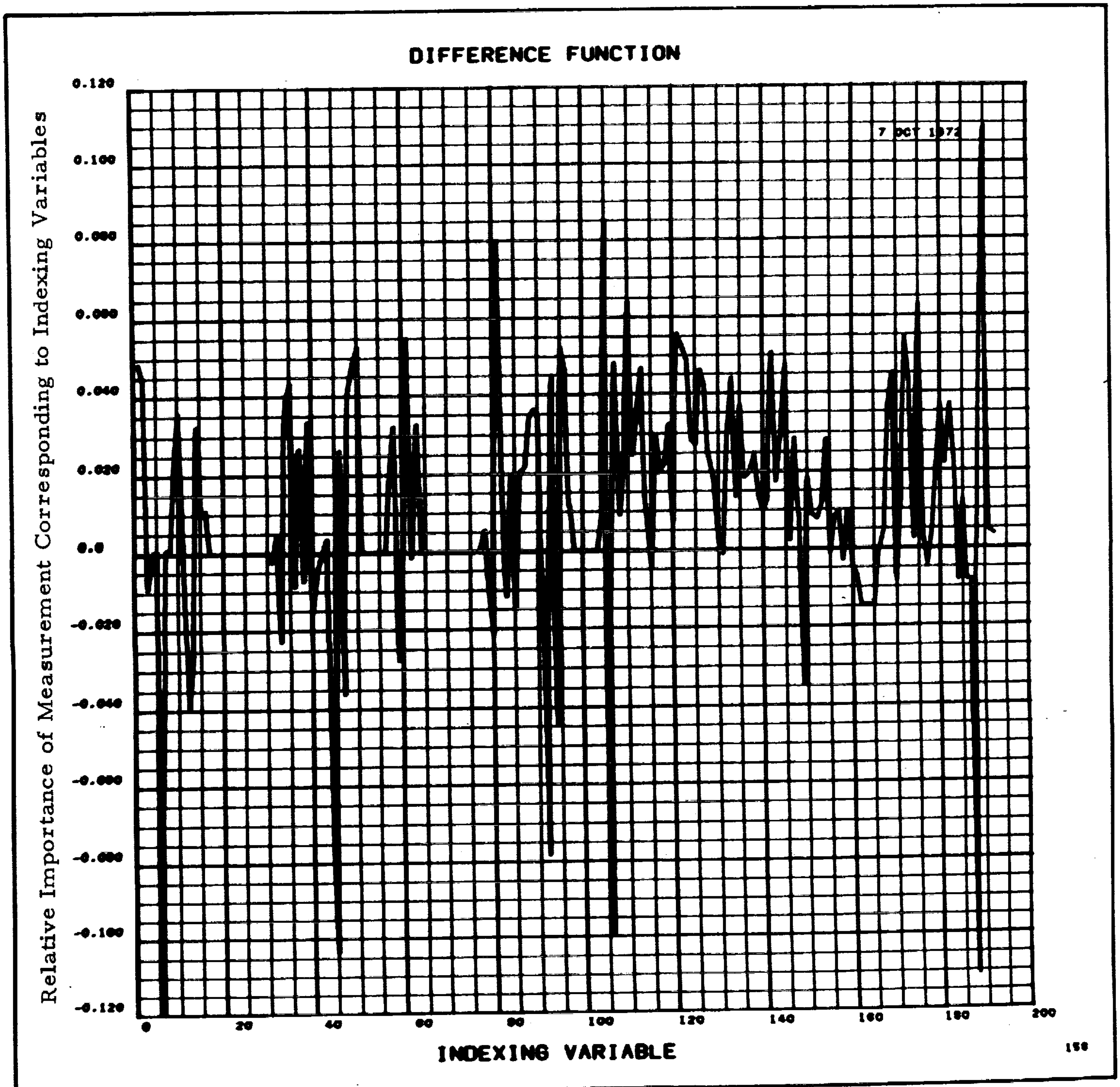
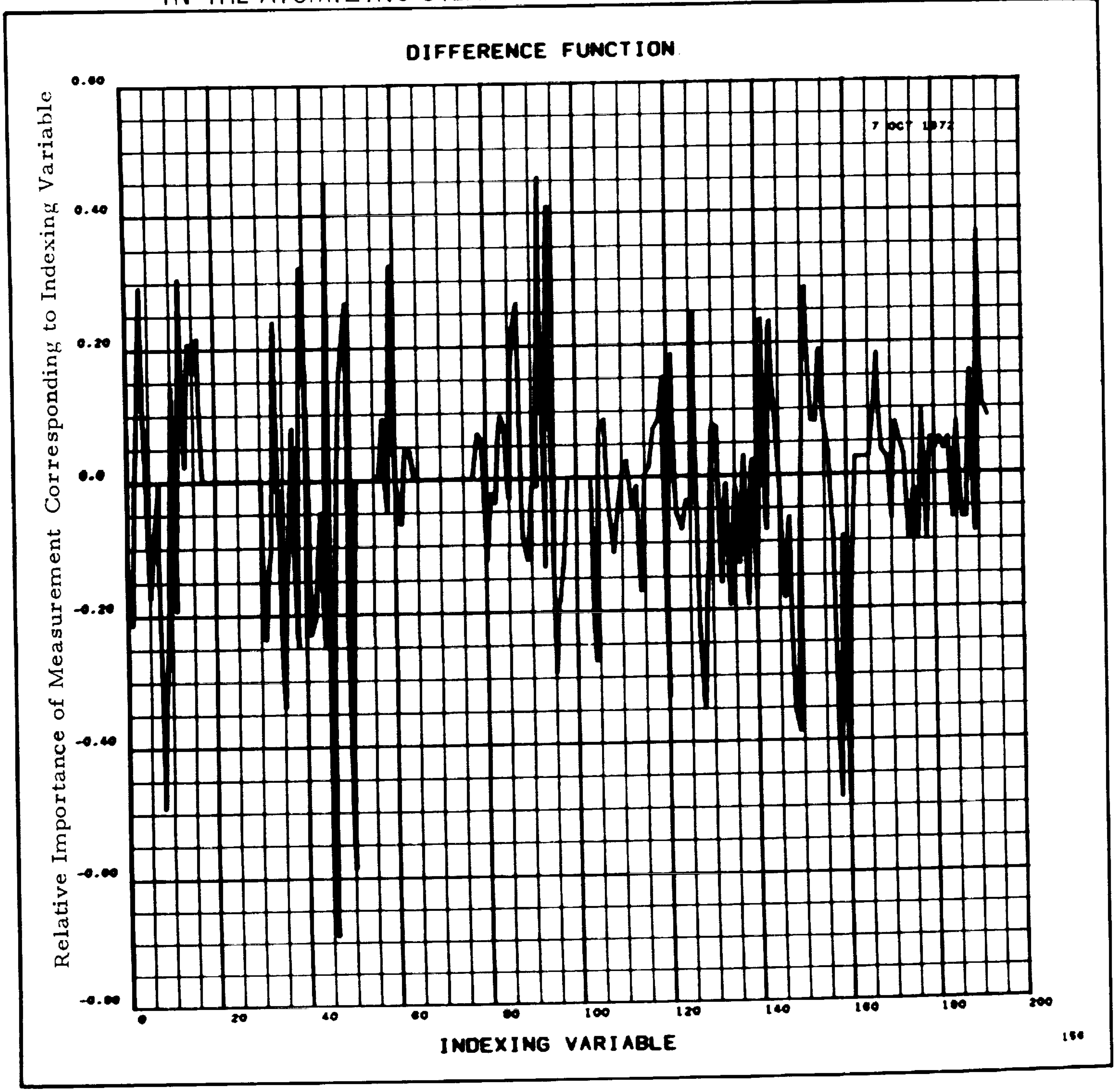
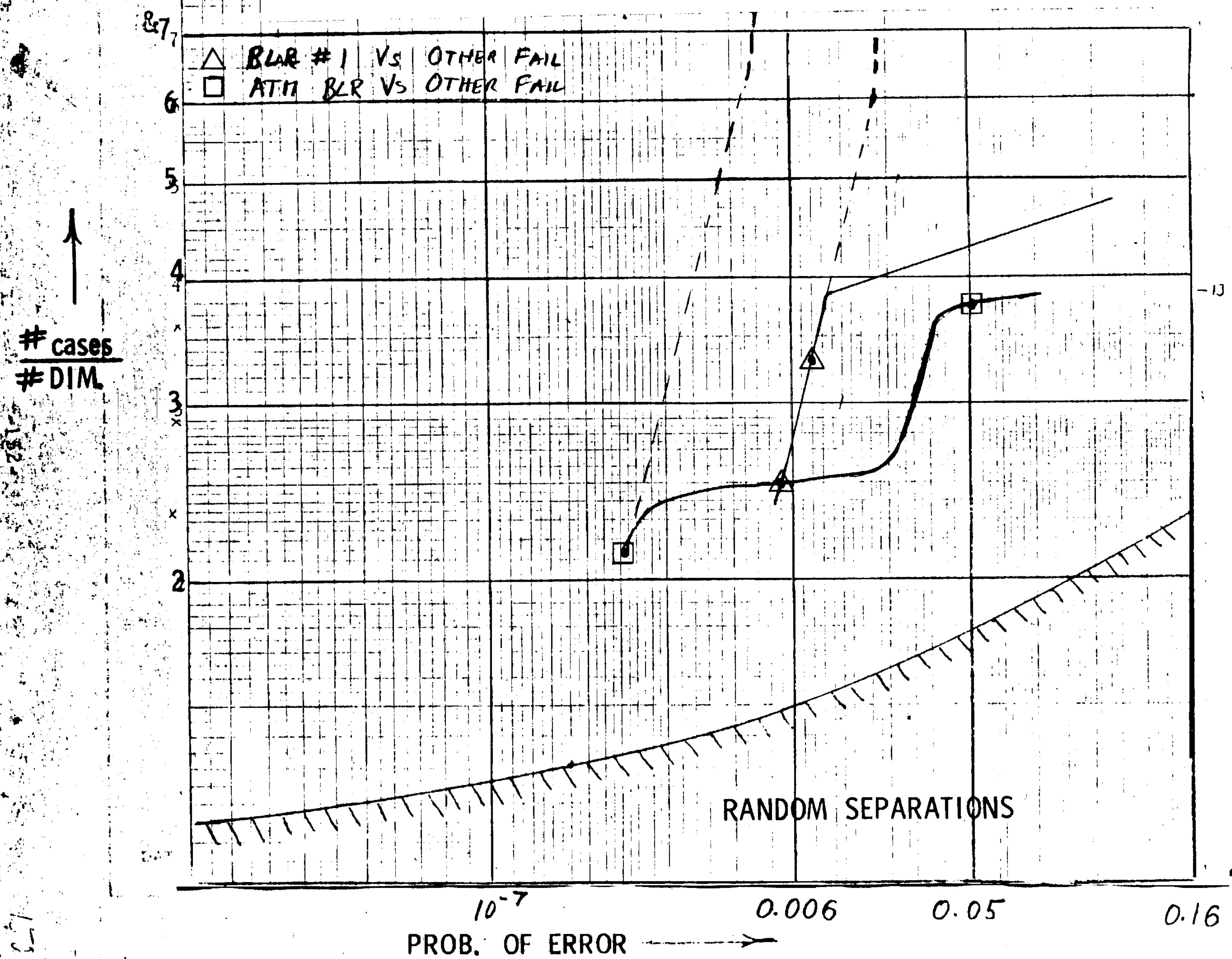


FIGURE 6.10 - RELATIVE IMPORTANCE OF MEASUREMENTS FOR DIAGNOSING FAILURES IN THE ATOMIZING STEAM BOILER USING 35 DIMENSIONS



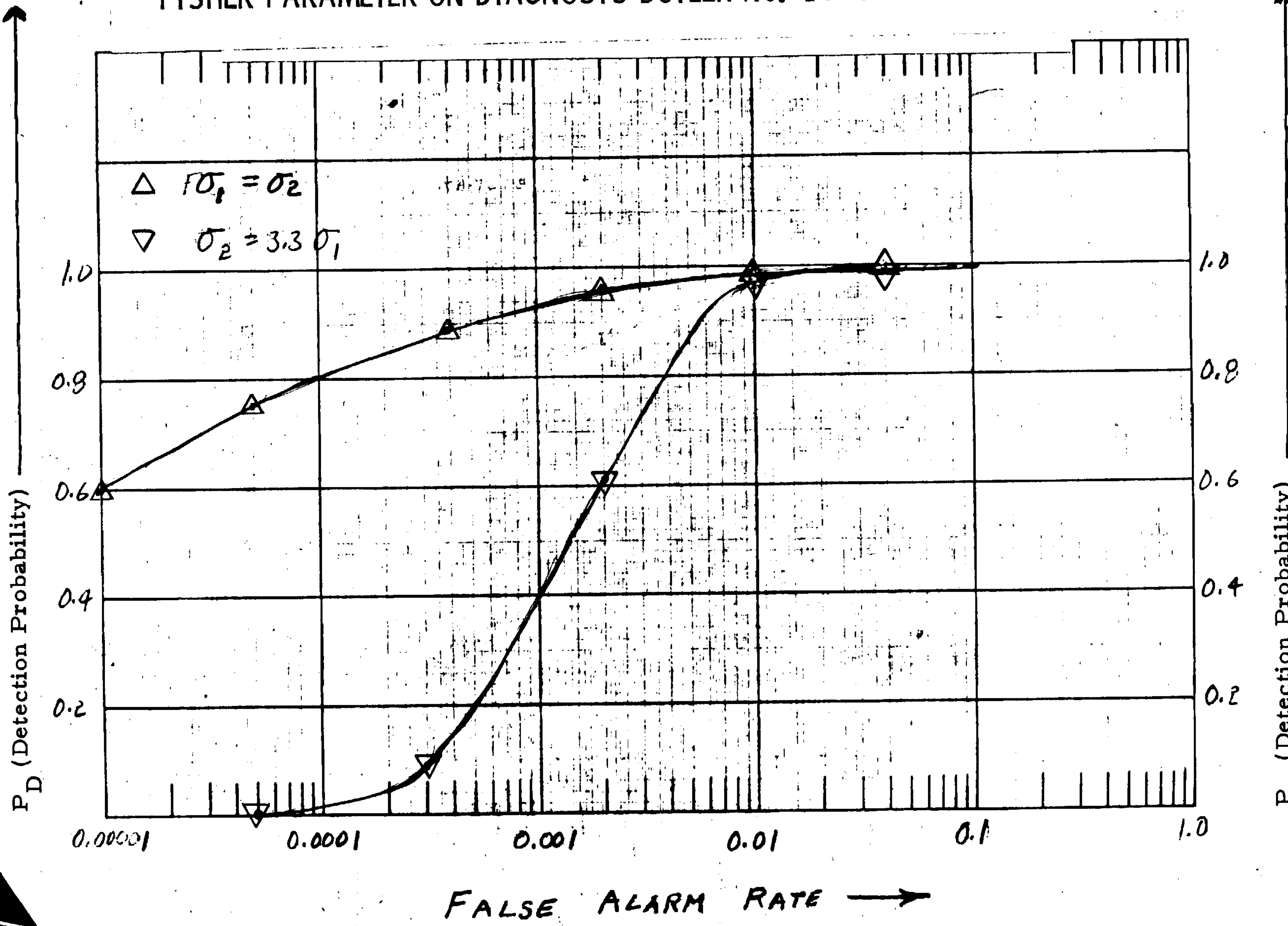
1 - 1690

FIGURE 6.11 - PERFORMANCE MAP FOR DIAGNOSTIC ALGORITHMS



359-91 5 CYCLES X 70 DIVISIONS

FIGURE 6.12
CLASSIFICATION PERFORMANCE TRADE-OFF CURVES SHOWING EFFECT OF FISHER PARAMETER ON DIAGNOSIS BOILER NO. 1 FAILURES



NOTE: USE TYPE B PENCIL FOR VUGRAPHS AND REPORT DATA.

7.0 TIME TO FAILURE ALGORITHM

When both the detection of an incipient failure in the system and the diagnosis of the failure has been completed, the remaining question is when will the failure occur. If this information is also available, the demand maintenance system cannot only alert the maintenance people that the failure will occur, define where the failure will occur, but also can schedule the corrective action to cause the least inconvenience to both the maintenance personnel and the users. The development of algorithms to predict the time at which failure will occur is the most difficult of the three types of algorithms considered. The major reason for this difficulty is the fact that a great deal of data is required for any given failure mode before one has sufficient information to derive such an algorithm. The compensation for this disadvantage is that it is exactly this failure that is most likely to occur. One series of failures for which such a set of data is available is the failure of the atomizing steam boiler. For this reason we have selected the atmozing steam boiler as a case to demonstrate the feasibility of deriving time-to-failure algorithms for those failures occurring sufficiently often to provide an adequate data base.

A separate base was constructed to be used for the derivation of the time-to-failure algorithm. This base was constructed using the 29 cases describing the Kennedy Space Flight Center central heating plant prior to the failure of the atomizing steam boiler. These 29 cases were processed through the ADAPT programs. The effect of dimensionality on this representation of these 29 cases is shown in Figure 7.1. The representation is nearly complete with only 15 terms. The first two of the optimum functions derived, using the time to failure data base are shown in Figures 7.2 and 7.3. When compared with Figure 5.37 and Figure 5.38, the most striking difference is the disappearance of the steam pressure for atomizing boilers A and B from these two optimum functions. This implies that for the atomizing steam boiler failures there was far less variation in the steam pressure associated with the atomizing boilers. If these two variables and their 12-hour average are deleted from Figure 5.37, the remainder of the variation is remarkably similar to that shown by Figure 7.2. The same is true for the second optimum function given in Figure 7.3.

Figure 7.4 presents a scatter plot representation of the atomizing steam boiler cases. This scatter plot shows four individual groups representing the four specific failures for which time-to-failure information was available. Examination of the distribution of these failures on the scatter plot represented in Figure 7.4 in comparison with the scatter plot for the universal detection base presented in Figure 5.36 shows that the failures of the atomizing steam boiler represent a reasonable cross section of the variation displayed by the entire data base.

Time-to-failure algorithms were developed using twelve and nine dimensions. The relative importance spectrum for these algorithms is shown in Figure 7.5. This figure shows that the most important dimension for the time-to-failure algorithm is dimension No. 9. It also shows that there is significant information in the eighth, tenth, and eleventh optimum functions.

Figure 7.6 shows this ninth optimum function. The most important variables to this function are the rainfall during the past hour, the change in the boiler No. 1 flow, the number of gallons of fuel used, the change in flow for Zone 2 makeup feed water meter and the number of gallons of fuel used by the atomizing steam boiler. Thus in part, this function is made up of a contrast between the amount of energy used by the boiler, the heating load being carried by the boiler and the amount of fuel being supplied to the atomizing steam boiler. This is a reasonable collection of information to make a significant contribution to the time-to-failure. The track of this algorithm on the performance map is shown in Figure 7.8. The two algorithms were presented by the circles on the solid line which shows the effect of dimensionality on the performance of this algorithm. The curvature of this line is based upon the relative importance spectrum shown in Figure 7.5. That is, very little information is lost as one decreases from twelve to eleven dimensions, a significant and approximately equal amount is lost in decreasing from eleven to ten and ten to nine measurements. Once one decreases below nine measurements, a very large amount of the information required predicting the time-to-failure is lost.

Figure 7.8 presents a comparison of the estimated time-to-failure with the actual time-to-failure for the 12-dimensional algorithm. The abscissa on this plot is the actual number of hours prior to the occurrence of the failure for that particular case and the ordinant is the time-to-failure as estimated using the time-to-failure algorithm presented in Table 2.4. Since this algorithm only has a ratio of number of cases to number of dimensions of about 2-1/2, one might expect this performance to be degraded somewhat. However, the performance indicated for this algorithm is a one-sigma accuracy of six hours. This accuracy is illustrated by the interrupted line of Figure 7.8. The accuracy which might be expected of a final algorithm should certainly be greater than that achieved by the 9-dimensional algorithm which was a one-sigma error of 9 hours. Thus, one can be quite confident that the kind of failure for the atomizing steam boiler can be predicted to within 6 to 9 hours and probably closer to 6 hours of the time-to-failure 72 hours in advance of the failure. Figures 7.8 and 7.9 present the relative importance vectors for the 9 and 12 dimensional time-to-failure algorithms respectively. These figures show the importance of each of the measurement to predicting the number of hours before failure will occur. As would be expected, many of the variables important to optimum function No. 9 are also important in the relative importance vector. Furthermore, the relative importance vectors for the 9 and 12-dimensional case are quite similar. This is encouraging in that it leads one to believe that the mechanisms upon which the algorithms are based are generally similar. Clearly, the 12 dimensional algorithms illustrated in Figure 7.9 must include some elements pertinent to the prediction of time-to-failure which are not included in the 9-dimensional algorithm. Examination of these two figures shows that in general the higher dimensional algorithm does not rely as heavily on the rain flow, the 12-hour average temperature or the flows in Zone 2.

Although the successful development of this time-to-failure algorithm for the atomizing steam boiler is not sufficient to insure that time-to-failure algorithm can be derived for all failures which are identified, it does show the time-to-failure algorithms are feasible and can be derived for at least some of the failure modes. The system utilized to implement these maintenance algorithms must be designed to operate both with and without this time-to-failure information.

FIGURE 7.1 - VARIATION OF INFORMATION RETAINED WITHIN DIMENSIONALITY FOR TIME TO FAILURE BASE

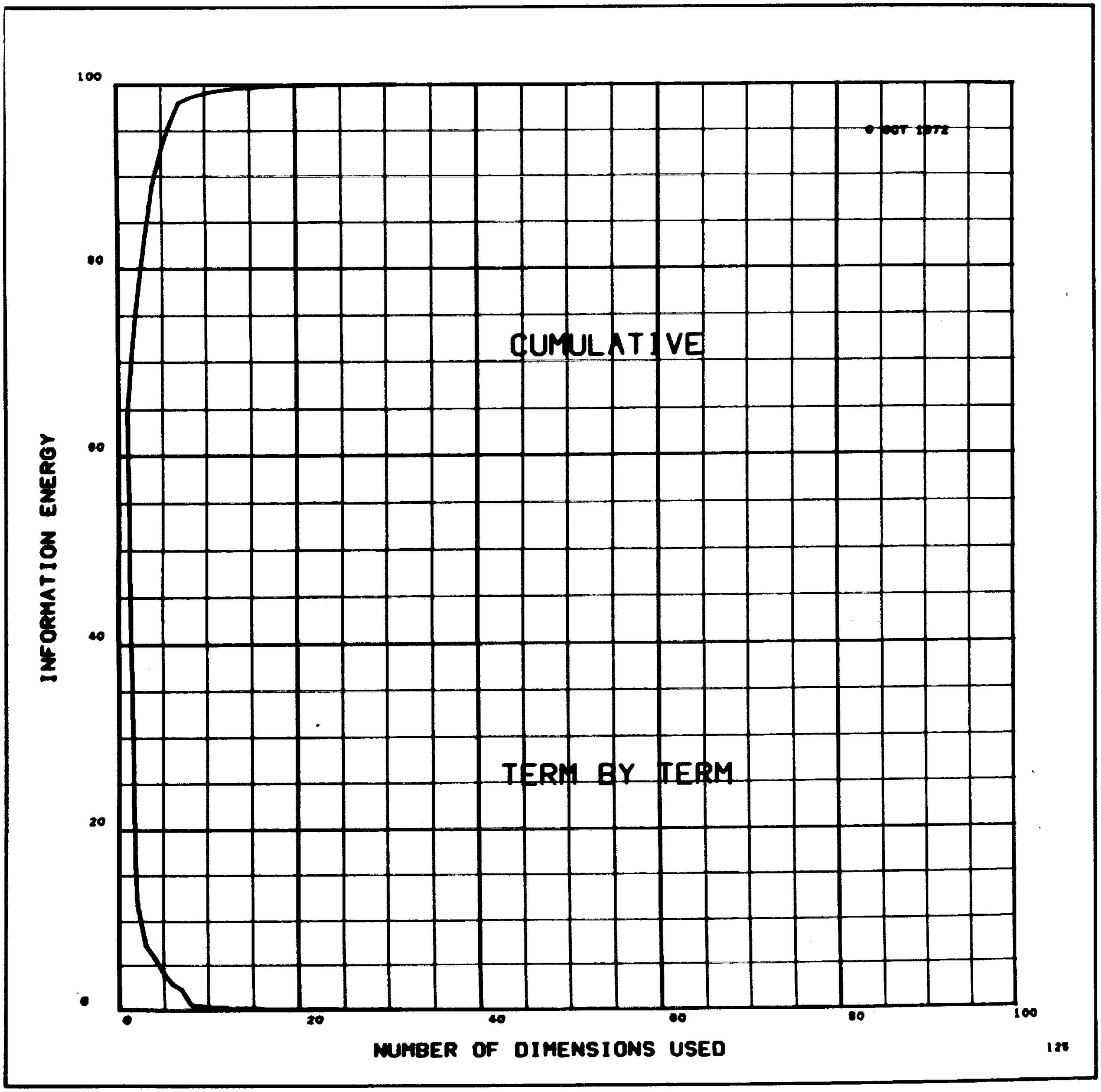


FIGURE 7.2 - FIRST OPTIMUM FUNCTION TIME TO FAILURE BASE

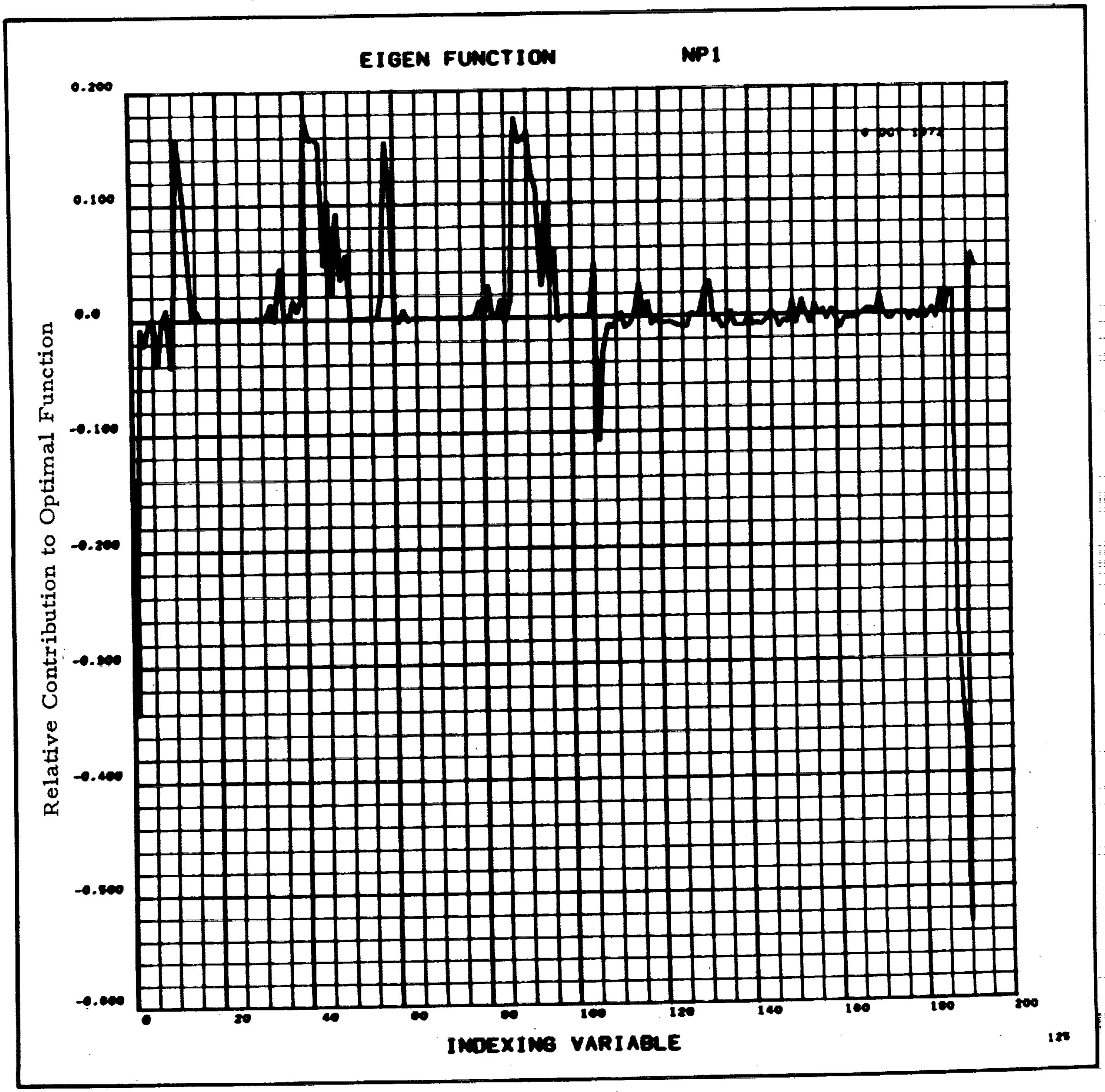


FIGURE 7.3 - SECOND OPTIMUM FUNCTION TIME TO FAILURE BASE

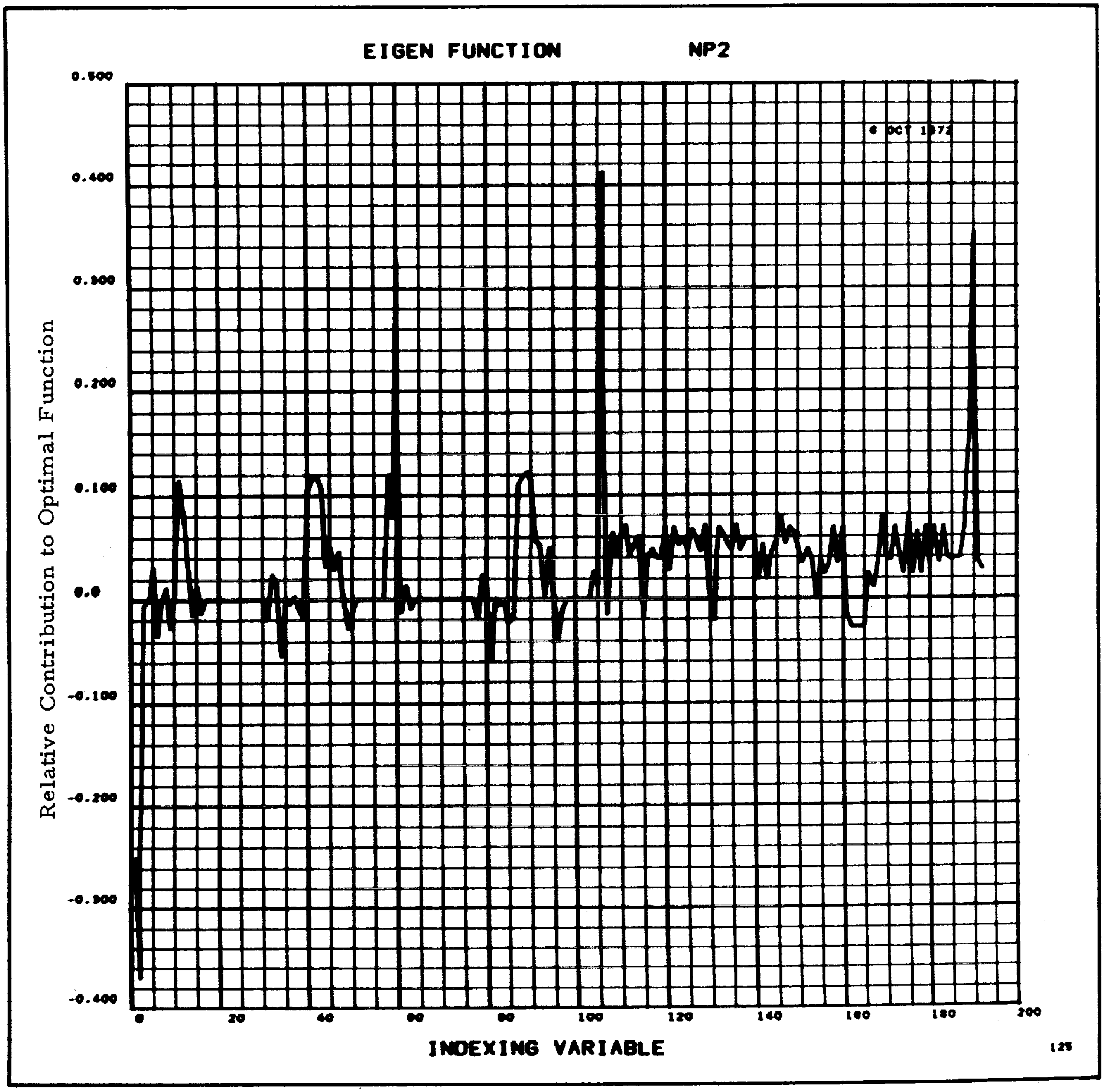


FIGURE 7.4 - SCATTER PLOT OF FIRST TWO COEFFICIENTS OF GENERALIZED FOURIER SERIES REPRESENTATION OF CASES USED TO DEVELOP TIME TO FAILURE ALGORITHM

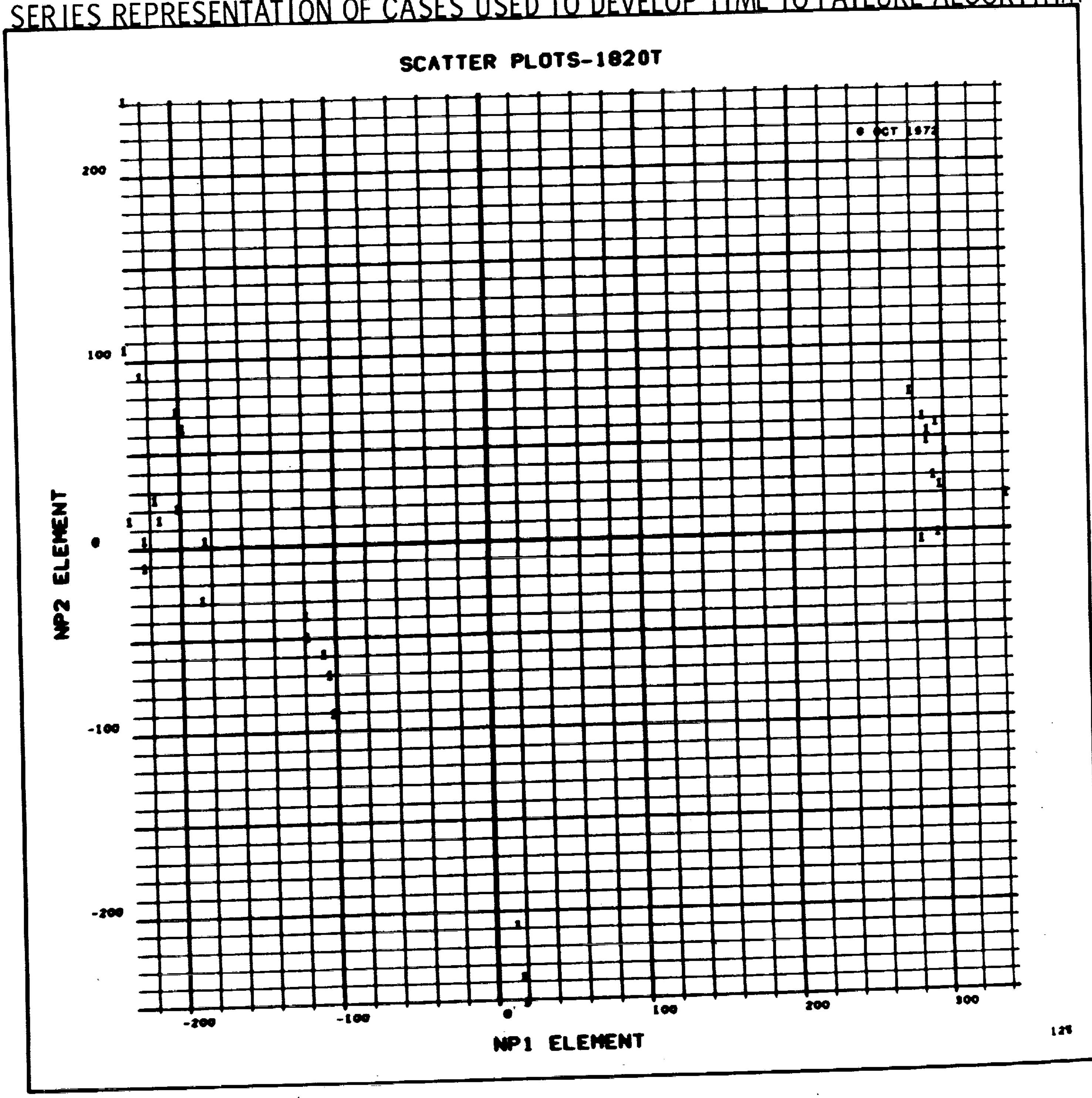


FIGURE 7.5

RELATIVE IMPORTANCE OF OPTIMUM DIRECTIONS TO PREDICTING TIME TO FAILURE

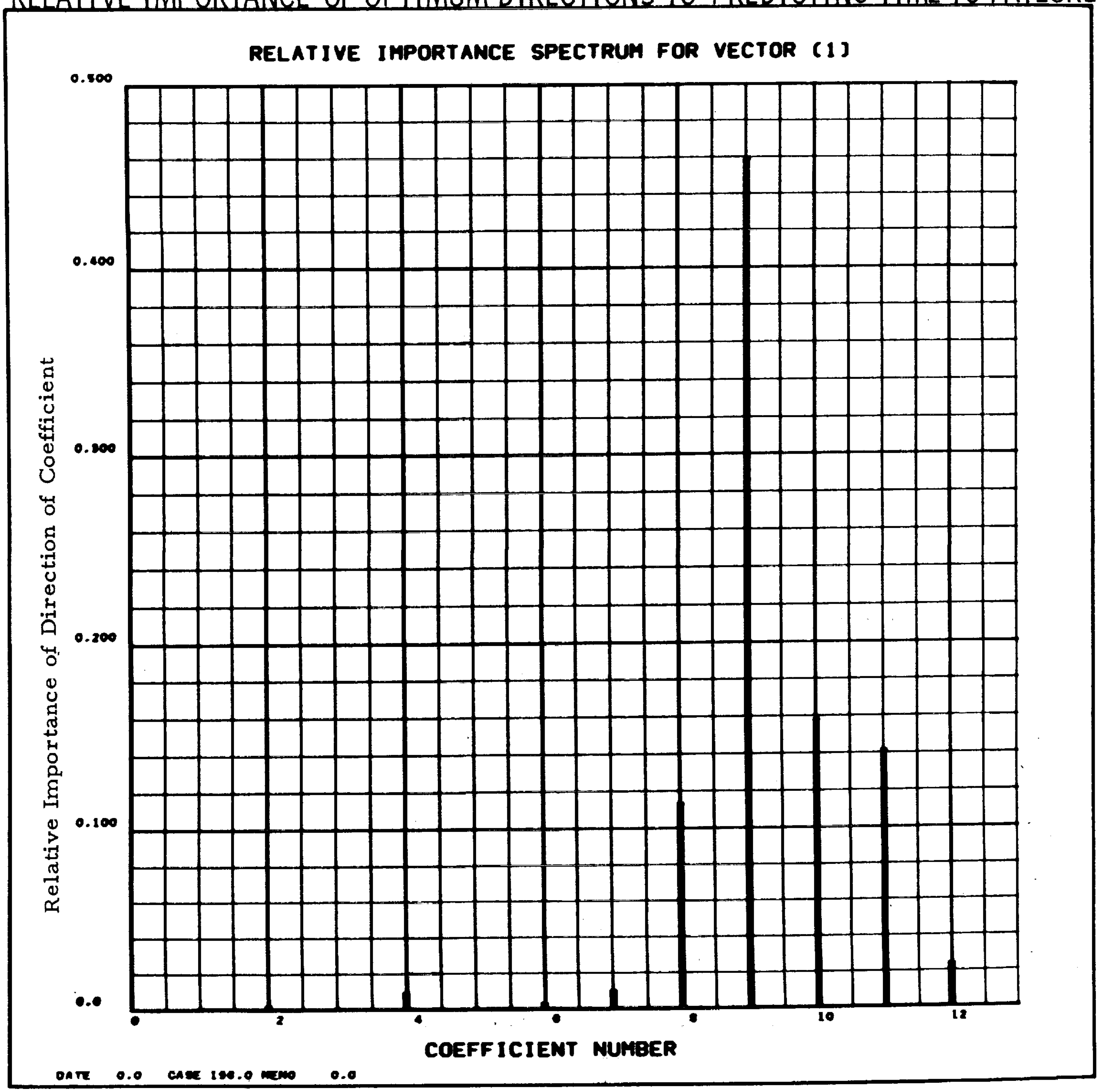


FIGURE 7.6 - NINTH OPTIMUM FUNCTION TIME TO FAILURE BASE

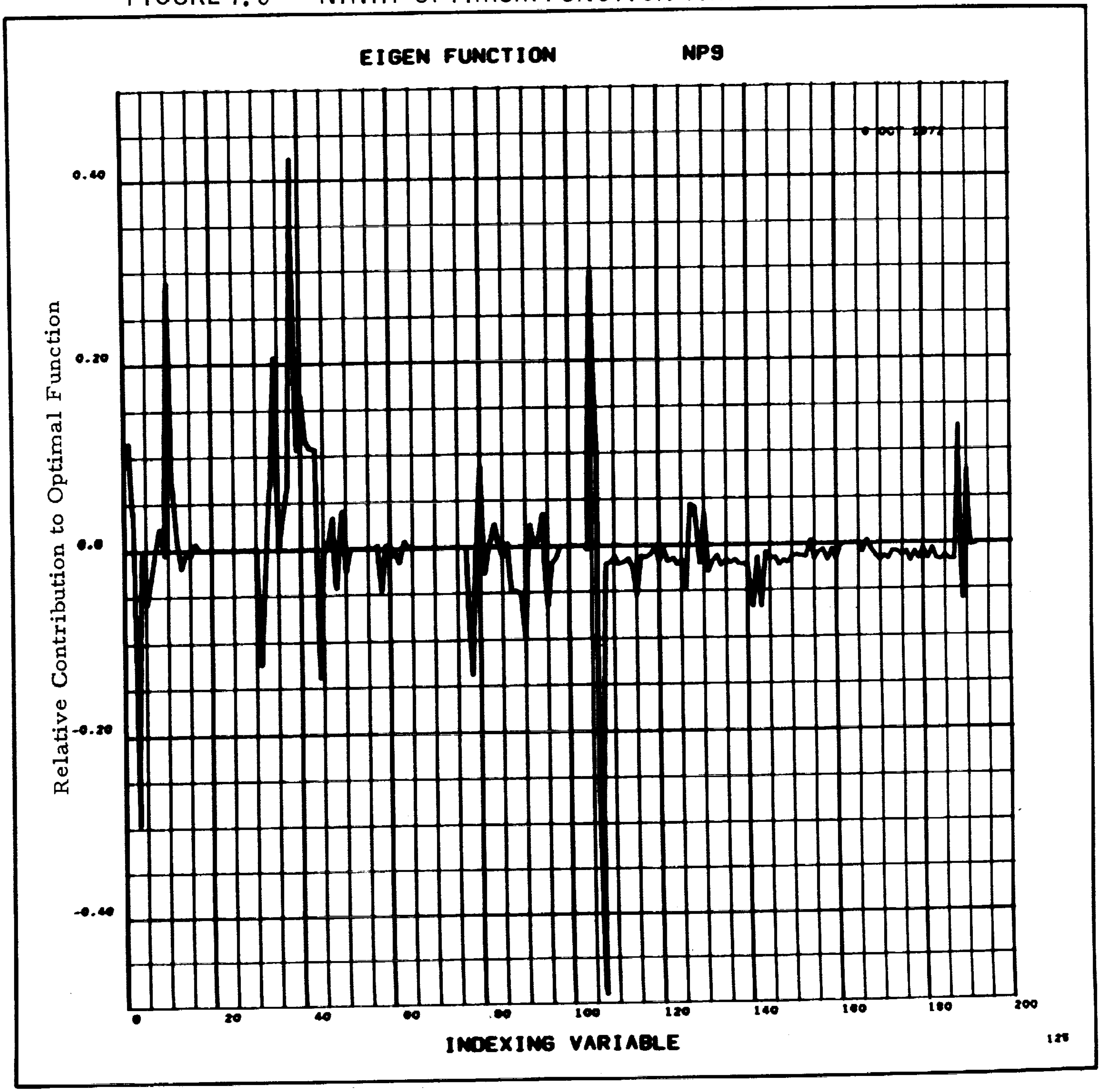
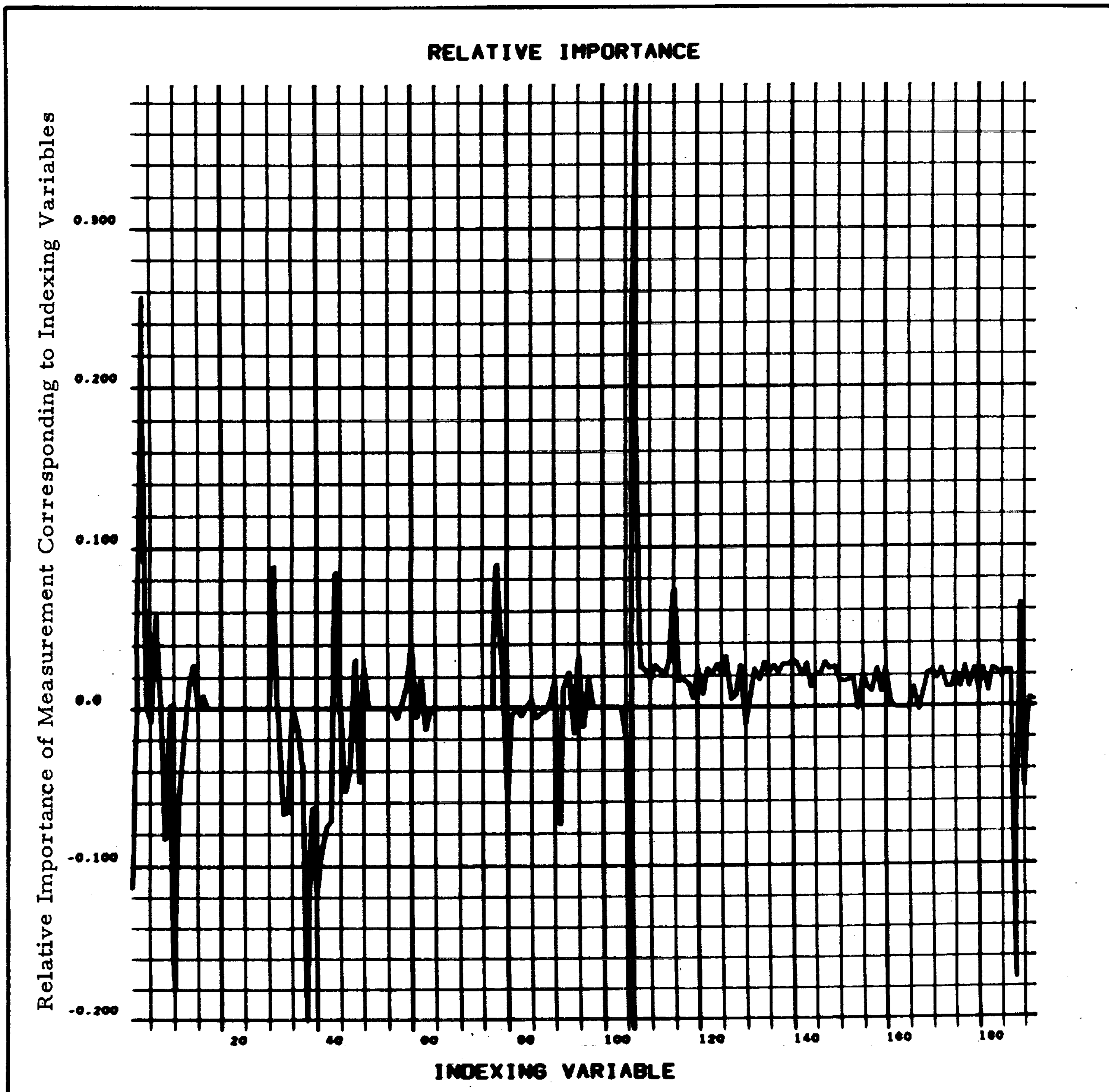
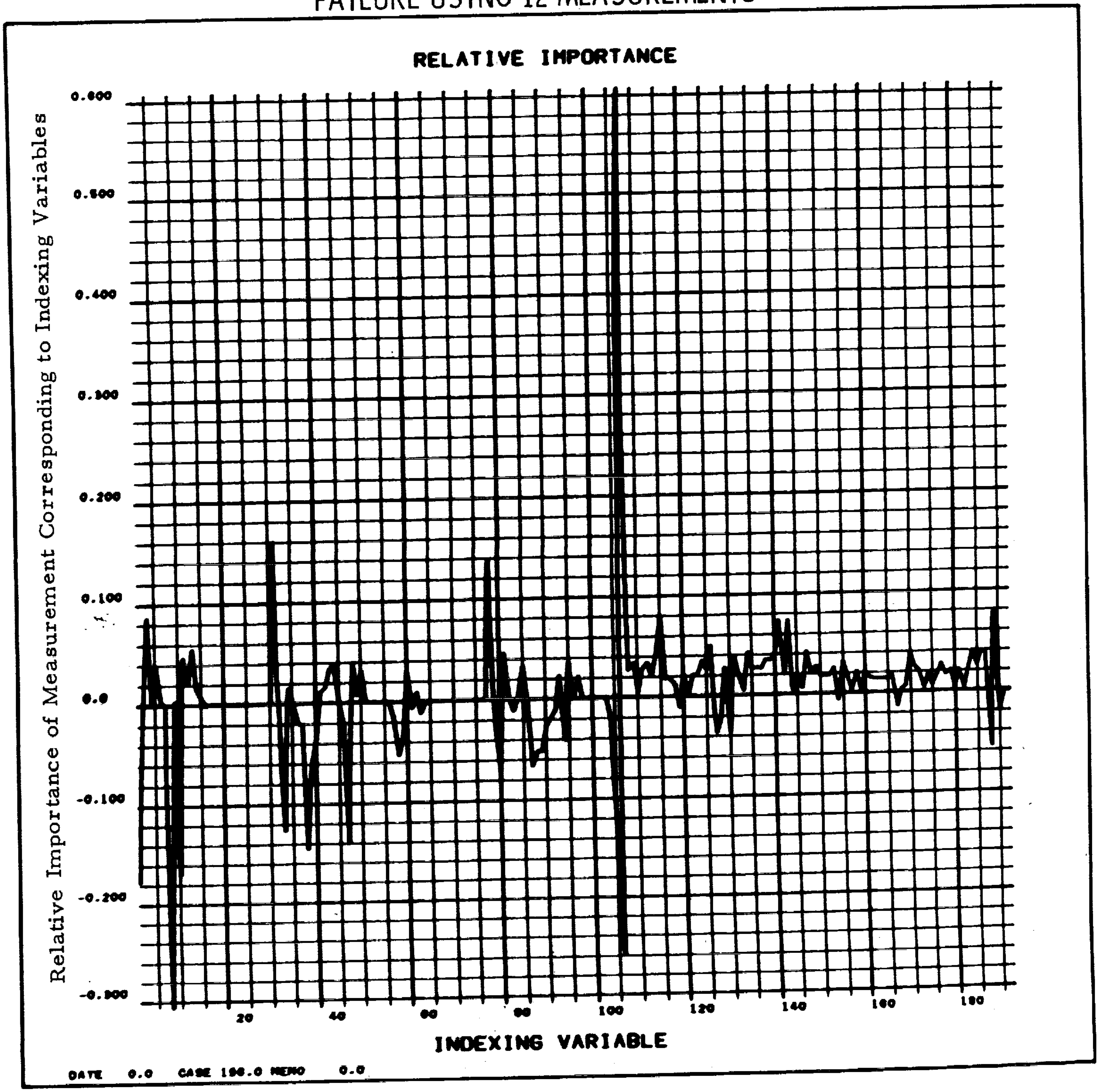


FIGURE 7.7 - RELATIVE IMPORTANCE OF MEASUREMENTS FOR PREDICTING TIME TO FAILURE USING 8 DIMENSIONS



NOR-9 t_f = 8/12

FIGURE 7.8 - RELATIVE IMPORTANCE OF MEASUREMENTS TO PREDICTING TIME TO FAILURE USING 12 MEASUREMENTS



References

1. Hunter, H. E., N. Kemp, "Application of Avco Data Analysis and Prediction Techniques (ADAPT) to Prediction of Cyclone Central Pressure and its Derivatives using NIMBUS HRIR Data", AVSD-0362-70-RR, August 1970.

2. Hunter, H. E., N. Kemp, "Application of Avco Data Analysis and Prediction Techniques (ADAPT) to Prediction of Cyclone Central Pressure and Its Derivatives Using NIMBUS HRIR Data", AVSD-0142-71-RR, March 1971.

3. Hunter, H. E., N. Kemp, "Application of Avco Data Analysis and Prediction Techniques (ADAPT) to Prediction of Cyclone Present Motion and 12 Hour Motion, and Re-centering Effects, Using NIMBUS HRIR Data", AVSD-0334-71-RR, July 1971.

4. Kemp, N. H, Hunter, H. E., Amato, R. A, "Application of Avco Data Analysis and Prediction Techniques (ADAPT) to Multi-Spectral Extra Tropical Cycle Accuracy Investigation", AVSD-0128-72-CR, March 1972.

5. Hunter, H. E., Kemp, N. H., "ADAPT Hurricane Data Selection and Performance Study", AVSD-0400-72-RR, Nov. 1972.

6. Hunter, H. E., J. Conway, "Demonstration of Feasibility of Using the Avco Data Analysis and Prediction Techniques (ADAPT) to Develop Algorithms for Automating the Identification of Solar Burst", AVSD-0255-71-RR, 21 May '71.

7. Hunter, H. E., Meixzell, L. M., Amato, R. A., "Final Letter Report ADAPT Solar Burst Compacting Study", AVSD-0209-72-CR, May 1972.

8. Hunter, H. E., Amato, R. A., "Application of Avco Data Analysis and Prediction Techniques(ADAPT) to Prediction of Sunspot Activity", AVSD-0287-72-CR, August 1972.

9. Kemp, N. H, Hunter, H. E., Amato, R. A., "Application of Avco Data Analysis and Prediction Techniques (ADAPT) to a Gauss-in-Gauss Detection Study", AVSD-0260-72-RR, July 1972.

10. Hunter, H. E., R. Amato, J. Conway, N. Kemp, "Demonstration of Applicability of Avco Data Analysis Technique to Sonar Signature, AVSD-0605-70-RR, December 1970.

11. Hunter, H. E., N. Kemp, "Demonstration of Feasibility of Avco Data Analysis and Prediction Techniques (ADAPT) for Sonar Detection" AVSD-0411-710RR, September 1971.

12. Hunter, H. E., L. Meixsell and J. Conway, "Feasibility Demonstration for Optically Implementing an Avco Data Analysis and Prediction Techniques (ADAPT) Algorithm for Recognizing Spirals", AVSD-0026-72-RR, Jan. 1972.

13. Hunter, H. E., Conway, J, "Use of the Avco Data Analysis and Prediction Techniques (ADAPT) to Develop Analytical Techniques for a Comprehensive Attack on Auto Theft and Burglary in Lawrence, Mass.", AVSD-0042-72-RR, 31 January 1972.

14. Hunter, H. E., "Discussion of Patterns Recognition Techniques Applied to Diagnosis", presented at the Society of Automotive Engineers, Mid-year meeting, May 18-20, 1970.

15. Jones, T. O., Grimes, D. M., Dork, R. A., "A Critical Review of Radar as a Predictive Crash Sensors" presented at the Second International Conference on Passive Restraints, Detroit, Michigan, May 22 - 25, 1972, SAE Report 720424, pages 22-24, 38.

16. Watanabe, S. "Karhunen-Loeve Expansion and Factor Analysis Theoretical Remarks and Predictions", Transaction of the 4th Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes, 1965, pages 635 thru 660.

17. Anderson, T. W., "Introduction to Multivariate Statistical Analysis, Chapter 11, New York, John Wiley & Sons, 1958.

APPENDIX A

FEATURES OF ADAPT ANALYSIS

The unique aspect of the ADAPT approach to empirical data analysis is preceding the analysis with the derivation of the optimal representation for the particular data set. The ADAPT programs provide a unique capability for determining this optimum representation for large data sets. However, regardless of the size of the data set, the availability of this optimum representation provides many significant benefits to any further empirical analysis. These benefits include: 1) definition of which variables dominate the variation, 2) ordering of the data by its general usefulness for extracting information, 3) reduction in the computation required to perform further analysis, 4) reduction in the amount of learning data required to perform any given analysis, 5) an improved ability to establish performance and validity criteria, and 6) the ability to perform special functions such as clutter subtraction and extrapolation.

The availability of the optimum functions for representing any given data set is analogous to having the governing differential equations for a classical physics problem. These optimum functions provide information regarding the nature of the physics which govern the phenomena associated with this data. In particular, these functions will define exactly where the greatest and most highly correlated variation from case to case occurs. This information can be extremely useful in selecting data to be used for the analysis and in understanding the mechanism governing the phenomena which produced this data.

In addition to simply having the optimum functions for representing the data, these functions are ordered such that each function explains successively less variation in the data. This provides the user with a capability to reject variables in an intelligent rather than a random manner, if the resources or available learning data require the use of fewer dimensions than would naturally be used to describe the data. This ordering allows one to throw away those variables which explain the smallest amount of variation and therefore in general should be least useful to any analysis. Although it might be more desirable to be selective based on the particular analysis to be performed, this is not usually possible until after the analysis has been performed, when it is obviously no longer useful. Thus, it is almost axiomatic that the apirori rejection of data for a particular analysis cannot be based on that particular analysis, so the rejection based on explained variation is an attractive approach to eliminating data when realities of the resources or available learning cases makes such an elimination necessary.

Regardless of any prior decision to reduce the dimensionality, the ADAPT approach to any real problem will automatically lead to a significant reduction

G

in dimensionality. When the information energy curves which are produced by the ADAPT programs are examined, it is almost always possible for the analyst to select some dimensionality after which it is inconceivable that further useful information is incorporated in the data. This criteria alone usually results in a reduction of dimensionality of more than an order of magnitude.

A reduced dimensionality obviously allows one to perform computations with smaller computer capabilities. Furthermore, the orthogonality of the optimum representation also provides simplifications in the computation. For example, in the optimal ADAPT space one can in some cases derive the Fisher discriminant without inverting the covariance matrix. This combination of reduction in quantity of computation required and simplification due to orthogonality also makes it feasible to update classification and regression algorithms in real time for cases where this might otherwise be impossible.

A more significant aspect of the lower dimensionality of the learning space follows from the requirement that the amount of learning data be large compared to the dimensionalty of the learning data. This requirement arises from the situation analogous to fitting a third order polynomial through a series of points. If the third order polynomial is to be fitted to three points, it will always fit perfectly and no physical relationship need be involved. However, if the third order polynomial is to fit a hundred points well then one knows that this third order polynomial must be related to the data in some physical manner. The same is true for empirical analysis in general. If the number of dimensions of the learning space is equal to the number of learning cases one can expect most empirical algorithms to provide a perfect fit to the learning data. However, this fit is normally based on differences between the population and the sample statistics and is not based on the physics of the problem. Experience has shown that the number of learning cases required to derive an empirical algorithm varies from 2 to 6 times the number of dimensions of the learning space. Thus, the usual ADAPT reduction of an order of magnitude or more in dimensionality of the learning space translates immediately into an equivalent reduction in the requirement for learning data. Since obtaining learning data is one of the most expensive aspects of empirical data analysis, this attribute of the ADAPT approach is often sufficient by itself to make the difference between feasibility and infeasibility of solving a given problem.

The ADAPT representation also provides an opportunity for establishing a necessary, although not sufficient, validity criteria. Validity criteria provide a method of determining whether a particular test case is from the same population as the learning data, and therefore determine the validity of applying the algorithms derives on the learning data to that particular test case. The ADAPT validity criteria consists of comparing the length of the test data vector in the original data space and in the ADAPT optimum space. If this transformation from the original data space to the optimum space results in a shortening of the test data vector by a factor considerably greater than the shortening which the learning data vectors suffered, one has an indication that the test data and learning data are from different populations. In addition to providing this validity criteria, the ADAPT programs have been designed to calculate performance criteria as part of the learning process. These performance criteria provide the analyst with a basis for immediately evaluating how well he can expect a given algorithm to perform on test data. The ADAPT programs provide the analyst with both the performance criteria and the experience factor required to determine whether the algorithm derived is overdetermined. If the algorithm is overdetermined, the analyst must adjust the dimensionality of the problem or increase the quantity of learning data to derive a physically meaningful algorithm.

The ADAPT approach of obtaining the optimum representation of the data prior to performing the analysis introduces the capability to perform clutter subtraction on the data prior to performing the analysis. The clutter subtraction can be used to eliminate any characterizable aspect of the signature from the data histories. This is accomplished by subtracting the coordinate directions corresponding to those characteristics to be eliminated from the space prior to the optimization procedure. Another unique capability resulting from the optimum representation step is the ability to do an extrapolation making use of both historical data from previous data histories and the available portion of any given data history. Conceptually this is equivalent to utilizing historical information to guide the interpolation over missing data points.

In addition to these advantages which accrue from the optimal representation, the ADAPT programs have been operational since approximately 1965. They have been applied to a great many different problems, and during this period part of the practical pit falls associated with empirical analysis have been encountered, overcome and the programs improved to take advantage of this experience. This experience has also provided Avco with the understanding of what diagnostic outputs are required to enhance the ability of the analyst to develop the required algorithms, and to provide the data necessary to reintroduce the physics to the problem at as many points as possible. The key areas where the physics may be reconsidered as part of the analysis are: 1) at the time of data selection and preprocessing decisions; 2) after the development of the optimum representation,it may be examined to insure that the variation is consistent with the expected variation based on the physics of the problem; 3) after the development of the algorithm, the relative importance vector may be examined to determine if the variables which appear important to the decision are consistent with the analyst's understanding of the physics and the relative importance spectrum may be examined to determined if the difficulty in obtaining the algorithm is consistent with the difficulty which would be expected based on the physics of the problem.

In summary, the capability to find the optimum representation for large data vectors has been combined with many years of experience in using this representation as a preliminary step preceding empirical data analysis. This unique combination has been used to prepare a set of computer programs for performing empirical data analysis. These programs provide the user with a fast and economical way to generate simple empirical algorithms for classification, regression, clustering and extrapolation and/or analysis of any given set of learning data.

APPENDIX B

OPTIMAL ORTHOGONAL EXPANSION FOR TWO FUNCTIONS

We wish to carry though the ADAPT expansion of each of two given functions in the series of the optimal orthogonal functions defined by these two functions, as described in the Introduction.

Suppose we are given the functions $u_1(t)$ and $u_2(t)$ of the independent variable t, over some domain $t_1 \leq t \leq t_2$. Let the functions be normalized, so that

$$\int u_1^2 dt = \int u_2^2 dt = 1$$

Then the only parameter is the product integral

$$c \equiv \int u_1 u_2 dt, \qquad |c| \leq 1$$

the last inequality being Schwärz' inequality for normalized functions.

First we construct an orthonormal set of 2 functions v_1, v_2 from the given ones by the Gram-Schmidt procedure.* These functions are easily seen to be

$$v_1 = u_1 \quad , \quad v_2 = (u_2 - c u_1)/\sqrt{1-c^2}$$

We now find the expansion coefficients of u_1, u_2 in a series of v_1, v_2:

$$u_i = x_{i1} v_1 + x_{i2} v_2, \qquad x_{ij} = \int u_i v_j dt$$

$$x_{11} = 1, \quad x_{12} = 0, \quad x_{21} = c, \quad x_{22} = \sqrt{1-c^2}$$

* Note that the Gram Schmidt procedure represents the continuous function Ui(t) by two discrete components, which may be treated similar to the components of the ADAPT data vectors.

The optimal orthogonal functions are now obtained by finding the eigenvalues and eigenvectors $\underline{d}$ of the two-by-two matrix

$$S = \frac{1}{2}\left[x_{1i}x_{1j} + x_{2i}x_{2j}\right]$$

(the factor in front corresponds to weighing by dividing by the number of functions, in our case 2.) They are easily found to be

$$\lambda_1 = \frac{1}{2}(1+|c|) \quad , \quad \lambda_2 = \frac{1}{2}(1-|c|)$$

$$\underline{d}_1 = (\sqrt{\lambda_1}, \sqrt{\lambda_2}) \quad , \quad \underline{d}_2 = (\sqrt{\lambda_2}, -\sqrt{\lambda_1})$$

The eigenvectors are the expansion coefficients of the optimal orthogonal functions h_1, h_2 in a series in v_1, v_2, i.e.,

$$h_i = d_{i1}v_1 + d_{i2}v_2 \quad , \quad \underline{d}_i = (d_{i1}, d_{i2})$$

Returning to the original u functions we find the associated optimal functions to be

$$h_1 = \frac{1/2}{\sqrt{\lambda_1}}\left(u_1 + \frac{c}{|c|}u_2\right) \quad , \quad h_2 = \frac{1/2}{\sqrt{\lambda_2}}\left(u_1 - \frac{c}{|c|}u_2\right)$$

and the expansions of the u functions in them are

$$u_1 = \sqrt{\lambda_1}\, h_1 + \sqrt{\lambda_2}\, h_2 \quad , \quad u_2 = \frac{c}{|c|}\left(\sqrt{\lambda_1}\, h_1 - \sqrt{\lambda_2}\, h_2\right)$$

It is sufficient to discuss the case of $c \geq 0$ because if $c < 0$, a change in the sign of u_2 returns to the first case. We note that the optimal function h_1 is proportional to the average of the input functions. The average is intuitively the best single function to represent two functions, so we see the best single function is associated with the larger eigenvalue λ_1. The optimal function associated with λ_2 is proportional to the difference of the given functions.

We also note that

$$\lambda_1 + \lambda_2 = 1, \quad \lambda_1 - \lambda_2 = c = \int u_1 u_2 \, dt \leq 1$$

The decrease in the eigenvalue from the first to the second is the product integral of the two functions. If the functions are closely correlated one would expect c to be near unity, and λ_2 would be much less than λ_1. But if the functions are nearly uncorrelated one would expect c to be small, and there is only a slight decrease in the eigenvalue, going from the larger to the smaller. Thus the rate of decrease of eigenvalues can be associated with the degree of correlation of the input functions.

APPENDIX C

PERFORMANCE EVALUATION OF FISHER DISCRIMINATE

1.0 INTRODUCTION

The task of developing useful empirical algorithms may be divided into the following three parts:

1. Generation of the algorithms,
2. Performance evaluation of the algorithm, (i.e. a goodness measurement for the algorithm), and
3. Establishment of the validity of applying the algorithm to the test data.

In the ADAPT programs, the most common technique for developing empirical classification algorithms is the use of the Fisher linear discriminant. This has been found to be one of the most useful techniques for generating classification algorithms. It is applicable to non-Gaussian data. For Gaussian data it is possible to define various optimum classifiers including various maximum likelihood separations, optimum quadratic classifiers, etc. However, experience has shown that Gaussian data is very rare in nature. For non-Gaussian data linear classifiers have the advantage that for sufficiently large data spaces the dot product operation normally falls within the criteria for application of the Central Limit Theorem and therefore produces projection values which have Gaussian distribution even when the input data is not Gaussian. This phenomena allows one to significantly improve the performance evaluation of the algorithm. Another advantage of the linear classifier is the extremely simple format making it easy to implement either as a subroutine in a larger program for use on a digital computer, or in a special purpose computer.

The classical approach to establishing the validity of applying a given empirical algorithm to test data is to reserve a certain portion of the available data as test cases. The algorithm is then developed using only that portion of the available data designated as learning data and then applied to the independent proof data. When the amount of data available is limited, which is usually the case, one technique which is often used is that known as "holding one out". In this technique one case is removed from the data set and the remaining cases used to develop the algorithm. The algorithm is then tested on the one remaining case and its performance noted. This case is then added back into the learning set and a different case withheld and the procedure repeated. Since in general algorithm development is considerably more expensive then testing, this approach is more expensive to implement than the approach of retaining a large proof test data but it does allow one to perform the evaluation using a smaller set of data. It should be pointed out that this classical approach is neither necessary nor sufficient in a rigorous sense for ensuring the applicability of the algorithm to a new set of data. In particular, care must be exercised in selecting the independent proof sample such that its selection reasonably models the selection of the entire sample from the population or universe of data.

The ADAPT programs in addition to providing the capability to implement this classical approach to establishing the validity of applying the algorithm to the test data, also utilizes a validity criteria to test each individual history for similarity to the learning data. If the test case is not sufficiently similar to the learning data then one cannot feel confident in applying an algorithm derived on the learning data to this particular test case. The ADAPT measure of similarity is the relative reduction in explained variation as one proceeds from the original data space to the optimum ADAPT space for the test data case as compared with the learning data cases. Clearly it is necessary, but not sufficient, that the test case be adequately represented in the ADAPT base derived from the learning data for any empirical algorithm to be valid. The ADAPT programs furnish the user with information to judge the degree of similarity which is required between the learning and test data. The ADAPT programs generate as part of the algorithm development a relative importance spectrum which defines how much of the explained variation is required to develop a meaningful algorithm. One criteria is that the representation of the test case on the learning base must explain at least as much variation as explained by the first "L" terms of the representation. Here "L" is defined as the maximum number of terms required to include all of the important terms, a defined by the relative important spectrum, for the particular algorithm being analyzed. A simpler, but significantly less rigorous criteria which is often used is that the minimum representation of the test case must be greater than the minimum representation observed on any learning data. Clearly, this requirement is a necessary but not sufficient requirement to ensure adequate representation.

The preceding portion of this appendix has reviewed the ADAPT approach to the generation of algorithms and to the establishment of the validity of applying the algorithm to a given test case. These have been quite general and are applicable to a large number of linear discriminants. The remaining sections of this appendix will develop procedures for evaluating the performance of separation algorithms derived using the Fisher discriminant. The great majority of ADAPT classification problems are solved using the Fisher discriminant. The procedures for defining its performance have been refined considerably further than for other discriminants included in the ADAPT programs. In general, these procedures could be used as a guide for establishing performance measures for many of the other discriminants included in the ADAPT programs. The following discussion is divided into two parts. The first part, discussed in the next section, is that of establishing a threshold for the ADAPT algorithm to achieve a special goal. The second part, discussed in the last section, is the measurement of the performance of the algorithm with a given threshold.

2.0 SETTING OF FISHER THRESHOLDS

The approach to setting the threshold to be used to classify the projection value obtained from applying the Fisher discriminant is based on the analysis presented by Anderson and Bahadur in Reference 1. Strictly speaking, this analysis requires that all possible projection vectors produce Gaussian projections. In general, this is only true if the input data is itself Gaussian. For the great majority of projection directions, in particular those directions which are normally determined by the application of the Fisher discriminant, the Central Limit Theorem will result in a Gaussian projection. Thus, although the theory is not rigorously applicable, it is usually applicable to a large percentage of the possible projection directions when the data space is sufficiently large to invoke the Central Limit Theorem. Thus, one suspects it may still be a valid guide as to the selection of the Fisher weighting parameter and the threshold to be used with the Fisher discriminant. Experience with a great variety of data has shown that this is indeed the case.

Reference 1 shows that if one desires to minimize total number of errors made by the Fisher classification algorithm one should select the Fisher weighting parameter, P, according to the following relation:

$$P\sigma_1 = (1-P)\sigma_2 \tag{1}$$

where σ_1 and σ_2 are the standard deviation of the projection values of the first and second classes respectively. Assuming that the origin has been selected mid-way between the means of the projection values of each class the threshold, TH, is given by:

$$TH = (\tfrac{1}{2} - P)V \tag{2}$$

Another criteria which one may wish to use, rather than minimizing the number of errors, establishes an algorithm which will achieve a desired false alarm rate. This special case is also discussed in Reference 1. Suppose one desires a probability $\overline{P}_N$, that there will be no false alarms in Class 1 when N Class 1 cases are examined (i.e. no Class 1 cases will be classified as belonging to Class 2.) The following relation will define the false alarm probability for Class 1, P^I_{FA}:

$$(1 - P^I_{FA})^N = \overline{P}_N \tag{3}$$

Solving this equation for the probability of false alarm for Class 1 under the assumption that P_N = is equal to 0.5 gives:

$$P_{FA}^{I} = 1 - \exp(\ln \bar{P}_N / N) \underset{N \to \infty}{\sim} \frac{0.693}{N} \tag{4}$$

Once the desired false alarm rate has been defined, Reference 1 shows that the proper Fisher weighting parameter to achieve this false alarm rate is given by:

$$P\sigma_1 = \beta^{I}(1 - P_{FA}^{I}) \tag{5}$$

where β^{I} is the variable in the cumulative standard normal distribution function of the probability $1 - P_{FA}$. The corresponding threshold is given by:

$$TH = \mu_1 - \beta^{I}\sigma_1 \tag{6}$$

where μ_1 is the mean of Class 1 and σ_1 is the standard deviation of Class 1.

The above equations, although strictly valid only for the case of Gaussian data, may be expected to give a good approximation even in the case where the data is not Gaussian, when the data space is relatively large. Experience with the utilization of these equations in a large number of real problems has verified that they do provide a good guidance for the selection of both the

Fisher weighting parameter and the best threshold to achieve either the goal of minimum errors or a predefined false alarm rate.

3.0 PERFORMANCE MEASURES

The simplest measure of the performance of a linear classification algorithm is to examine the projection values actually obtained on the learning and/or test data by applications of the discriminant. The ADAPT programs present a bar chart plot of these projection values for each of the learning cases, which can be used to visualize the performance of an algorithm on the learning data. However, these plots are extremely inconvenient for evaluating a large number of algorithms. Although the information required to determine the trade-off between detection probability and false alarm rate is on these bar charts, they are not very convenient for visualizing this trade-off. The most desirable way to evaluate the performance of a large number of algorithms is to obtain a single number which measures the quality of the algorithm. Since the Fisher discriminant is the result of a maximization of the quantity V, which can be defined by

$$V = \frac{|\mu_1 - \mu_2|^2}{P\sigma_1^2 + (1-P)\sigma_2^2} \tag{7}$$

it is clear that the maximum value of V is itself a good measure of the performance of the algorithm. The maximum value of V, over all possible projections, turns out to occur when the denominator of Equation 7 is equal to the square root of the numerator, which means V becomes, geometrically, the distance between the means of the projection of the two classes on the Fisher direction. Thus for the Fisher discriminant Equation 7 provides a relationship between the projection of the means of the two classes, the standard deviation of each class, and the Fisher weighting parameter.

It is interesting to consider the special case in which the standard deviation of each of the classes is equal. For this case,

$$\sigma^2 = V = |\mu_1 - \mu_2| \tag{8}$$

and

$$(\sigma_1 + \sigma_2) V = 2/\sqrt{V} \tag{9}$$

Thus the special case of equal standard deviations allows us to get a good physical comprehension of the parameter $2/\sqrt{V}$. This parameter is used as a measure of the goodness of performance of the discriminant. Regardless of the relationship between the standard deviations of the two classes, the smaller this parameter (the larger V) the better the performance of the algorithm. In the particular case where the standard deviations of both classes are equal this parameter is just equal to the sum of the standard deviations divided by the distance between the mean. The resulting simplification is very instructive for both methods of setting the threshold.

For the case where one wishes to minimize the number of errors, the situation is shown in Figure C-1. The threshold is set half way between the mean projections of the two classes, because the criterion requires that the errors for the two classes are the same. Then the probability of error is the shaded Area A_1 which is the value of the cumulative normal distribution centered on μ_1, up to μ_1 -V/2. If G is the standard cumulative normal distribution, this is

$$P_E = G\left(\frac{-V/2}{\sigma}\right) = G\left(\frac{-1}{2/\sqrt{V}}\right) \tag{10}$$

P_E is the probability of making an error in either class, and

$$P_D = 1 - P_E \tag{11}$$

is the probability of correctly identifying a member of either class.

For the case where one wishes the maximum detection of Class 2 for a specified error probability P_{FA}^{I} of Class I, the threshold is set by Equation 6. Again, for equal standard deviations the situation is quite simple. Take the zero of projections half way between the mean projections of the two classes. Then μ_1 = V/2, and Equation 6 becomes

$$TH = + V/2 - \beta^{I}\sigma \tag{12}$$

or

$$\beta^{I} = (V/2 - TH)/\sigma \tag{13}$$

This is the standard normal deviate at which

$$P_{FA}^{I} = G(\beta^{I}) \tag{14}$$

The detection probability P_D^{II} of Class 2 is the area under the normal curve centered on $\mu_2 = -V/2$ up to TH. The normal deviate for this curve at that point is

$$\beta^{II} = (TH + V/2) \tag{15}$$

and

$$P_D^{II} = G(\beta^{II}) \tag{16}$$

But TH can be eliminated from (13) and (15) to give

$$\beta^{I} = \sqrt{V} - \beta^{II} = \frac{2}{2/\sqrt{V}} - \beta^{II} \tag{17}$$

Thus for the case of equal standard deviations, the detection probability of Class II depends only on the false alarm probability of Class I, and the Fisher maximum through $2/\sqrt{V}$. Figure C-2 presents these ROC curves for various values of parameter $2/\sqrt{V}$. Thus we see that the parameter $2/\sqrt{V}$ is useful for both visualizing the bar charts of the performance of the algorithm as well as for visualizing a Receiver Operating Characteristic for this particular algorithm. It therefore has considerable intuitive value for rapidly judging the performance of the classification algorithm. For these reasons it is used in ADAPT as the parameter for evaluating the performance of the ADAPT derived Fisher discriminant. In addition to obtaining an understanding of the trade-off between detection probability and false alarm rate, it is important to have a measure of algorithm performance to evaluate the effect of dimensionality of the space in which the algorithm is derived. This is extremely important since the use of too large a dimensionality in the derivation of an algorithm will result in the algorithm being derived by fitting the learning data according to special characteristics of the particular learning sample, and not according to characteristics of the population sample. That is, the major basis for the separation will be the difference between the population and sample means, rather than the difference between the means of the two populations being classified. This phenomena is quite an analogous to the fitting of a third order polynomial through a set of data. If a third order polynomial is fit to 3 data points, there is no reason to believe that a general law has been derived. However, if this same third order polynomial makes a reasonably good fit to 100 points, there is little doubt that these 100 points are related by some phenomena which is well expressed by a third order polynomial.

Thus, it is important to understand the capabilities of a Fisher discriminant to derive classification algorithms simply on the difference between sample and population means. In many years of ADAPT experience, this was evaluated by performing separations of odd cases versus even cases from both classes for each problem being considered. The performance of these separations were then compared with the performance of the classification algorithm derived between the desired classes. If the algorithm derived for separating the odd versus even gave a similar performance to the desired algorithm then one concluded that the algorithm was not based on physical characteristics but rather on the differences between the sample and the population means. This experience can be summarized in a plot such as presented in Figure C-3. This figure plots the number of cases divided by the number of dimensions versus the performance measure obtained for separations of odd from even (i.e. random separations) for a large variety of problems and data. The extrapolation of this curve for low values of the performance measure was accomplished by making a similar plot on a linear scale and noting that for a number of cases over number of dimensions of unity the performance measure should to go 0. It is interesting to compare Figure C-3 with the results of a similar analysis presented in Reference 2 indicated that when a number of cases to number of dimensions exceeded six, one could have confidence in the performance of the algorithm. Figure C-3 clearly shows why this is the case. Remembering that we may relate $2/\sqrt{V}$ to the probability of error we note that for a performance measure of 2 the probability of error is approximately one in three. Since a random process for selecting a class has a probability of error of one in two, it is clear that an algorithm whose performance measure is two or greater is probably not of very great interest. Thus this curve shows that any algorithm of interest derived in a space such that the number of cases divided by the number of dimensions is greater than six lies to the left of all of the data shown in this figure.

A performance map can be defined which combines all of the characteristics of this performance measure into a single plot. Figure C-4 presents a sample of such a plot. The ordinate of this plot is the ratio of the number of cases to number of dimensions used to derive the algorithm. The abscissa is either the performance measure or the probability of error depending which scale we wish to read. Thus when an algorithm is derived using the Fisher discriminant it may be placed at some point in this figure simply by noting the number of cases used in the learning data, the number of dimensions of the space in which the algorithm is derived, and the performance measure for that algorithm. All of these parameters are available in the ADAPT output for the deviation of the Fisher discriminant. If the algorithm occurs to the right of the cross-hatched region in this figure, one knows that it cannot be applied to test data and is not a general algorithm. If it falls near but to the left of the cross-hatched area, one realizes that the performance of this algorithm on the learning data is significantly better than one can expect on the test data. Only if the algorithm falls to the left of and reasonably far away from this cross-hatched area does one have an algorithm whose learning data performance is indicative of the performance which can be expected on a test case.

References for Appendix C

1. Anderson, T. W. and Bahadur, R. R., "Classification into Two Multi-Variate Normal Distributions with Different Co-Variance Matrices" Annals of Mathematical Statistics, Vol. 33, p. 420, 1962.

2. Foley, Don H., "Probability of Error in the Design Set As A Function of A Sample Size Dimensionality", Thesis, Syracuse University, 1971.

FIG C-1 DISTRIBUTION FUNCTIONS FOR PROJECTIONS ON FISHER DIRECTION FOR $\sigma_1 = \sigma_2$

$G\{\beta\}$ = CUMULATIVE STD. NORMAL DISTRIBUTION CURVE WHERE : $\beta = \frac{W_K - \mu}{\sigma}$

$$= \int_{\infty}^{\beta} f\{Z\}\, dZ$$

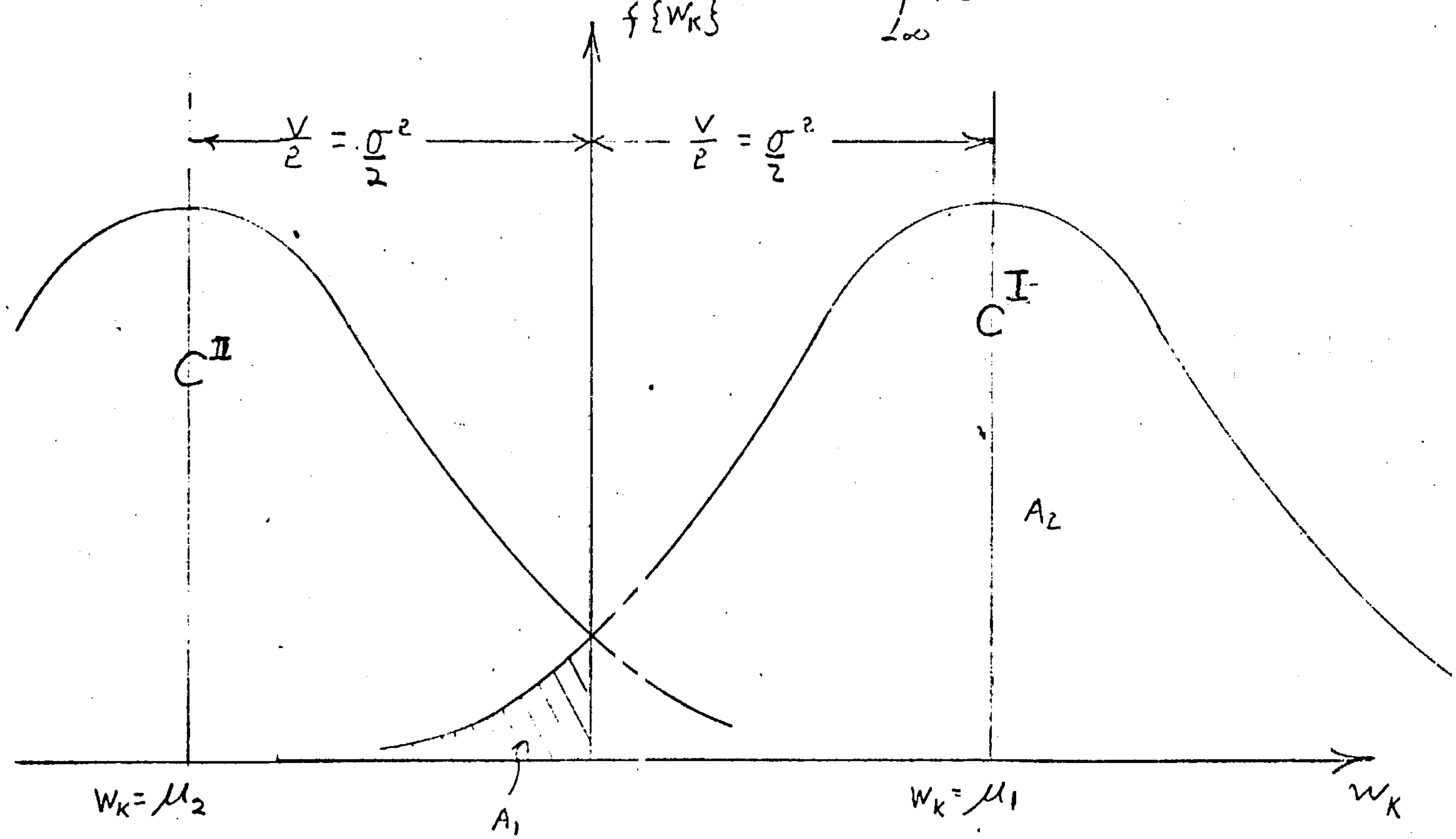
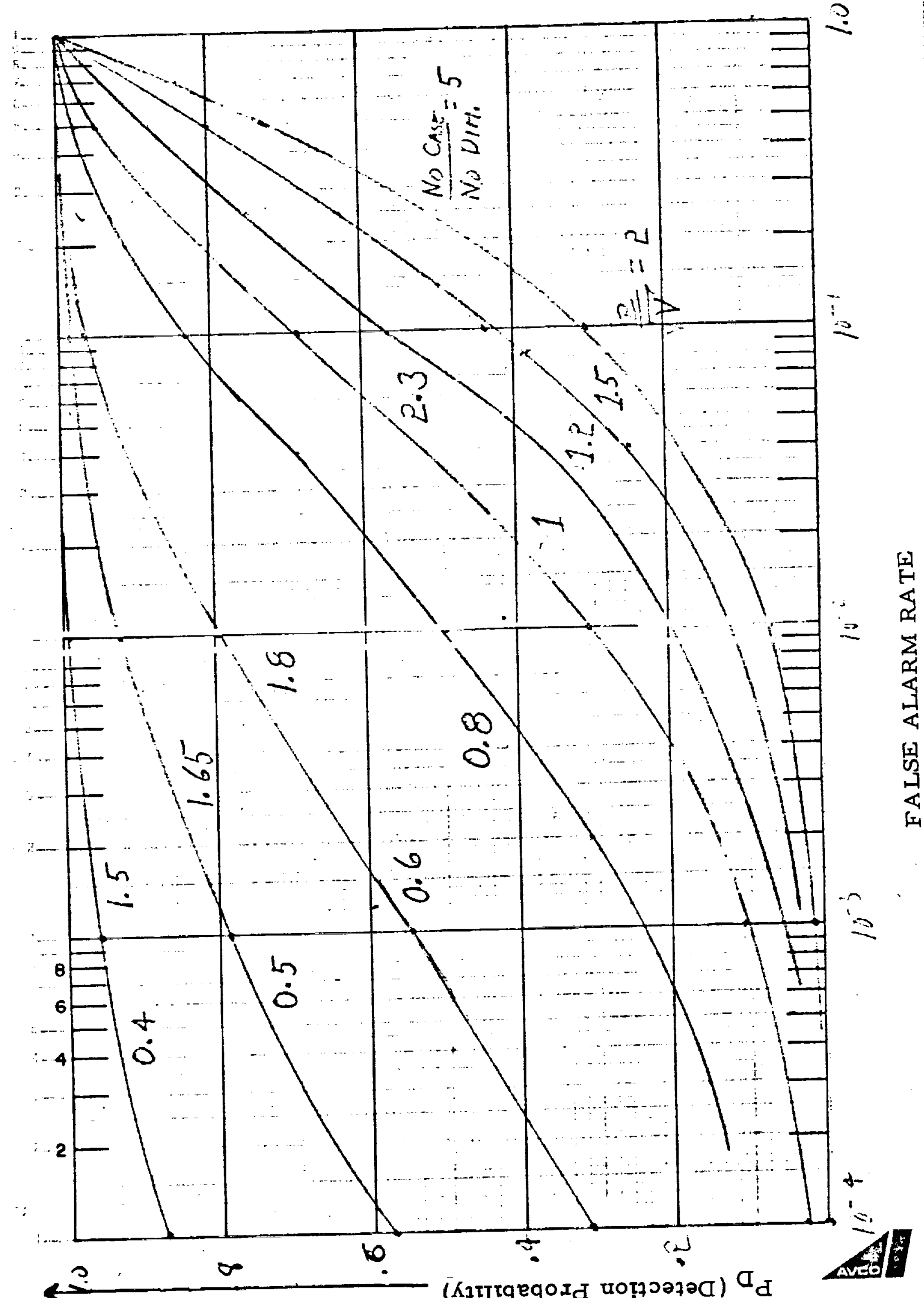


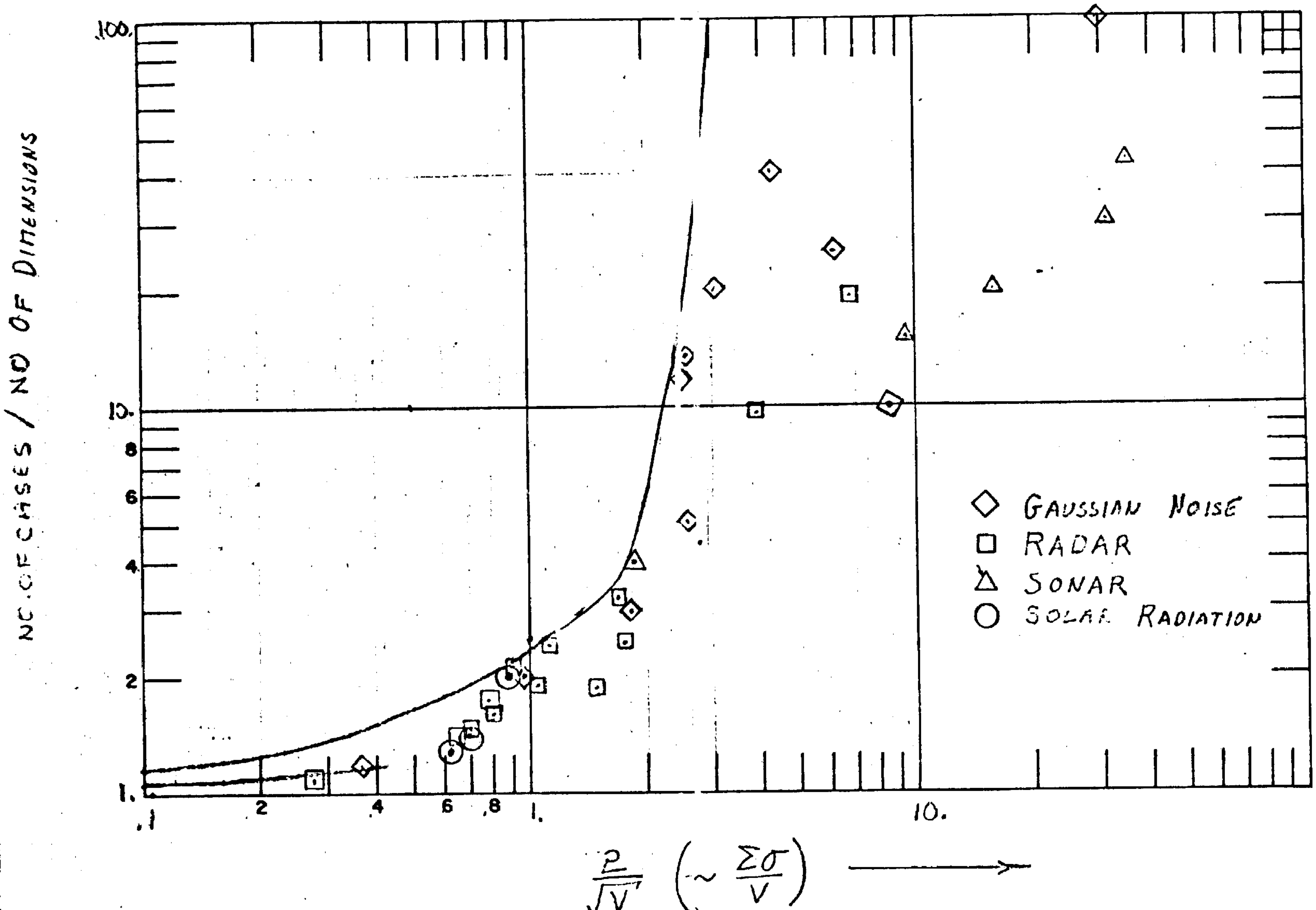
FIGURE C-2 FISHER R.O.C. CURVES VS $\frac{\Xi}{V}$ FOR $\sigma_1 = \sigma_2$



DATE

NOTE: USE TYPE B PENCIL FOR VUGRAPHS AND REPORT DATA.

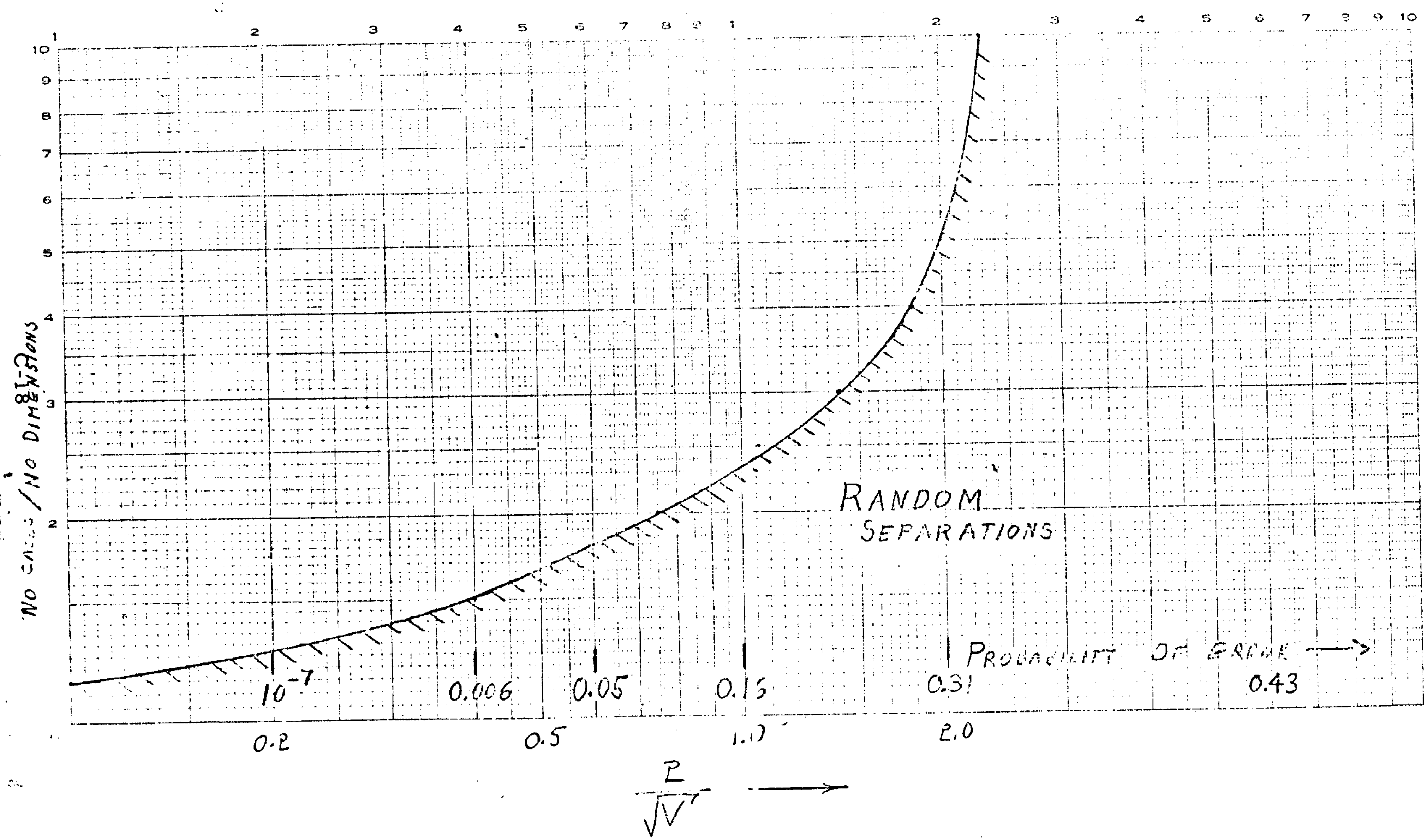
FIG C-3 PERFORMANCE OF RANDOM SEPARATIONS



NO. 340-L21 DIETZGEN GRAPH PAPER
LOGARITHMIC
2 CYCLES X 1 CYCLE

EUGENE DIETZGEN CO.
MADE IN U.S.A.

FIG C-4 PERFORMANCE MAP



APPENDIX D

PROCEDURE FOR IMPLEMENTING VALIDITY CRITERIA

The validity criteria is that the value of Q_K must exceed Q_K-Min. where

$$Q_k = \frac{\sum_{\ell} Y_{K\ell} Y_{\ell K}}{\sum_{i} V^T_{Ki} V^T_{iK}} \tag{1}$$

and

$$Y_{\ell K} = \sum_{i} (\text{Vector}-\ell)_i \, V^T_{iK} \tag{2}$$

Q_K-Min. is determined by analysis of the learning and/or proof test data. The V^T_{iK} are the components of the test case being evaluated and the Vector- ℓ are given in Table D-1 for the universal boiler #1 detection algorithm.

TABLE D-1

EIGEN VECTORS

VECTOR 1	-1.3717549D-03	2.3665290D-02	-1.0721605D-02	2.3370850D-02	1.3476998D-02	2.4861606D-01	-9.8730117D-03
-5.3214399D-02	4.5163736D-01	-4.4563836D-01	1.0037895D-02	-3.0621530D-03	-5.7553546D-03	7.7911449D-02	1.5764171D-01
1.1266433D-01	-1.7515253D-02	-8.9193578D-03	-3.7022946D-02	4.4272770D-01	-4.4640796D-01	1.0102627D-02	-8.3743103D-05
-5.3650425D-03	2.6860250D-01	7.8475590D-02	1.2313579D-02	2.0476068D-02	1.6342735D-02	-2.4888561D-02	3.5009952D-02
3.3870247D-02	3.0663217D-02	7.2277489D-02	3.7653969D-02	1.1642118D-02	1.4458296D-02	3.1546779D-02	3.5899100D-02
6.2689327D-03	-3.5272832D-03	1.2031091D-02	-2.4208591D-02	-1.7046216D-02	1.7726603D-02	1.3230371D-02	2.7596543D-02
2.1086143D-02	4.6622000D-02	-8.4336171D-02					
VECTOR 2	-2.7128345D-02	-5.5572372D-02	-2.0480621D-02	-3.4664515D-02	6.1018128D-02	2.9138691D-01	3.8737493D-02
3.4485552D-01	-1.5332519D-01	1.4885600D-01	1.3551870D-01	1.7701172D-02	1.2629562D-01	1.7550571D-01	2.8500339D-01
3.9912383D-01	2.2193442D-02	3.5831503D-02	2.7302161D-01	-1.3835460D-01	1.4653707D-01	1.3072665D-01	3.6999152D-03
1.2627675D-01	3.1327324D-01	1.6943301D-01	9.7213397D-02	1.1275627D-01	-2.9713966D-02	-2.9572129D-02	-5.6284962D-02
-3.1421472D-02	2.8515067D-02	-5.6899344D-03	-6.9206564D-02	1.0597706D-02	-1.0630659D-02	-7.3410487D-02	-1.0197468D-02
1.2220739D-02	4.9269998D-03	-1.8718884D-02	3.2204282D-02	4.9868336D-02	-1.0274715D-02	-4.8303986D-02	-1.3921678D-03
1.3096429D-03	4.5547533D-02	-3.3638262D-01					
VECTOR 3	-6.6606269D-03	-6.9341353D-02	2.6626019D-02	-1.0051304D-01	5.9018285D-02	-4.7116323D-02	-2.3647739D-02
5.0293045D-01	7.6102880D-02	-9.8628321D-02	-8.4711302D-02	1.4991260D-02	-2.7915167D-02	-6.5733574D-02	-1.8657143D-01
-5.0916459D-01	-3.0072643D-02	-1.7565660D-02	3.6152038D-01	8.3501327D-02	-9.3681342D-02	-9.2489215D-02	6.0713191D-03
-3.7171695D-02	-3.1161466D-02	-7.9189923D-02	8.9368156D-02	1.6273166D-01	-7.4855468D-02	2.9007331D-01	-3.7875657D-02
-8.3265714D-02	-2.0150536D-02	-1.7154686D-02	2.1383254D-02	-6.0199570D-02	-7.5128530D-02	3.9499607D-02	-6.6791565D-02
-2.4168493D-02	-2.0201440D-02	-2.2476440D-02	-2.3108632D-02	-2.5804530D-02	-1.7947532D-02	2.5870089D-02	-3.7627824D-02
-3.4664695D-02	-8.7301411D-02	-2.6302252D-01					
VECTOR 4	5.0977141D-02	-4.0889181D-02	7.1817016D-02	8.5109350D-03	7.5306166D-02	5.7996329D-02	1.1568481D-02
3.7769771D-01	-1.3762060D-02	-1.4780760D-02	4.0056557D-02	2.8845016D-02	2.2113735D-02	8.7019893D-02	6.2552228D-02
5.4711502D-03	-2.6092065D-03	1.7992908D-02	2.2584730D-01	2.1721020D-02	-2.4121403D-02	3.2863983D-02	2.7202321D-02
2.5529367D-03	6.4949640D-02	1.1179152D-01	7.3917478D-02	1.2618208D-01	-1.2378209D-02	1.5937246D-02	4.0399694D-02
3.0599497D-02	-2.2297630D-02	1.1931597D-01	6.6828536D-02	-4.3213386D-02	3.2794597D-02	1.3773825D-02	7.2562531D-02
1.8900084D-02	1.0254196D-02	7.6785333D-03	1.0360919D-02	1.7590528D-02	2.2110703D-02	3.0710748D-02	4.3976056D-02
2.9556715D-02	5.4155470D-02	9.5461604D-01					
VECTOR 5	-1.3630625D-02	-2.3774461D-02	7.2713079D-02	-7.2330623D-02	-2.3686177D-02	-3.6653707D-01	5.8260295D-02
1.0177073D-01	9.2224941D-02	-1.4900637D-01	1.1939435D-01	-4.2773209D-02	7.7239551D-02	-4.4952721D-02	-1.8012432D-01
4.6006125D-01	1.6393206D-02	6.1115709D-02	1.4587261D-01	1.3513799D-01	-1.2271514D-01	1.2723267D-01	2.2806888D-02
9.3942797D-02	-3.3742227D-01	-4.3823369D-02	5.8665703D-02	1.1114559D-01	-1.9925548D-02	-1.2262357D-01	-1.9750208D-01
-1.2920493D-01	4.9902524D-02	-1.1827675D-01	-1.0645751D-01	-1.2239902D-01	-7.7354256D-02	-1.7587103D-01	-1.0859232D-01
-1.2688381D-01	-1.3744048D-01	-1.1759512D-01	-3.5234791D-02	-2.9323727D-02	-8.3719577D-02	-1.0775595D-01	-1.8656687D-01
-1.4737733D-01	-1.4155734D-01	5.5626429D-02					
VECTOR 6	4.9626430D-02	2.9859996D-02	1.4620725D-02	6.9723747D-02	-7.6502099D-02	-4.4310074D-02	2.0141948D-02
-1.9624174D-01	2.9625013D-03	5.7994778D-03	5.7054821D-02	-8.9731307D-02	4.0193512D-03	4.2906128D-03	-3.9973275D-02
2.9102837D-01	7.6439267D-02	4.8895662D-03	6.7263823D-02	-5.9430261D-03	2.9642627D-03	8.3179797D-02	-4.2752610D-02
2.9210774D-02	-5.2157050D-02	2.3614082D-02	6.2966592D-02	-2.6554522D-02	5.5563931D-03	8.8152473D-01	2.1645662D-02
1.1205502D-01	-2.0206756D-02	7.4796270D-02	4.1187830D-02	4.4154592D-02	6.3982087D-02	5.0555024D-02	3.2867712D-02
1.0499682D-02	-7.2007359D-02	7.1350067D-02	-3.2860290D-02	-4.6883995D-02	3.3391932D-02	-9.3772676D-03	5.7049310D-02
1.3212691D-02	-1.0521671D-02	2.3844032D-03					
VECTOR 7	3.8293545D-02	7.4238178D-03	-1.1525437D-01	4.0250209D-03	6.8317760D-02	-1.3106956D-01	-1.6613750D-01
1.0663157D-01	-4.0115777D-02	-4.3920993D-02	2.5069655D-02	-1.1202827D-01	1.5642103D-01	-7.1486629D-03	-3.5769204D-01
7.3000707D-01	-7.5479114D-02	-1.4156634D-01	1.1292345D-01	2.9395052D-02	-2.6255401D-02	1.0494971D-02	-1.3105511D-01
1.2251056D-01	-1.3627631D-01	-7.4112056D-03	2.9412692D-02	-6.6687573D-02	3.3427982D-02	-2.3705331D-01	2.2060333D-01
2.0614142D-01	-1.2040785D-01	2.1925301D-01	1.0016935D-01	1.3542408D-01	1.5768539D-01	1.6169738D-01	8.9331343D-02
1.3443290D-01	1.2449703D-01	1.0076902D-01	-1.1016303D-01	-1.3574366D-01	1.4478537D-01	1.1269000D-01	2.2954661D-01
1.1300722D-01	-4.5224962D-02	-1.5507980D-01					
VECTOR 8	-6.2557477D-02	1.4746944D-02	5.4792609D-02	-1.2207816D-01	2.3079724D-02	1.5701579D-01	-2.9761181D-02
-2.1281063D-01	-1.6619111D-01	-3.0062421D-02	-3.7461115D-02	-9.5905703D-02	1.5013477D-02	4.4510561D-02	-1.3322628D-01
-2.5744973D-02	-4.4692106D-02	-3.9037394D-02	6.0771106D-01	2.9906985D-02	-2.6533980D-02	-1.4349673D-02	-2.8308934D-02
7.4157583D-02	3.7[illegible]69250D-02	-1.5750467D-02	1.4029613D-01	2.6850164D-01	-6.7331504D-02	-1.3572734D-01	-1.0428503D-02
-1.3395127D-04	-7.5201026D-02	-1.8431367D-02	-2.8421910D-02	3.3973073D-02	1.4967806D-02	5.0450377D-02	-7.6865197D-02
-1.1121215D-02	-2.2[illegible]D-03	2.1136174D-02	1.3109983D-02	2.4175947D-03	2.9337566D-02	-4.1389611D-02	2.7304537D-03
2.2000267D-02	8.4527778D-03	2.3271784D-02					
VECTOR 9	-7.0276990D-02	1.3149863D-01	-1.5477119D-02	-2.4674554D-02	7.4570259D-02	-1.0565129D-01	9.1870176D-02
-7.[illegible]D-02	-4.0995707D-02	5.2761144D-03	4.3260367D-02	1.3905125D-01	1.2175529D-01	-2.0702353D-01	1.9019894D-01
-5.[illegible]D-03	-6.3473196D-02	7.7574901D-02	-3.9611230D-02	1.3017197D-02	-6.5031385D-04	1.4003026D-02	6.4597829D-02
[illegible]D-02	-3.2296767D-02	-2.1420236D-01	5.0060554D-02	3.4020850D-01	4.8398202D-02	2.5735424D-02	2.2185053D-01
-1.1[illegible]D-01	4.1964952D-02	2.7533619D-01	2.8963972D-01	-6.9491459D-02	-4.3307781D-02	-2.6804249D-01	3.4839421D-01
-1.5[illegible]D-01	-9.7713755D-02	-2.2653458D-01	-3.2554049D-03	4.2816512D-02	1.1331055D-01	1.8520354D-01	1.3210537D-01
-9.1[illegible]D-02	9.5975102D-02	-7.5033506D-02					
VECTOR 10	3.7[illegible]4665D-02	1.3247007D-02	-4.0508158D-02	-1.1039553D-01	7.0894970D-02	-2.9594941D-01	-1.2470386D-01
-1.[illegible]D-02	-2.0[illegible]1950D-02	-2.1915167D-02	-9.7049996D-02	-7.7529031D-02	-4.7252663D-03	1.3651251D-02	7.5455917D-01
-1.1[illegible]D-02	-7.0793440D-02	-1.5746508D-01	1.0918825D-01	1.9079745D-02	-3.3298236D-02	-1.0563928D-01	-7.5849618D-02
-1.1440191D-02	-2.0515452D-01	1.4219761D-02	7.0914495D-02	-5.6731565D-03	-1.0030674D-01	-2.3920899D-02	4.9970900D-02
2.121[illegible]D-02	-1.4[illegible]D-01	1.9359145D-02	3.9349050D-02	-1.9763761D-02	-2.7111739D-02	2.1706517D-01	-1.1992258D-01
5.[illegible]D-02	3.6020110D-02	1.1347008D-01	-6.7727191D-02	-1.2433405D-01	-6.6795323D-03	1.1210117D-02	-7.5861956D-03
-1.1614733D-02	-6.2560002D-02	-3.1965490D-02					
[illegible]	-1.4598262D-01	-7.2022347D-01	2.2219150D-01	-1.3331399D-01	-4.0020572D-01	-1.0945693D-01	-7.3946511D-02
			-2.2086648D-01	-2.5624131D-01	-1.3429925D-02	1.2573294D-01	1.4010318D-02
				-1.0159458D-01	8.1374533D-02	-9.0854819D-02	-5.6712412D-02

```
 1.5772665D-01  1.3440528D-02  1.1023090D-01 -2.9344445D-02  1.4212133D-01  [illegible]  [illegible]
-1.1970047D-01 -9.8806062D-04 -5.5602378D-02 -9.0524040D-02  7.4247056D-02  6.2941024D-02 -2.1640664D-01  2.5818097D-02
 2.2735064D-01  2.8709464D-01 -1.7896128D-01  8.4589729D-02  1.1623036D-01  1.6076086D-01  1.8839515D-01  9.0385886D-03
 2.5562741D-02  9.2682943D-03  5.5352994D-03
VECTOR 12      -1.9411331D-02 -1.1194059D-01
-1.9355106D-01 -1.2218450D-01 -7.0840296D-02  9.7505574D-03 -2.4850161D-01  5.2718450D-01 -7.0959758D-02 -1.5137202D-01
-3.9030456D-03 -6.1340699D-02 -1.5478826D-01  8.5770835D-02  2.5400145D-01  1.6396777D-01 -1.1093232D-01 -6.8722790D-03
-1.4395746D-02 -1.4187530D-01  1.1834041D-02 -7.5139334D-02  3.5128386D-02 -4.4208364D-02 -1.4255734D-02 -1.1457497D-02
 3.6045327D-02  3.5271330D-02 -6.3496736D-02 -5.5972757D-03 -1.4946481D-01 -1.2975079D-03  1.1972814D-01 -1.9959645D-01
 2.3846957D-01 -2.0687928D-01 -2.8444143D-01 -7.4356621D-02  8.6753175D-02  1.3129200D-01 -2.0194845D-01 -1.0966562D-01
 8.9391922D-02 -1.1924651D-01  4.1869988D-02 -3.7920967D-03 -2.6506531D-02  6.4475111D-02  1.7095066D-01 -1.2373848D-01
VECTOR 13      -3.0203547D-01  2.2263984D-01
 5.2239066D-02  4.8123036D-03 -5.0178993D-02 -3.5087409D-01 -1.6714637D-01 -2.0305554D-01 -1.3883695D-01  1.2003063D-01
-1.0398343D-01 -1.7631243D-01 -9.3510276D-02  6.3864312D-02 -6.7287138D-02 -1.1555532D-01 -2.4597091D-03  1.0848463D-01
-4.1372249D-02 -4.1318996D-02 -8.1651496D-02 -1.8354790D-01 -3.9434846D-02 -1.9056144D-02  9.7304495D-02  7.7810092D-02
 1.6421633D-01  2.5741567D-02 -3.9066286D-02 -4.2428864D-02 -1.5248133D-01  7.2984892D-02 -3.3256648D-03 -3.1921773D-02
-2.4544806D-02 -2.3708518D-02 -2.2619977D-01 -1.5074986D-01  2.7475046D-01  2.5396338D-01 -2.2570038D-01 -1.5009875D-02
 2.1059631D-01  1.5373710D-01 -2.2232467D-02  1.4606769D-01  1.6058903D-01  1.3713641D-02 -1.7948804D-01  7.5607484D-02
VECTOR 14       3.4532921D-01  6.4210368D-02 -3.7639157D-01 -2.4450097D-01  1.9139637D-02 -2.2367925D-02  3.2545894D-03

-1.0812397D-01  1.6349016D-01  9.8221106D-02 -1.6788456D-02 -5.4644657D-02 -1.1493812D-01  2.6288102D-01 -1.1477183D-01
-3.4225810D-04  2.0195572D-01 -3.5247775D-03  1.8720259D-02 -5.6514666D-02  6.3915704D-02 -3.4084664D-02 -9.1721410D-03
-9.0366913D-02  6.0244108D-02  3.2091504D-01  4.7543643D-02 -6.2810646D-02 -1.8635585D-02 -2.6314900D-02 -8.5524680D-02
-2.3217527D-01 -1.4981839D-01  1.3424317D-01  2.4277764D-01 -8.6466725D-02 -9.6873583D-02 -6.2040269D-02  9.6757913D-02
-3.2306648D-02 -3.8621907D-02 -7.8732365D-02 -6.7819668D-02 -7.2890623D-02 -1.4247263D-01  2.3959148D-01 -4.6451119D-02
-1.1988521D-01 -2.7978568D-01 -6.9928111D-03
VECTOR 15       4.1122962D-02 -8.9551800D-02
-8.0632471D-02  3.2868420D-01  9.0231819D-02 -1.7973806D-01  2.8677192D-01  2.1478170D-01 -6.8807992D-02  1.8591255D-01
-1.8377029D-01 -1.4369739D-01  1.7563512D-01  1.6936507D-02 -1.4796945D-01  3.4752092D-01  3.7929989D-02  2.8926833D-02
 2.5514705D-01  7.1923443D-02  7.7231003D-02 -1.0604323D-01 -1.2315763D-01  1.1859565D-01 -1.2953087D-02 -2.0625833D-01
 1.7171405D-01  7.6626099D-03 -4.2314265D-02  1.3792924D-02  2.7966699D-01  1.9817449D-02  4.9895149D-03 -1.7477510D-01
-6.6798575D-02 -1.5971191D-01 -5.0705104D-02  2.4656656D-02  1.4848652D-01  1.4193410D-01  9.8222595D-02 -9.9327676D-02
 6.8363462D-02 -4.2344597D-02  1.5663957D-02 -7.8394592D-02 -7.0167204D-02 -9.2438161D-02 -2.2109020D-02 -1.3867031D-01
VECTOR 16       1.7359603D-02 -1.3109978D-01
-9.3780374D-02  2.6495544D-01  6.0596252D-02  3.6414653D-01  9.7707345D-02  2.3003343D-01 -2.4979606D-01  4.9562931D-02
-1.7423185D-01 -1.7902812D-01 -3.0191179D-02  2.5168984D-01  1.3271638D-01 -2.8730569D-02  2.1722841D-01 -9.0855489D-03
-9.4234724D-02 -1.1445256D-01  3.3041779D-01  1.2478763D-01 -3.2020203D-02  7.6915464D-02  1.7598647D-01  5.0385554D-02
-3.3417287D-02 -2.8832299D-02 -3.3597612D-02  5.2367101D-02 -1.6124656D-01  2.7223560D-02 -3.8839660D-02  2.7981568D-01
-8.0776223D-02 -1.2870942D-03 -5.0625555D-02 -5.7330449D-02  7.0952050D-02  6.4585602D-02 -9.6291269D-02  2.9082320D-02
-3.4371907D-02  1.9968470D-01 -1.2013792D-01  9.7268288D-03  2.3496952D-02  4.7043270D-02 -1.0843085D-01  2.2205228D-01
VECTOR 17       6.9800518D-01  3.9111509D-02
-1.4101283D-02 -2.9333436D-01 -1.1950594D-01  4.1770347D-01  2.5642084D-01 -1.0618684D-02 -4.6696126D-06  2.5125560D-02
-5.0254803D-03 -2.1353065D-01 -4.6235176D-03 -7.7602454D-02 -3.6769594D-02  1.1120661D-01  8.0218201D-02  1.6505457D-02
 9.1285944D-02  1.0160113D-02  4.8410642D-03  2.3044308D-02  5.3350080D-02 -4.8023696D-02 -8.6598949D-02 -6.3615434D-02
-1.1015182D-01  4.6427015D-02 -1.7080176D-02 -1.5285615D-01 -6.9244145D-03  8.9990091D-02 -1.2033750D-02 -3.4320644D-02
 3.9352689D-02  3.4440735D-02  7.7147508D-03  1.6478777D-02 -2.1297589D-02 -1.8073304D-02 -1.7076717D-02  5.7847038D-02
-7.5643353D-02  2.7431726D-04 -8.5300586D-02  5.9509332D-03  1.1816767D-02 -1.4987282D-02  1.3824003D-01 -8.0475278D-02
VECTOR 18       1.4053465D-01 -1.9007727D-01
-3.2396146D-02 -1.3991218D-01 -7.2211701D-02 -4.3667896D-01  2.7834876D-01  9.3915900D-03 -1.8622371D-01 -1.6669806D-01
 1.1977867D-02  2.3352156D-02 -1.4291630D-01 -2.7879816D-02  8.3363581D-02  1.0539480D-02  6.0565458D-02 -6.2769123D-02
-3.9300143D-03  3.8350355D-02  1.4064525D-01 -9.2061126D-02  4.2467511D-02 -5.4232856D-02 -1.8603631D-02  5.4513082D-02
-7.2428261D-03 -7.6675464D-02 -2.4588035D-01  4.7679618D-02  3.4540351D-01 -5.3545149D-02  4.8376366D-02  1.7966044D-01
 1.1391057D-01  2.0480746D-01 -3.7535440D-02 -1.6472092D-01 -2.3586278D-01 -2.2990193D-01 -4.3585959D-02 -3.2544158D-02
 3.7152390D-02 -2.2552857D-01  1.2741914D-02  4.0068007D-02  8.4196304D-02  4.2984433D-03 -6.5348952D-02  1.6618473D-01
VECTOR 19      -3.8705760D-01  1.2255022D-01
 2.4355394D-02 -3.4462166D-01 -1.5287899D-01 -1.1860472D-01  2.9984379D-01  4.5393653D-02 -1.3409079D-01  1.0618023D-01
-1.2925412D-01  1.9487319D-01  1.2157030D-01 -2.4335384D-02 -3.1431783D-01  1.6006698D-01  2.1187026D-01  3.9063962D-02
 1.3154622D-01 -2.0050707D-02  1.2047782D-01  1.1211794D-01  1.4223566D-01 -1.3900604D-01 -3.1501634D-02  6.2191680D-03
-1.4341674D-01 -9.1425383D-02  3.0813595D-02 -1.7735563D-01 -2.4259809D-01  1.0172947D-01  1.6927350D-02 -6.5894632D-02
 6.7470268D-02  9.4470905D-02 -7.7597795D-02  1.5623695D-02  7.1665809D-02  3.0957264D-03 -8.7526549D-02  9.7050972D-02
-3.5434012D-02  3.3947109D-02 -1.6988054D-03 -7.6029329D-02  8.1344641D-02 -1.3466621D-01  1.6117751D-01 -4.2180226D-02
VECTOR 20       7.1417565D-02 -2.5361725D-01
-1.4276823D-02  7.7903047D-02  9.3178331D-03 -6.1481624D-02  4.8557584D-01  1.2599019D-01 -4.4182436D-02  5.7920232D-02
 1.3716373D-01  1.3311612D-01  1.8173397D-02 -3.8378228D-02  4.6936019D-03 -2.8174999D-01 -1.5720288D-01  4.5022870D-02
-2.6123532D-01  2.6689294D-03 -1.5490166D-01  1.2047424D-01 -1.4350781D-02  6.4104310D-03 -5.0573137D-02  2.6889694D-02
-6.0746376D-04 -1.2450323D-01  2.4460301D-01  1.9750592D-01 -2.1267279D-02 -1.2369060D-01 -9.0047108D-02 -2.4834850D-01
 4.9440851D-02  5.5870397D-03  9.7651946D-02  1.8539138D-02  7.5010782D-02  8.1203700D-02 -1.2738932D-01  2.0729083D-02
 5.1554420D-02 -2.1991094D-01 -5.8604649D-02  2.1381484D-01  1.9832352D-01  1.9781141D-01  4.5268405D-02 -1.5345555D-02
VECTOR 21       2.0210475D-01 -1.7120665D-01
-6.8722965D-02  2.9467319D-01  2.9391273D-02 -1.3033819D-01 -1.2902420D-01  2.7498741D-02  9.3531788D-02  1.1602658D-01
-3.9731136D-02  3.1157230D-01  9.6341095D-02  5.7775567D-02 -3.0775571D-02  3.0042079D-01 -3.1959344D-01  2.9714461D-02
 2.2151359D-01 -4.1079364D-02 -2.3264928D-01  1.6836076D-01 -2.1027218D-02  2.5490353D-02  6.8684354D-02 -5.5782577D-02
-1.4794033D-01 -7.3889974D-03  7.9117141D-02 -1.6865495D-01 -2.6977767D-01 -1.3677783D-01 -2.4005199D-02 -4.0616209D-02
 8.5352275D-02  1.9843312D-01  1.6313402D-01 -1.3786585D-01 -1.0722219D-01 -1.8428782D-01  4.2029645D-02  7.3855121D-02
 [illegible]  [illegible]  [illegible]  1.2374551D-01  5.1841179D-02  3.2366899D-02 -1.3019037D-02  2.2895429D-02
```

D-3

3

```
8.6352275?D-02        1.8843312?D-01        1.0312462?D-01        1.2374551D-01
-5.3476774?D-02       1.6254654D-01         4.9439720D-02
VECTOR 22             -1.7693527D-01        7.1290896D-01         -8.0712763D-02
-1.9509743?D-02       2.6240794?D-01        7.8559072D-02         1.5819107D-02
9.7981905D-02         -4.6334728D-01        7.9843814D-02         2.6792509D-01
-4.7933904D-02        -1.6710272?D-02       -1.1159346D-01        -2.9559767D-02
-3.1171409D-02        3.9062453D-02         -4.1620032D-02        3.8273460D-02
1.9970139D-01         2.0703341D-01         1.4511265D-02         -8.2902952D-02
-3.0295733?D-02       7.1911455D-02         4.6016052D-03
VECTOR 23             3.2437138D-02         5.7500824D-01         -5.7377294D-03

5.1941100?D-02        3.2[illegible]D-02     -1.4049[illegible]D-02  2.2[illegible]D-02
2.3170691D-01         -1.4504366D-02        -5.9982675D-02        1.2812213D-01
3.2407794D-01         -4.4512950D-02        7.7544992D-02         -2.0066844D-02
-6.5199202D-02        1.1947639D-01         2.2333619D-02         1.9125565D-01
-1.9892145D-01        -3.6295189D-02        4.7041229D-03         -3.4946436D-02
-1.0778857D-01        -1.3956573D-01        1.3945191D-01         6.5799790D-02
-2.7147728?D-02       -2.2697397D-01        2.0849138D-01         -4.7966527D-02
8.4979309D-02         3.4323228D-02         2.8979561D-02         -1.2763536D-01

5.7008039?D-02        1.4099699D-01         7.4917340D-02         -2.0348031D-02
8.7470126?D-03        -3.2731994D-01        -2.6971649D-02        -4.0119962D-02
3.3500653D-02         -1.0281238D-01        -6.0420185D-02        7.0974802D-02
-2.0732833?D-01       -9.3303379D-02        -7.1027957D-02        -1.9746490D-01
1.5170610D-01         1.1224781D-01         1.6296194D-01         8.4382322D-02
-1.3490222D-01        1.6036890D-03         -7.7684603D-03
VECTOR 24             -5.2195753D-04        -2.1096983D-01        -1.0552096D-01
-3.5776359?D-02       -6.2232099?D-02       2.4583978D-02         1.9978528D-01
-1.5231642?D-01       7.4850190D-02         -2.6130663D-01        8.7411227D-02
1.5771799D-01         9.6840988D-02         1.6998408D-02         7.6387426D-02
-7.6511675D-02        1.6051335D-01         -1.4967379D-02        3.6539353D-02
-1.1918445D-01        -1.2483972D-01        -9.7017744D-02        -1.7207899D-01
-2.2609653?D-01       -6.2759369D-02        3.0487715D-02
VECTOR 25             -3.6573746D-02        -2.3220532D-01        -1.6872153D-01
-5.0707352D-03        -3.3030692D-02        1.2872866D-02         -1.6337200D-02
-1.1321395D-01        -3.7893669?D-01       1.2469871D-02         -8.0557631D-02
-1.6664733?D-02       -6.9251362D-02        -7.7891827D-02        -1.4777913D-01
-4.5437747?D-01       -5.7843164D-02        3.8181819D-02         -9.3795577D-02
-1.7232697D-01        -1.5923743D-01        1.2525562D-01         1.9711015D-01
-2.2244521D-01        -5.5321446?D-02       2.6322254D-02
VECTOR 26             -5.1333994D-02        1.5024294D-01         -1.2552259D-01
4.1393870?D-03        -5.7599443D-02        -4.0319909D-02        6.9608233D-02
-1.2572981D-01        -2.9185599D-01        -1.9655195D-01        -7.4446152D-04
1.6476657D-02         -1.8896610D-01        -3.9465469D-02        1.2393271D-01
1.3179896D-02         5.0447684D-02         5.4912161D-02         -5.9303952D-02
-1.3891883D-01        -5.3494235D-02        1.4716939D-01         3.5069787D-02
-1.2173533D-01        7.4002655D-02         3.7352390D-02
VECTOR 27             -5.0631225D-03        2.9221753D-01         -8.3977253D-02
4.9739041D-03         1.7265133D-01         -1.3928472D-02        -9.3468813D-02
4.3600797?D-02        -7.1906543D-02        -2.9540569D-01        1.0449121D-01
1.3715357D-01         2.4402110D-01         -3.6546378D-02        -1.7431793D-01
-6.4629901D-03        -2.6421861D-01        -1.4685589D-01        3.0339212D-02
-1.0743227D-01        -2.3332515D-01        -2.2882281D-01        6.8246016D-02
1.1663960D-01         -2.7610393D-01        7.5055800D-02
VECTOR 28             -1.5602894D-02        8.3927444D-02         1.0044625D-01
-1.4315319D-02        3.5603590D-02         3.3235754D-02         2.0028225D-02
-1.1797915?D-02       -3.1549475D-02        -2.1426531D-01        -8.1381597D-02
7.7592655D-02         1.4973672?D-01        -2.8629352D-02        4.8600857D-02
-5.5507330?D-02       -2.9011174D-01        2.3707906D-01         4.2437047D-02
-1.3019264?D-01       -1.1222541D-02        1.9260619D-01         -2.2330973D-01
7.5970093?D-02        2.0200564D-01         -3.6300351D-02
VECTOR 29             -7.6622174?D-02       -9.8955977D-02        -5.5367591D-02
-2.7302557?D-02       -1.0356663D-02        -2.3643931D-02        -2.3402651D-01
4.5206760?D-02        -5.5869159D-02        5.6420325D-02         1.4788746D-01
4.5320192?D-02        -2.4609763?D-01       3.4494597D-01         -3.3907144D-01
1.3129124?D-01        4.4374794D-02         3.6416348D-02         -3.3179891D-02
-2.7149657?D-02       -1.0073924D-01        9.5676851D-02         1.3538984D-01
1.3254443?D-01        -1.4772033?D-01       -2.9169168D-02

1.6592001D-01         6.5851017D-02         6.5337343D-02         -3.4012022D-02
-1.4969004D-02        1.9408697D-02         -4.4787991D-02        3.6202341D-02
2.5820534D-01         8.3283611D-02         -1.9878609D-02        4.9058297D-02
2.0554086D-01         2.1275357D-01         3.5459107D-02         -1.7499876D-01
6.1320303D-02         1.9097418D-01         1.3912014D-01         -9.4882295D-02
2.2654349D-01         -4.1449054D-01        -9.7117377D-03        -2.7573354D-01
2.0040508D-01         6.0021572D-02         -1.3534567D-01        7.6898462D-02
3.2358607D-02         -3.8641883D-02        2.3967232D-01         4.3746607D-02
-1.4280755D-01        -6.5399780D-02        -7.6246773D-02        -2.1354278D-02
1.7805293D-01         2.2750618D-01         3.0254548D-02         -1.3950199D-02
-1.4882332D-01        -3.9252336D-02        1.5975756D-01         -1.0325476D-01
-7.4727708D-02        1.8729467D-01         7.8468472D-02         -4.9175684D-02
-2.6526320D-02        5.3079434D-02         1.1398861D-01         -6.4894032D-02
5.6543555D-02         -1.7277954D-02        -7.2308832D-02        9.1610601D-02
-2.8969632D-02        -1.0163266D-01        3.5613225D-02         2.2078590D-01
2.6801160D-01         1.6426836D-01         1.3124448D-01         -1.3616778D-01
1.3162920D-01         1.1137887D-03         5.0325906D-03         1.7188439D-01
7.5954130D-02         8.8176661D-02         1.9682978D-01         -1.9065107D-01
-1.0078108D-01        9.6232171D-02         1.9625354D-01         -4.1591597D-03
-5.5567273D-02        -2.8964589D-02        6.7173527D-02         -2.1857976D-01
-2.1942744D-01        -8.9301104D-03        3.6669914D-02         -9.0699155D-02
-2.5890323D-01        -1.0033451D-01        -2.9485067D-01        1.4818300D-01
-1.3393891D-02        4.2986611D-01         -1.3053666D-01        -1.8468215D-01
1.9431834D-01         3.1475998D-02         -9.6241944D-03        -1.9743415D-01
-1.9430580D-01        8.9203592D-02         -1.4290171D-01        2.6560970D-03
-1.0176806D-01        1.0293573D-01         -7.3725618D-02        7.2734387D-02
-1.3226983D-01        -7.0424927D-02        2.5786299D-02         1.4410663D-01
-8.2197978D-02        -1.7512476D-01        -1.1383077D-01        -1.9575162D-01
8.6719440D-02         -1.3687658D-02        -1.0459116D-01        2.3054187D-01
-8.5031696D-02        1.8158031D-02         -2.4683923D-01        -3.2510555D-01
7.3398139D-02         5.3647327D-02         3.1573642D-02         -7.0712937D-02
7.4025494D-02         3.1252932D-02         8.0240420D-02         -3.1369881D-02
8.9223266D-02         2.0428036D-02         1.0826789D-02         -3.4516029D-01
7.1015933D-02         -3.2315688D-02        1.0942740D-01         2.0552729D-01
-2.1136350D-01        -3.3535466D-01        -1.6959812D-01        -8.0687957D-02
-2.2521789D-02        -1.4905284D-01        3.7539341D-02         -2.0096612D-02
5.4551662D-01         1.1560542D-01         -1.6090327D-03        1.7305277D-02
3.2000162D-02         -1.8790850D-02        -2.2283393D-01        -1.1866562D-01
8.2299034D-03         1.0525596D-01         3.1574771D-02         -5.8169866D-02
1.8153159D-02         -2.1514577D-03        -3.3741110D-02        1.0193958D-01
1.9490277D-01         8.2002690D-02         -4.2591641D-02        -1.3884042D-02
```

APPENDIX E

EQUATIONS FOR UPDATING THE FISHER DISCRIMINANT

Desired Algorithm:

$$W_K = A_i^c \; V_{i,K}^T + C \tag{1}$$

To update algorithm, desire to use new learning histories, $V_{i,K}^L$, to compute $A_i^{c'}$

Step 1: Transform Learning Data, $V_{i,K}^L$, to optimum space

$$Y_{\ell,K}^L = H_{i\ell}\left(V_{iK} - \frac{1}{K_{MAX}} \sum_{K=1}^{K_{MAX}} V_{iK}\right) \tag{2}$$

Where:

$Y_{\ell,K}^L$ = Components of learning history "K" in optimum space

$H_{i\ell}$ = Transformation matrix derived by ADAPT representation programs and supplied by Avco on "H-Tape"

i = 1, 2, ... No. of Meas. used

ℓ = 1, 2, ... Dimensionality of Optimum Space

K = 1, 2, ... No. of Learning Cases

Step 2: Derivation of Fisher Discriminant, A_ℓ^c, in optimum space

$$A_\ell^c = D_{\ell\ell}^{-1}\left[W_\ell^{c1} - W_\ell^{c2}\right]$$

$$D_{\ell\ell} = \rho\, B_{\ell\ell}^{c1} + (1-\rho)\, B_{\ell\ell}^{c2}$$

ρ = Input = Fisher Weight Parameter

$$B_{\ell\ell}^{c_j} = \begin{bmatrix} Y_{1,1}^{Lc_j} - W_1^{c_j} & \ldots\ldots & Y_{1,M_j}^{Lc_j} - W_1^{c_j} \\ \vdots & & \\ Y_{\ell max,1}^{Lc_j} - W_{\ell max}^{c_j} & \ldots\ldots & Y_{\ell max,M_j}^{Lc_j} - W_{\ell max}^{c_j} \end{bmatrix}^{2}$$

Note: $B_{\ell,\ell}^{c_j}$ = VAR-COVAR Matrix of Class j in optimum Space!

$Y_{\ell,K}^{Lc_j}$ = $Y_{\ell,K}^{L}$ Assigned to Class j

j = 1, 2

M_j = No. of cases in class j

$$W_\ell^{c_j} = \frac{1}{M_j} \sum_{K=K_0}^{M_j} Y_{\ell K}^{L}$$

Note: K_o = (1 for j = 1
(M1 + 1 for J = 2

Step 3: Transform Fisher Discriminant back to data space:

$$A_i^c = H_{i\ell} A_\ell^c$$

Step 4: Find C = -TH . . . . . Where

TH = Fisher Threshold Determined as Described in Section 2 of Appendix C