INTRUSION DETECTION VIA AN ADAPTIVE DIGITAL PREDICTOR CHI-SQUARE TEST COMBINATION

by

RADEN DJAFAR SUMANTRI

B. S., Tokai University, 1976

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Electrical Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1979

Approved by:

Donald R. Hummels

Major Professor

Table of Contents

List of Figures

Page No.

Page No.

List of Tables

Chapter 1

INTRODUCTION

The problem considered is that of detecting an intruder by using an adaptive digital predictor (ADP) Chi-square test combination. Figure 1.1 shows a block diagram of the approach. The inputs used in this experiment were obtained at 128 sps via sensors which are in the form of buried cables such sensors responsed not only to the presence of the intruder, but also to a wide variety of other stimuli which result in a poor signal-to-noise ratio.

The algorithms for the adaptive digital predictor used in this experiment are Widrow's least mean square (LMS) algorithm, the weighted sequential regression (WSER) algorithm, and the long term sequential regression (LTSER) algorithm. The main function of the ADP is to decorrelate the ambient noise portion of the input signal [1].

The general input signal consists of two components, (1) ambient noise, and (2) an intruder, if any. Since the response time of the ADP is not fast enough to cancel an intruder, the over all problem is reduced to that of determining the presence or the absence of an intruder in decorrelated noise.

The approach taken here is to test the hypothesis that output samples from the ADP are Gaussian random variables using a Chi-square test. It was found that for relatively short sequences of the output samples, the hypothesis is true with good confidence level, unless an intruder is present. It was further found that the confidence level for the Chi-square test is a possible detector output signal. An advantage of this approach is that its performance depends primarily on the statistics of the signals received and thus, variations in

THIS BOOK CONTAINS NUMEROUS PAGES WITH DIAGRAMS THAT ARE CROOKED COMPARED TO THE REST OF THE INFORMATION ON THE PAGE.

THIS IS AS RECEIVED FROM CUSTOMER.

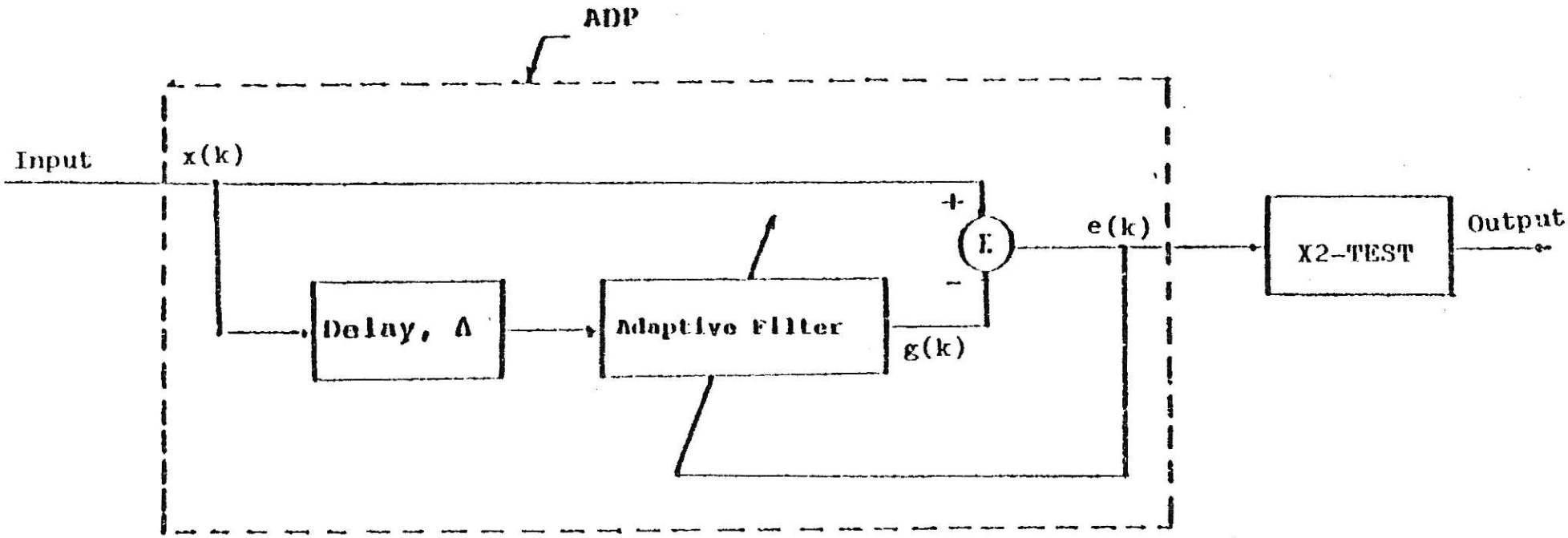


Fig. 1.1 A predictor - X^2-TEST comb. intrusion detection model

amplitude are less likely to result in false alarm.

Various aspects of the development of the Chi-square test procedure and ADPs are considered in Chapters 2 and 3 of this report. In Chapter 4 the experimental results are presented. Since the predictor and moving average filter (MAF) combination shown in Figure 1.2 has been studied previously, the experimental results include comparisons with that approach.

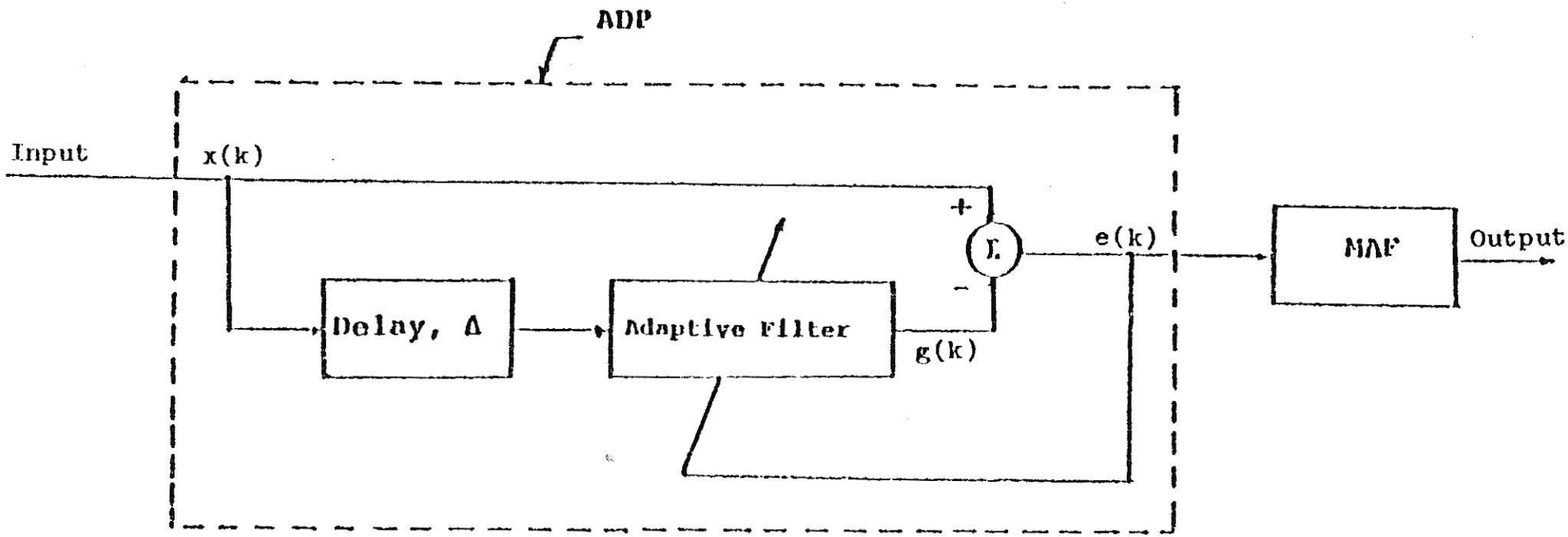


Fig. 1.2 A predictor - MAF Comb. intrusion detection model

Chapter 2

THE CHI-SQUARE TEST AND ITS APPLICATION TO INTRUSION DETECTION

A. CHI-SQUARE TEST

The Chi-square test was first introduced by Karl Pearson in early 19th century. Chi-square tests (χ^2 tests) are routinely used as a means of testing the agreement between observed and expected occurences of a random experiment. The element common to these tests is the comparison of the number of outcomes actually observed to fall into any number of classes with the number which upon some hypothesis is expected. If E(i) is the number expected, and F(i) is the number observed in i-th class then

$$\chi^2 = \sum_{i=1}^{K} \frac{F(i) - E(i)^2}{E(i)} \qquad (1)$$

where K is the number of classes, and χ^2 is the value of the Chi-square test for K classes. It is clear that the more closely the observed number agree with those expected, the smaller will be the value of χ^2. Also, the value of χ^2 is always positive or zero. The probability density function of χ^2 is well approximated by the curve [2]

$$Y(G,X) = c\, X^{\frac{G-2}{2}}\, e^{-X^2/2} \qquad (2)$$

$$G = K - P - 1$$

where P is the number of parameters of the test estimated from the samples, G is the number of degrees of freedom and c is a constant. The value of c depends on the value of G and is determined in such a way that the total area under the probability curves shown in Figure 2.1 is equal to 1. These curves show the probability density function of variable χ^2 plotted as Y(X,G) versus X for different values of G. The variable X corresponds to the value of χ^2. The Table 2.1 gives the

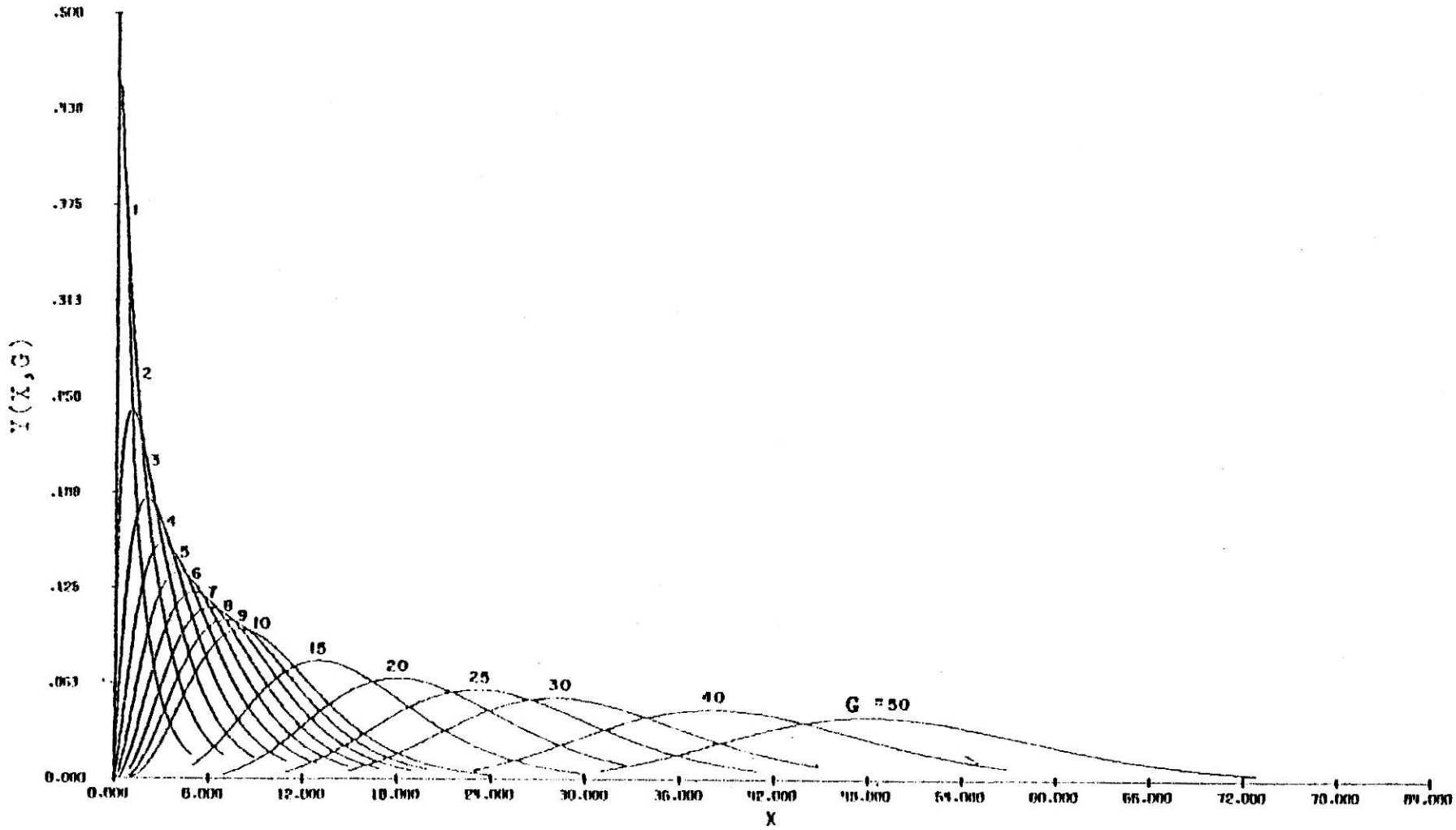


Fig. 2 . 1 The probability density function of the Chi-square distribution plotted as Y(X,G) versus X for different value of G, the number of independent degree of freedom, and X = the value of X^2 .

Table 2 . 1 Chi-square distribution.
A denote the right tail area for the values of χ^2 given below. G denotes the number of degrees of freedom (df).

G or df	A = 0.99	A = 0.98	A = 0.95	A = 0.90	A = 0.80	A = 0.70	A = 0.50	= 0.30	A = 0.20	A = 0.10	A = 0.05	A = 0.02	A = 0.01	A = 0.001
1	.00016	.00063	.0039	.016	.064	.15	.46	1.07	1.64	2.71	3.84	5.41	6.64	10.83
2	.02	.04	.10	.21	.45	.71	1.39	2.41	3.22	4.60	5.99	7.82	9.21	13.82
3	.12	.18	.35	.58	1.00	1.42	2.37	3.66	4.64	6.25	7.82	9.84	11.34	16.27
4	.30	.43	.71	1.06	1.65	2.20	3.36	4.88	5.99	7.78	9.49	11.67	13.28	18.46
5	.55	.75	1.14	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.07	13.39	15.09	20.52
6	.87	1.13	1.64	2.20	3.07	3.83	5.35	7.23	8.56	10.64	12.59	15.03	16.81	22.46
7	1.24	1.56	2.17	2.83	3.82	4.67	6.35	8.38	9.80	12.02	14.07	16.62	18.48	24.32
8	1.65	2.03	2.73	3.49	4.59	5.53	7.34	9.52	11.03	13.36	15.51	18.17	20.09	26.12
9	2.09	2.53	3.32	4.17	5.38	6.39	8.34	10.66	12.24	14.68	16.92	19.68	21.67	27.88
10	2.56	3.06	3.94	4.86	6.18	7.27	9.34	11.78	13.44	15.99	18.31	21.16	23.21	29.59
11	3.05	3.61	4.58	5.58	6.99	8.15	10.34	12.90	14.63	17.28	19.68	22.62	24.72	31.26
12	3.57	4.18	5.23	6.30	7.81	9.03	11.34	14.01	15.81	18.55	21.03	24.05	26.22	32.91
13	4.11	4.76	5.89	7.04	8.63	9.93	12.34	15.12	16.98	19.81	22.36	25.47	27.69	34.53
14	4.66	5.37	6.57	7.79	9.47	10.82	13.34	16.22	18.15	21.06	23.68	26.87	29.14	36.12
15	5.23	5.98	7.26	8.55	10.31	11.72	14.34	17.32	19.31	22.31	25.00	28.26	30.58	37.70
16	5.81	6.61	7.96	9.31	11.15	12.62	15.34	18.42	20.46	23.54	26.30	29.63	32.00	39.25
17	6.41	7.26	8.67	10.08	12.00	13.53	16.34	19.51	21.62	24.77	27.59	31.00	33.41	40.79
18	7.02	7.91	9.39	10.86	12.86	14.44	17.34	20.60	22.76	25.99	28.87	32.35	34.80	42.31
19	7.63	8.57	10.12	11.65	13.72	15.35	18.34	21.69	23.90	27.20	30.14	33.69	36.19	43.82
20	8.26	9.24	10.85	12.44	14.58	16.27	19.34	22.78	25.04	28.41	31.41	35.02	37.57	45.32
21	8.90	9.92	11.59	13.24	15.44	17.18	20.34	23.86	26.17	29.62	32.67	36.34	38.93	46.80
22	9.54	10.60	12.34	14.04	16.31	18.10	21.34	24.94	27.30	30.81	33.92	37.66	40.29	48.27
23	10.20	11.29	13.09	14.85	17.19	19.02	22.34	26.02	28.43	32.01	35.17	38.97	41.64	49.73
24	10.86	11.99	13.85	15.66	18.06	19.94	23.34	27.10	29.55	33.20	36.42	40.27	42.98	51.18
25	11.52	12.70	14.61	16.47	18.94	20.87	24.34	28.17	30.68	34.38	37.65	41.57	44.31	52.62
26	12.20	13.41	15.38	17.29	19.82	21.79	25.34	29.25	31.80	35.56	38.88	42.86	45.64	54.05
27	12.88	14.12	16.15	18.11	20.70	22.72	26.34	30.32	32.91	36.74	40.11	44.14	46.96	55.48
28	13.56	14.85	16.93	18.94	21.59	23.65	27.34	31.39	34.03	37.92	41.34	45.42	48.28	56.89
29	14.26	15.57	17.71	19.77	22.48	24.58	28.34	32.46	35.14	39.09	42.56	46.69	49.59	58.30
30	14.95	16.31	18.49	20.60	23.36	25.51	29.34	33.53	36.25	40.26	43.77	47.96	50.89	59.70

right tail area A for given value of χ^2 and the number of degrees of freedoms G. The information contained in Figure 2.1 and Table 2.1 are available from several sources for example see reference [2].

In using the Chi-square test, the expected value E(i) in Equation (1) for each class must be at least 5 [2]. If the value of right tail area A in Table 2.1 for a given value of χ^2 is between 0.1 and 1.0 there is certainly no reason to suspect the hypothesis tested, if it is below 0.02 it is strongly indicated that the hypothesis is not true.

B. APPLICATION OF THE CHI-SQUARE TEST TO INTRUSION DETECTION

It has been shown that the output of an ADP tends to be band limited white noise [1]. This property leads one to conjecture that the output samples may also have a Gaussian distribution. One of the objectives of the work reported here was to test this hypothesis using a Chi-square test. As will be shown later, the test confirms this hypothesis for relatively short sequences of the ADP output samples when there is no intruder present. It was subsequently found that the presence of an intruder leads to a negative result from the Chi-square test in most cases. Experiments were then performed to evaluate the use of the Chi-square test as an algorithm for an intrusion detection system.

The procedure for carrying out the χ^2 test is as follows

1. Divide the data signal (output of the predictor) into M blocks, where each block contains N data points.
2. Calculate the mean value m, and standard deviation σ of the data in each of the blocks using

$$m = \frac{1}{N} \sum_{i=1}^{N} Y(i) \tag{3}$$

$$\sigma = \left[\frac{\sum_{i=1}^{N} (Y(i) - m)^2}{N} \right]^{1/2}, \quad i=1, 2, 3, \ldots, N \tag{4}$$

where Y(i) is the value of the data at i-th point. The data samples in each block are then put in classes.

3. Compute the boundary points Z(i) of the classes using

$$Z(1) = -\text{ infinity}$$

$$Z(i) = m - .5(K/2+1-i)\sigma, \quad 1 = 2, 3, 4, \ldots, K \tag{5}$$

$$Z(K+1) = +\text{ infinity}$$

where K is the number of classes. Then compute the observed values F(i) for each class. Since two parameters, mean value and standard deviation were estimated from the data at hand, the number of degrees of freedom G is

$$G = K - 1 - P = K - 3 \tag{6}$$

As an example, Figure 2.2 shows the frequency reading F(i) for each class using a block of one of the signals processed by an LMS predictor.

4. Normalize the boundary points Z(i). The normalized boundary points $Z_n(i)$ are determined by

$$Z_n(i) = \frac{Z(i) - m}{\sigma} \tag{7}$$

Figure 2.3 shows the normalized boundary points, and classes for the same data block used in Figure 2.2.

5. Calculate the expected value E(i) for each class using an approximation [4] to Gaussian distribution function $P_x(i)$

$$P_x(i) = .5 + .5\ [1 - \text{Exp.}(-2Z_n(i)/\pi)] \tag{8}$$

where $Z_n(i) \geq 0$

$$E(i) = N\ P_x(i+1) - P_x(i) \tag{9}$$

which i = K/2+1, K/2+2, K/2+3, ..., K+1, and K the number of classes. Because the Gaussian distribution is an even function, we can then use

$$E(i) = E(K+1-i) \quad , i = 1, 2, 3, \ldots, K/2 \tag{10}$$

to find E(i) for those classes which correspond to negative data points. Figure 2.4 shows the expected value E(i) of each class and Figure 2.5 shows the differences between E(i) F(i) of each class of the same data block used in Figure 2.2

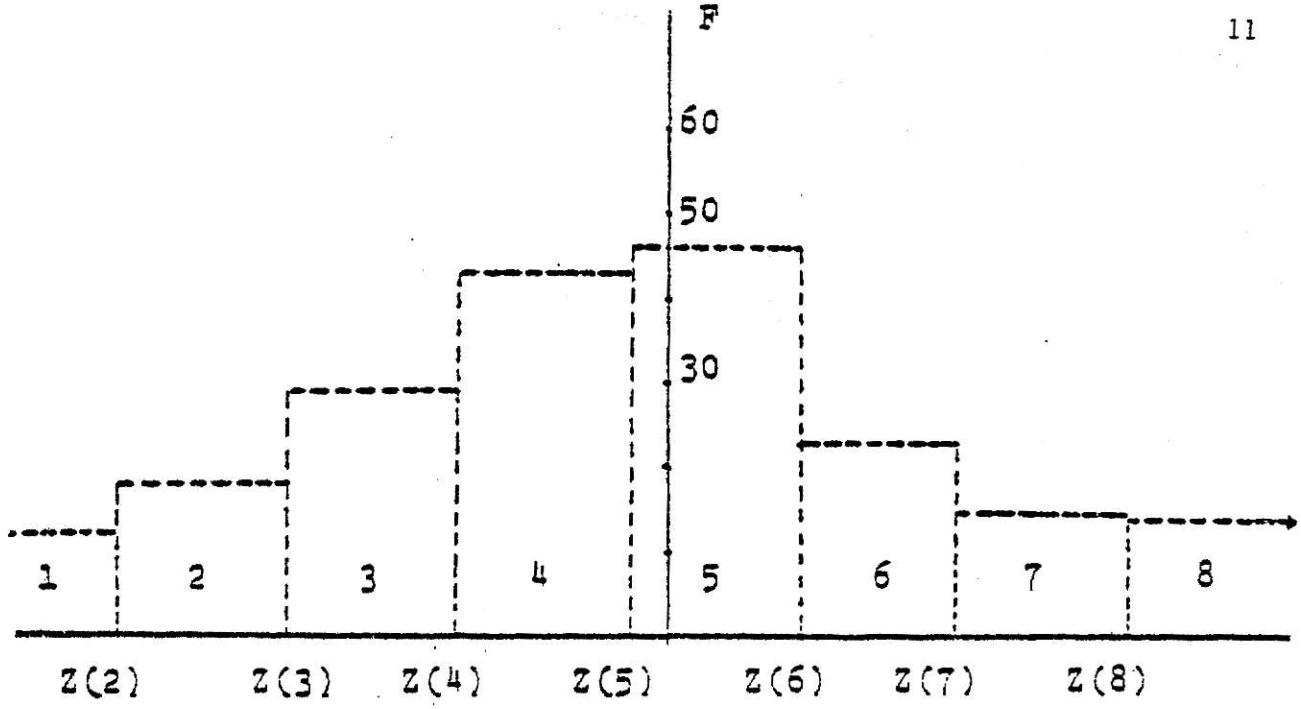


Fig. 2 . 2 The observed values F of each class of a block of the input file 7Bo9B6.FP.

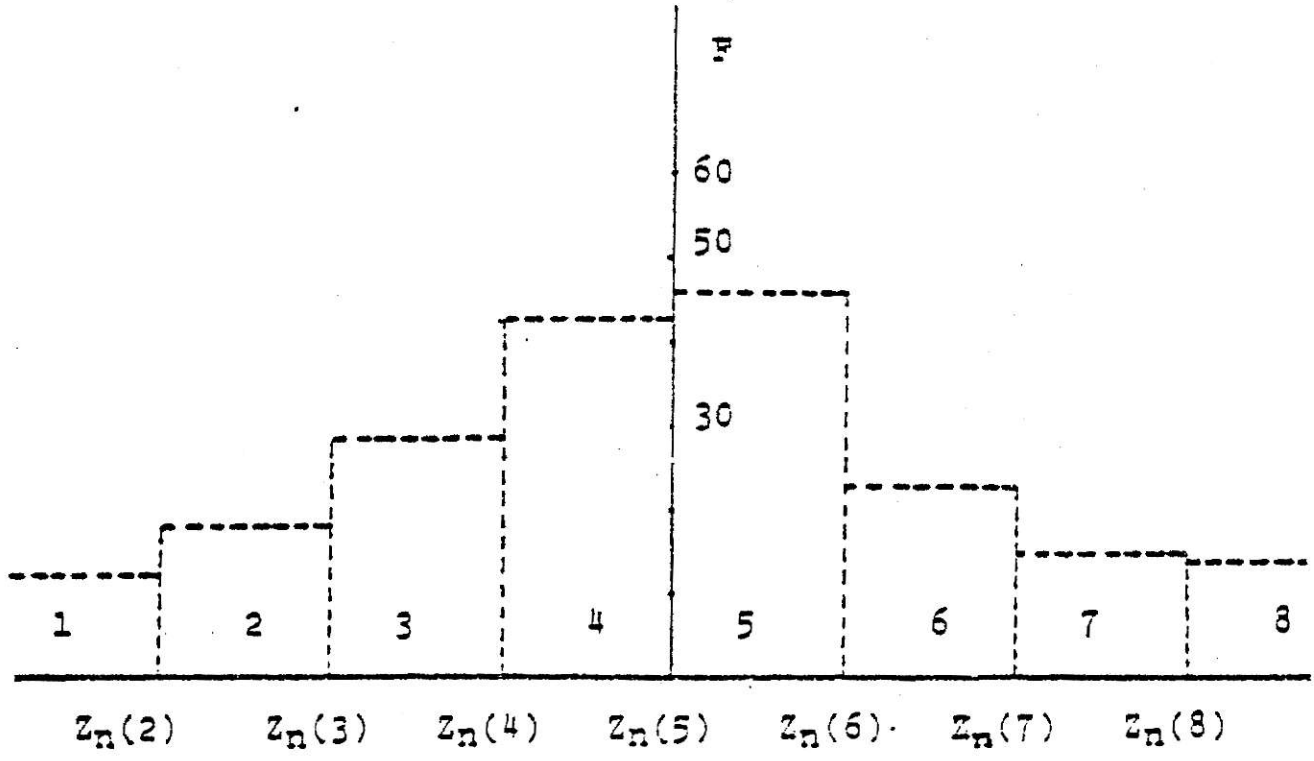


Fig. 2 . 3 The normalized boundary points $Z_n(i)$ of Fig. 2.2 .

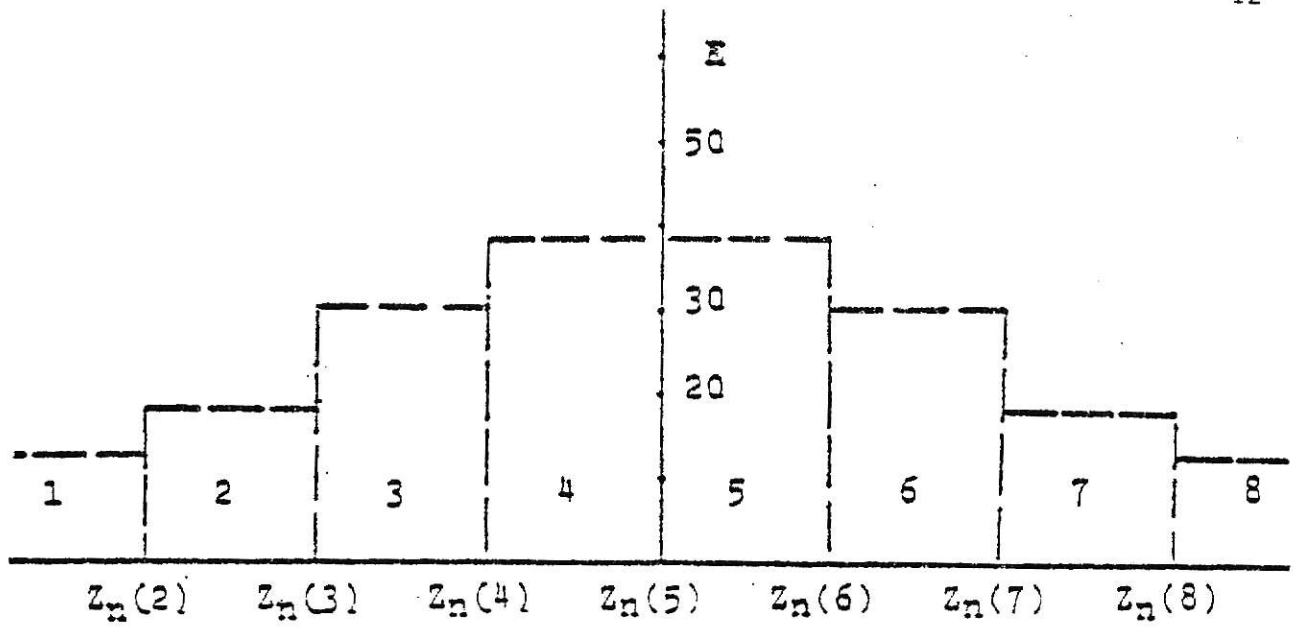


Fig. 2 . 4 The expected values E of each class of the same input file as in Fig. 2.2.

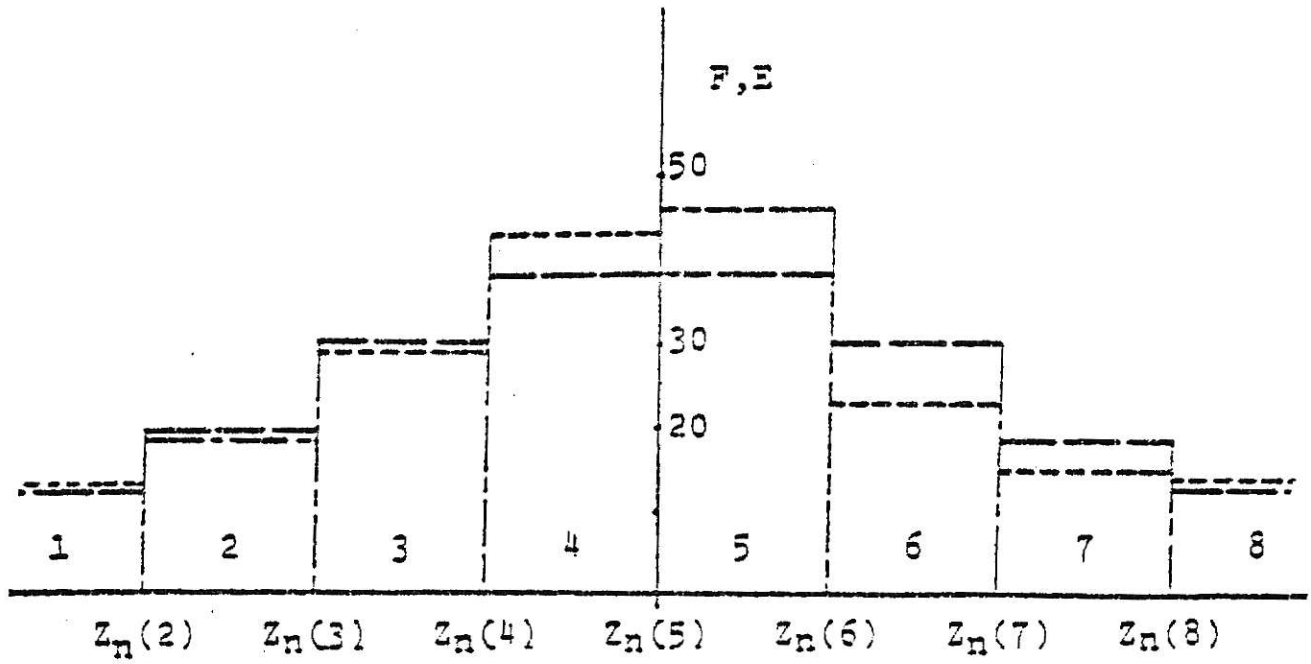


Fig. 2 . 5 The over all values of E and F of the same input file as in Fig. 2.2 .

6. Calculate the over all value of χ^2 for the block by summing the value of $\chi^2(i)$ of each class where

$$\chi^2(i) = \frac{F(i) - E(i)}{E(i)}^2 \quad , i=1, 2, 3, \ldots, K \tag{11}$$

then

$$\chi^2 = \sum_{i=1}^{K} \chi^2(i), \quad , i=1, 2, 3, \ldots, K \tag{12}$$

where K is the number of classes. Then using the χ^2 distribution function subroutine (CDTR) [5], calculate the level of confidence of that particular block. The CDTR subroutine is a program to calculate the confidence level of given values of χ^2 and G.

The computations outlined in steps 1 through 6 are performed on every block of the data samples, the result being a set of confidence levels which are a measure of the Gaussian character of each of the data blocks. Since the data are time samples of a random signal, graphing the confidence level versus time will show how the Gaussian nature of the ADP output signal varies with time. Because of the large variations in this signal, it has been found more convenient to plot the logarithm of this signal. In all of the experiments described in later sections of this report, the outputs of the χ^2 test algorithm are shown in DB.

Chapter 3

ADAPTIVE DIGITAL PREDICTOR

A predictor is a device which uses past input samples to estimate the current input sample. The system considered is as shown in Figure 3.1, and the estimate of the current sample g(k) is computed as follows:

$$g(k) = a(1,k)\, x(k-\Delta) + a(2,k)\, x(k-1-\Delta) + \ldots + a(M,k)\, x(k+1-M-\Delta) \qquad (13)$$

where a(i,k) is the i-th filter coefficient (weight) at time k, and x(k), x(k-1), x(k-2), ..., x(k-M) is the input sequence to the digital filter, and Δ denotes a delay time. It is clear from the above equation that the only parameters needed to calculate g(k) are the filter coefficients a(i,k). In the experiments described in this report, we use three different adaptive algorithms to compute the coefficients; (1) Widrow's LMS, (2) LTSER, and (3) WSER. In the following sections the equations defining the algorithms are presented. Detailed derivations can be found elsewhere [6, 7, 8].

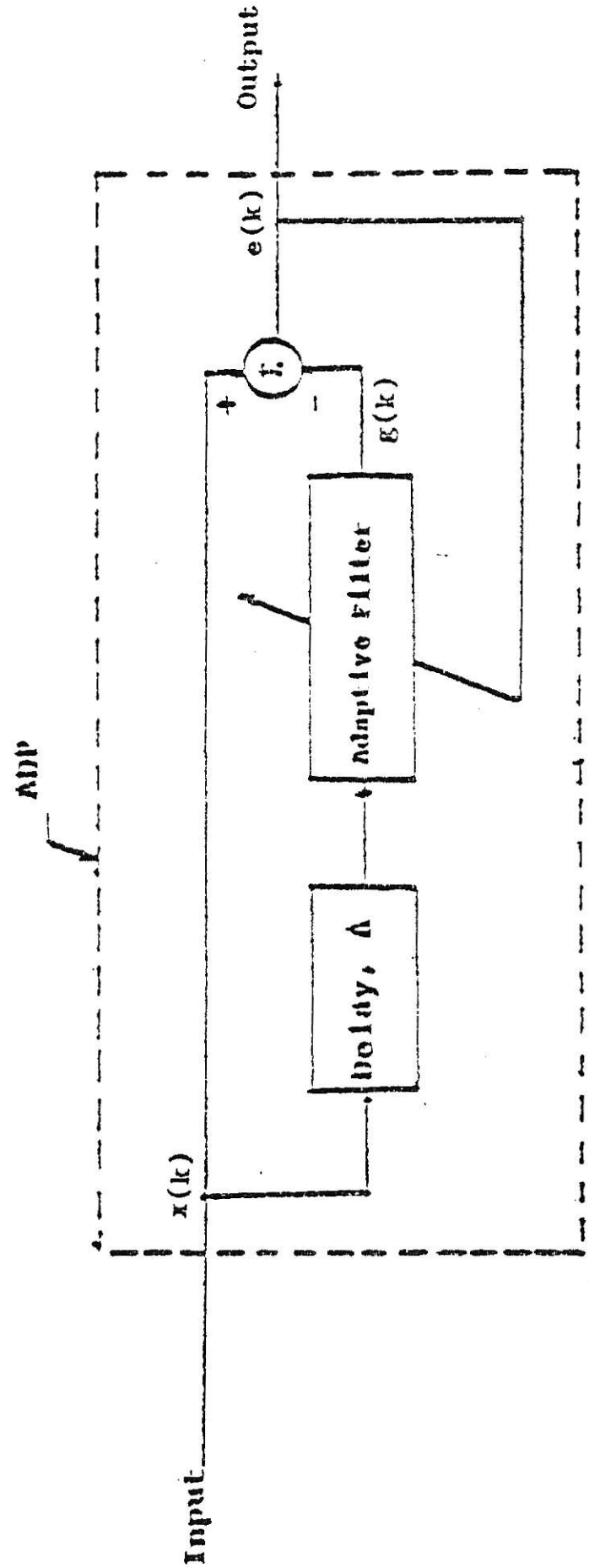


Fig. 3.1 Adaptive digital predictor configuration.

A. WIDROW'S LEAST MEAN SQUARE (LMS) ALGORITHM

The LMS algorithm is the simplest of the three considered. The coefficients a(i,k) are simply updated by the relation [7]

$$a(i,k+1) = a(i,k) + v\, x(k-i-\,)\, e(k)$$

$$, \; i=1, \; 2, \; 3, \; \ldots, \; M \qquad (14)$$

where v is convergence parameter, x(k) is the input to the system at time k, and e(k) is the output at time k, The main advantage of the LMS algorithm is the computation advantage gained by its simplicity.

B. WEIGHTED SEQUENTIAL REGRESSION (WSER) ALGORITHM

Let A_k be a vector with the filter coefficients arranged as it components

$$A_k^T = [a(0,k)\ a(1,k)\ a(2,k)\ \ldots\ a(M,k)]$$

then the desired estimate of the filter coefficient vector A_k can be computed using the WSER algorithm as follows [7]:

$$A_k = A_{k-1} + \delta P_k^{-1} [A_0 - A_{k-1}] + P_k^{-1} X_k\, e(k) \qquad (15)$$

where $X_k^T = [x(k)\ x(k-1)\ x(k-2)\ \ldots\ x(k-M)]$

$\delta = u(1-q)$

The parameter u and q are constants, x(k) and e(k) are the input and the output of the system, respectively. And A_0 represents the initial or priori filter coefficients. The (MxM) matrix P_k^{-1} can be updated recursively using the following steps [7]

1. $Q_0^{-1} = 1/q\ P_{k-1}^{-1} - \frac{1}{q\gamma_0} P_{k-1}^{-1} X_k X_k^T P_{k-1}^{-1}$

   where $\gamma_0 = q + X_k^T P_{k-1}^{-1} X_k$

2. $Q_1^{-1} = Q_0^{-1} - \frac{\delta}{\gamma_1} Q_0^{-1} e_1 e_1^T Q_0^{-1}$

   where $\gamma_1 = 1+\delta\ e_1^T Q_0^{-1} e_1$

3. $Q_2^{-1} = Q_1^{-1} + \frac{\delta}{\gamma_2} Q_1^{-1} e_2 e_2^T Q_1^{-1}$

   where $\gamma_2 = 1 + \delta\, e_2^T Q_1^{-1} e_2$

The above procedure is continued, until we obtain

$$Q_n^{-1} = Q_{N-1}^{-1} - \frac{\delta}{\gamma_N} Q_{N-1}^{-1} e_N e_N^T Q_{N-1}^{-1}$$

where $\gamma_N = 1+\delta\ e_N^T Q_{N-1} e_N$

$$P_k^{-1} = Q_{N-1}^{-1} - \frac{\delta}{\gamma_N} Q_{N-1}^{-1} e_N e_N^T Q_{N-1}^{-1} \tag{16}$$

where $N = M + 1$, and e_i is a unit vector whose components are zero, except for the i-th component, which is equal to 1. For the computation, the initial value P_0^{-1} is needed. For all cases considered here, $P_0^{-1} = I$ was used. Where I is the (MxM) identity matrix.

C. LONG TERM SEQUENTIAL REGRESSION (LTSER) ALGORITHM

The desired estimate of the filter coefficient vector A_k is computed as follows [8]:

$$A_k = A_{k-1} + P_k^{-1} X_k \, e(k) \tag{17}$$

where e(k) is the output of the system at time k, and

$$X_k^T = [x(k) \; x(k-1) \; x(k-2) \; \ldots \; x(k-M)]$$

The (MxM) matrix P_k^{-1} is updated recursively using the following relation [8]:

$$P_k^{-1} = P_k^{-1} - 1/\delta \; P_{k-1}^{-1} \; X_k \; X_k^T \; P_{k-1}^{-1} \tag{18}$$

where $\delta = 1 + X_k^T P_{k-1}^{-1} X_k$ is a scalar. The initial value P_0^{-1} in Equation (18) is chosen as $P_0^{-1} = cI$, where c is a constant, and I is the (MxM) identity matrix. The desired LTSER predictor is defined by Equation (17), where the matrix P_k^{-1} is updated using Equation (18).

Chapter 4

EXPERIMENTAL RESULTS

A. THE INFLUENCE OF BLOCK SIZE ON PERFORMANCE

The objective of the experiment described in this section was to determine the effect of block size on the ability to detect an intruder and to choose a good block size for the experiments described in subsequent sections of this report. The approach taken was to select two representative data files and process these data files using the Chi-square test approach previously described with several different choices for block size. The results were displayed in graphic form and a subjective judgement was made as to the block size which results in the best performance as an intrusion detector. The data files used in this experiment are shown in Figure 4.1. These two files were selected because they are some what different with respect to the character of the background noise but are both typical of the types of data encountered in practice. The file shown in Figure 4.1(a) is a case where the intruder is a walker, and the background is caterpilar noise. The data file in Figure 4.1(b) is a case where the intruder is a nonmagnetic walker and the background is road grader noise.

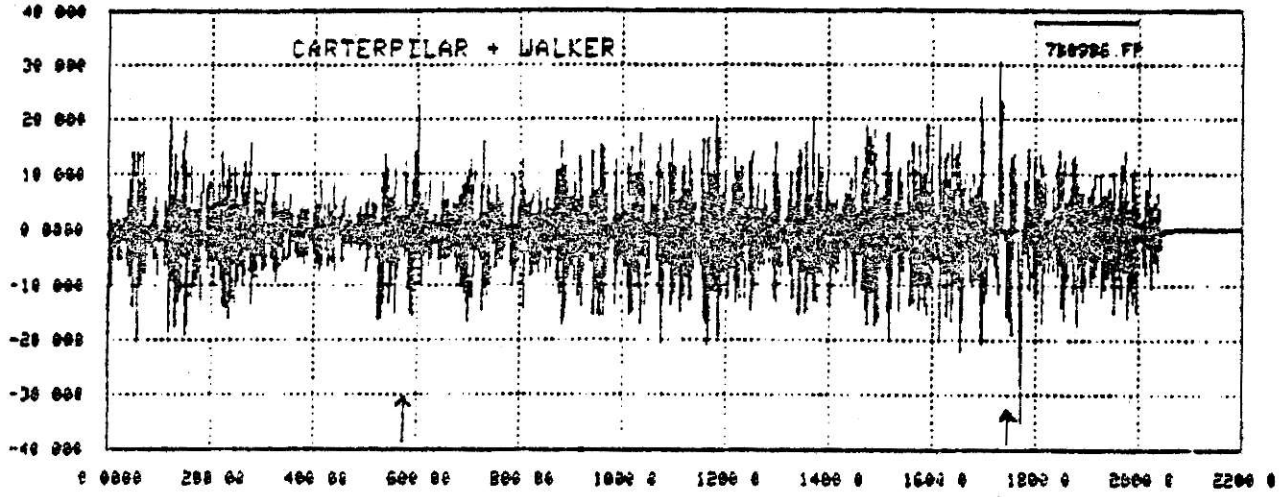
The results of this experiment are summarized in Figures 4.2 through 4.7. The tests were run for five different block sizes. These are N=100, 150, 200, 250, and 300 points.

In this experiment three different algorithms are used; the LMS, WSER, and LTSER. For the LMS predictor, the values used for convergence parameter v, the length of delay Δ, and the number of weights M are 0.0001, 1, and 4, respectively. For the WSER predictor, the values for the initial weights value A_0, the length of delay Δ, and the number of weights M are 0, 1, 4 respectively. The value of u is 0.2, and the

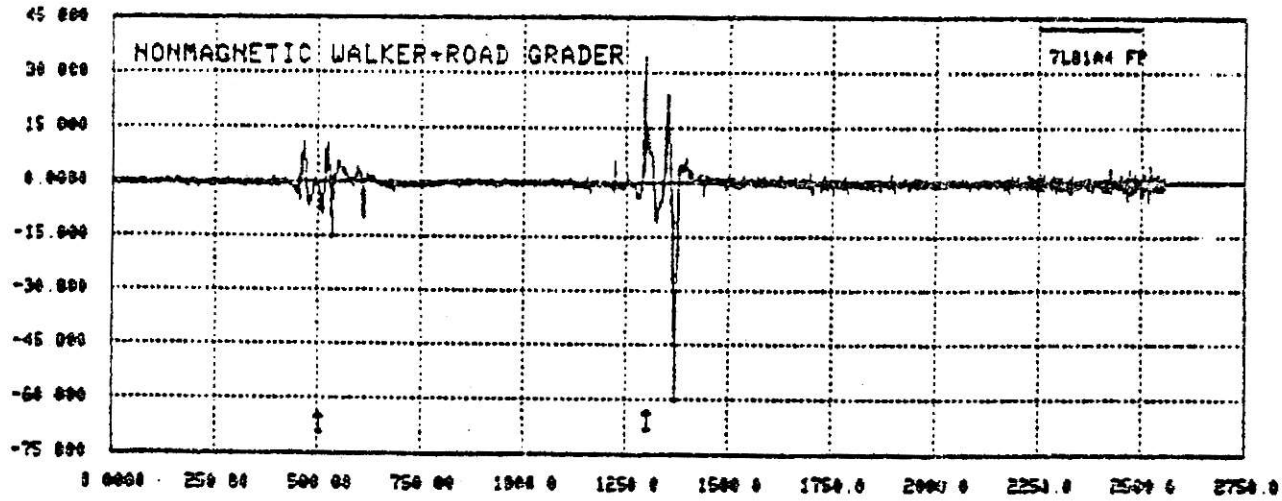
ILLEGIBLE DOCUMENT

THE FOLLOWING DOCUMENT(S) IS OF POOR LEGIBILITY IN THE ORIGINAL

THIS IS THE BEST COPY AVAILABLE

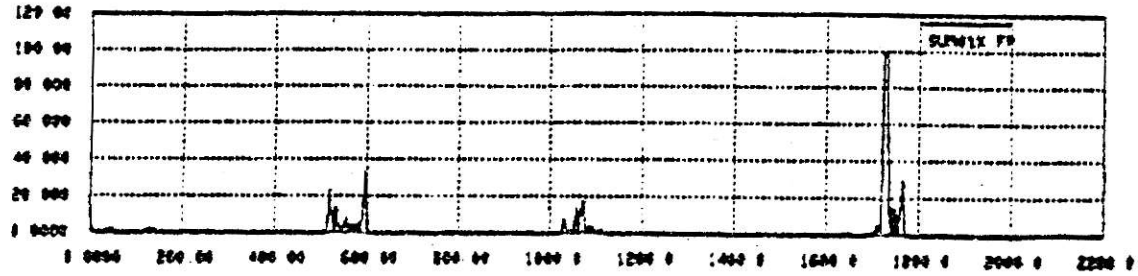


(a) INPUT FILE , NO: 7B09B6. FP

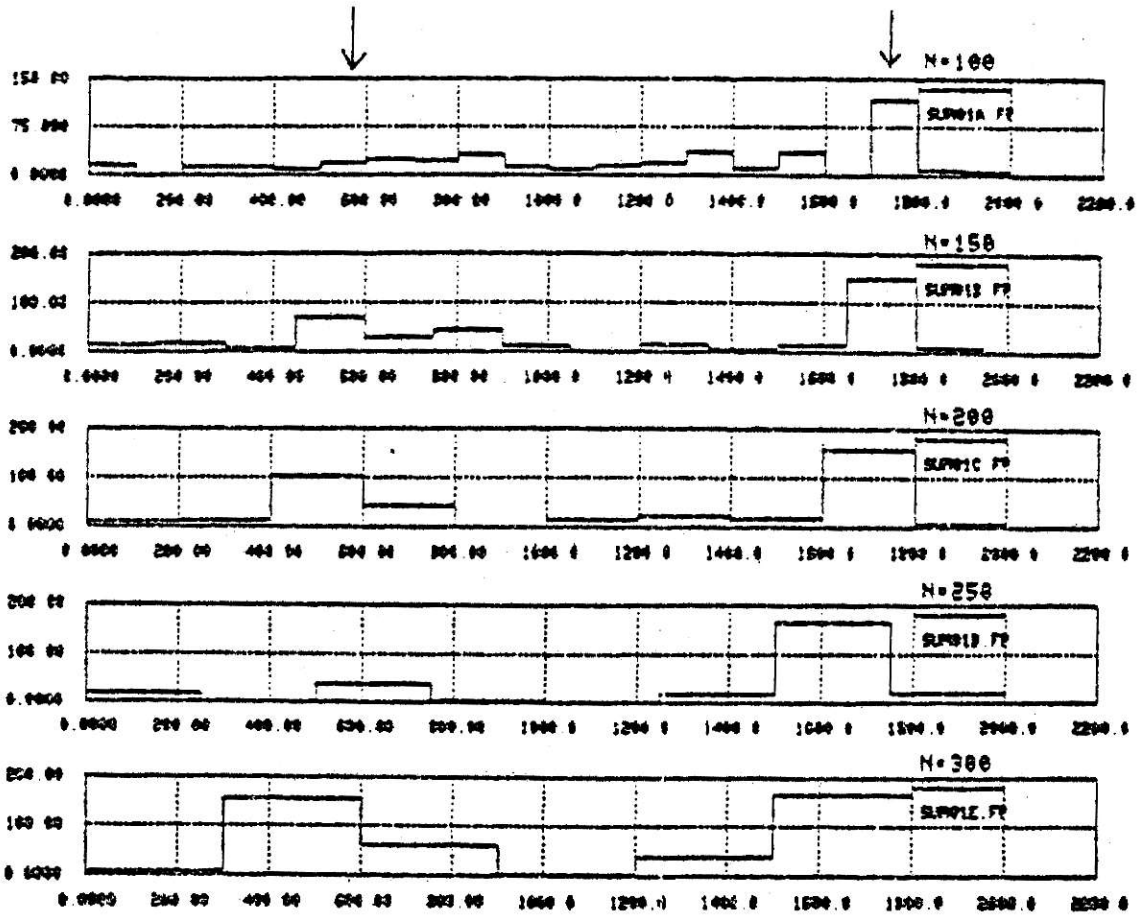


(b) INPUT FILE , NO: 7101A4. FP

Fig. 4 . 1 Input files used for block size experiment.

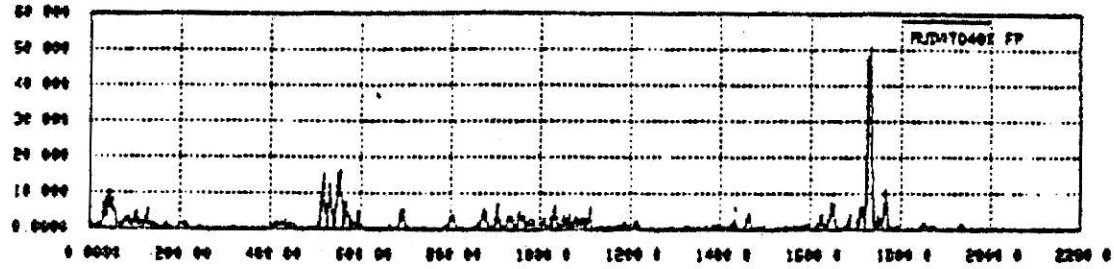


(a) (LMS) pred.- MAF comb. output of 7B09B6. FP

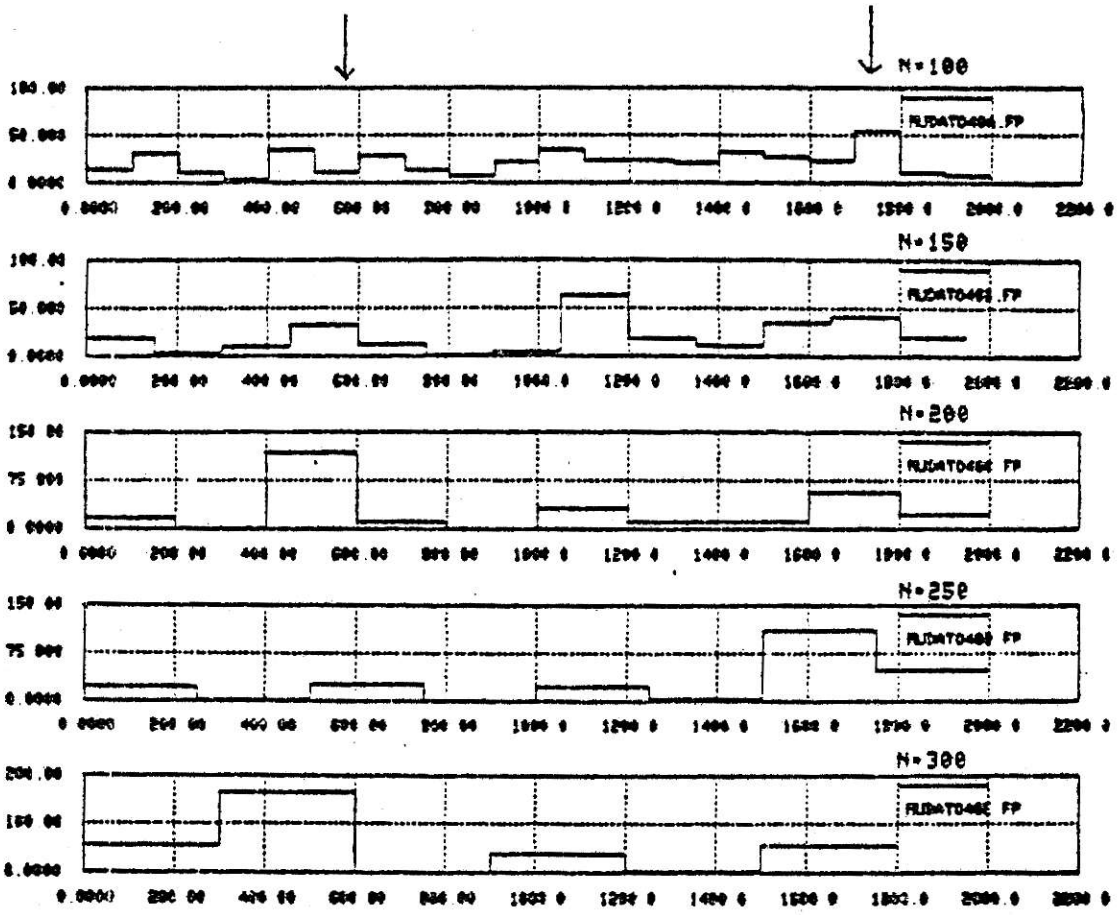


(b) (LMS) pred.- χ^2 test comb. output of 7B09B6. FP

Fig. 4 . 2 Chi-square test response versus block size (LMS predictor, input file 7B09B6.FP).

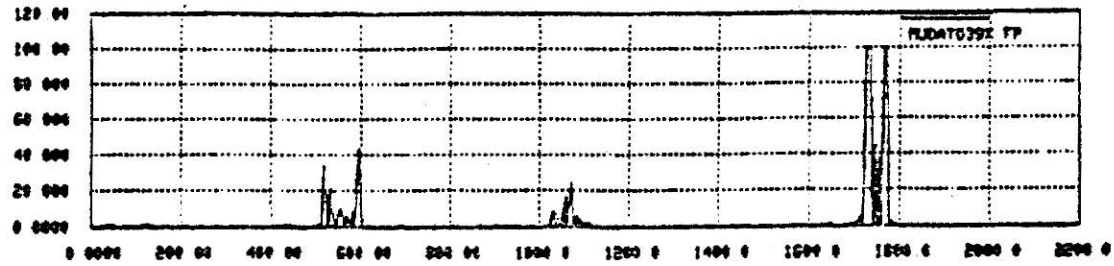


(a) (WSER) pred.- MAF comb. output of 7B09B6

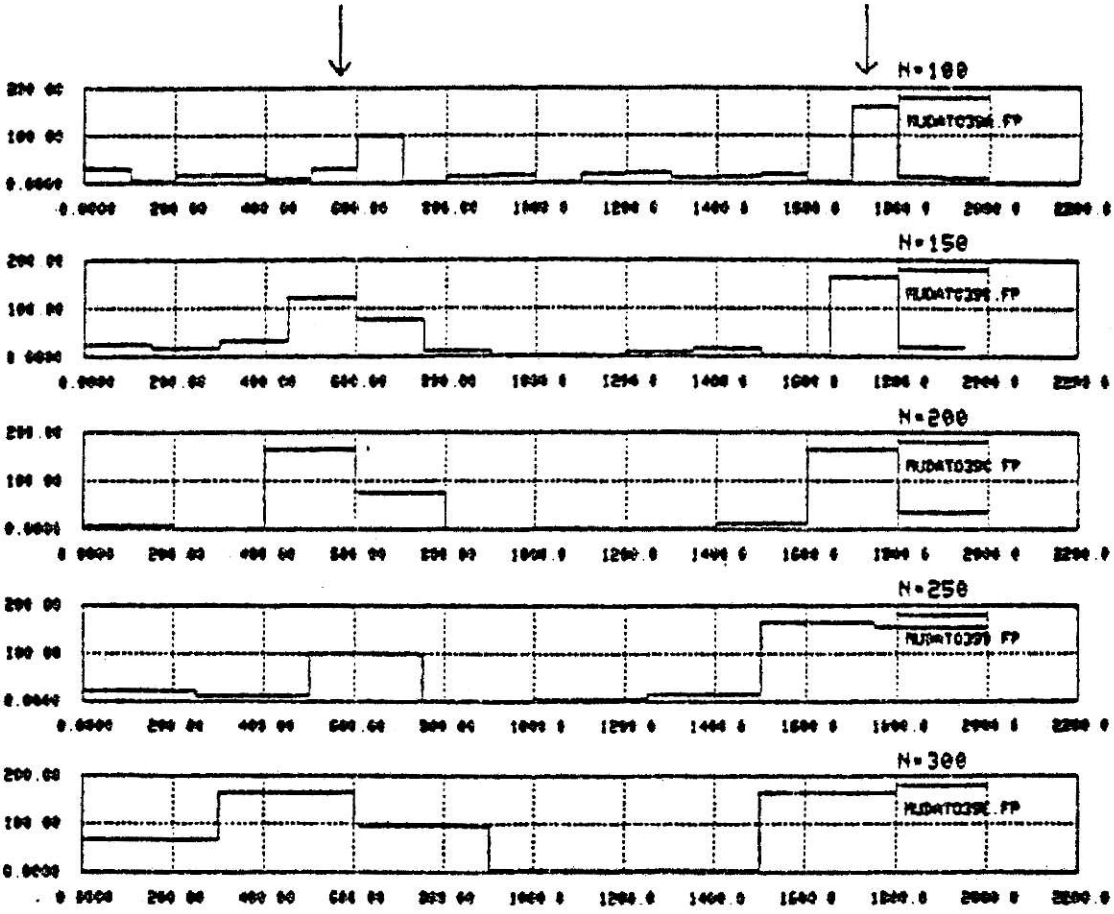


(b) (WSER) pred.- X^2 test comb. output of 7B09B6. FP

Fig. 4 . 3 Chi-square test response versus block size (WSER predictor, input file 7B09B6.FP).

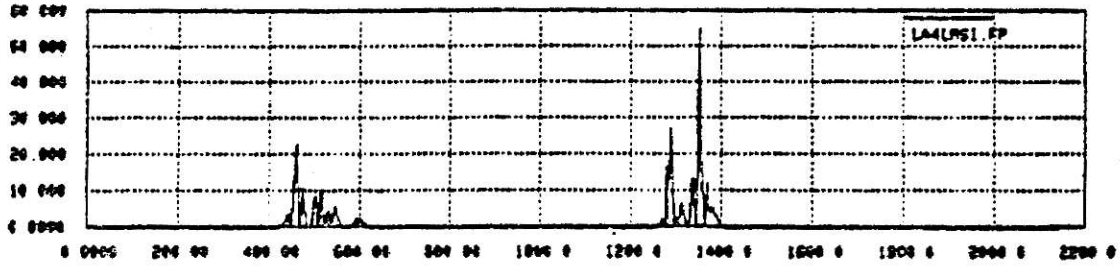


(a) (LTSER) pred.- MAF comb. output of 7B09B6. FP

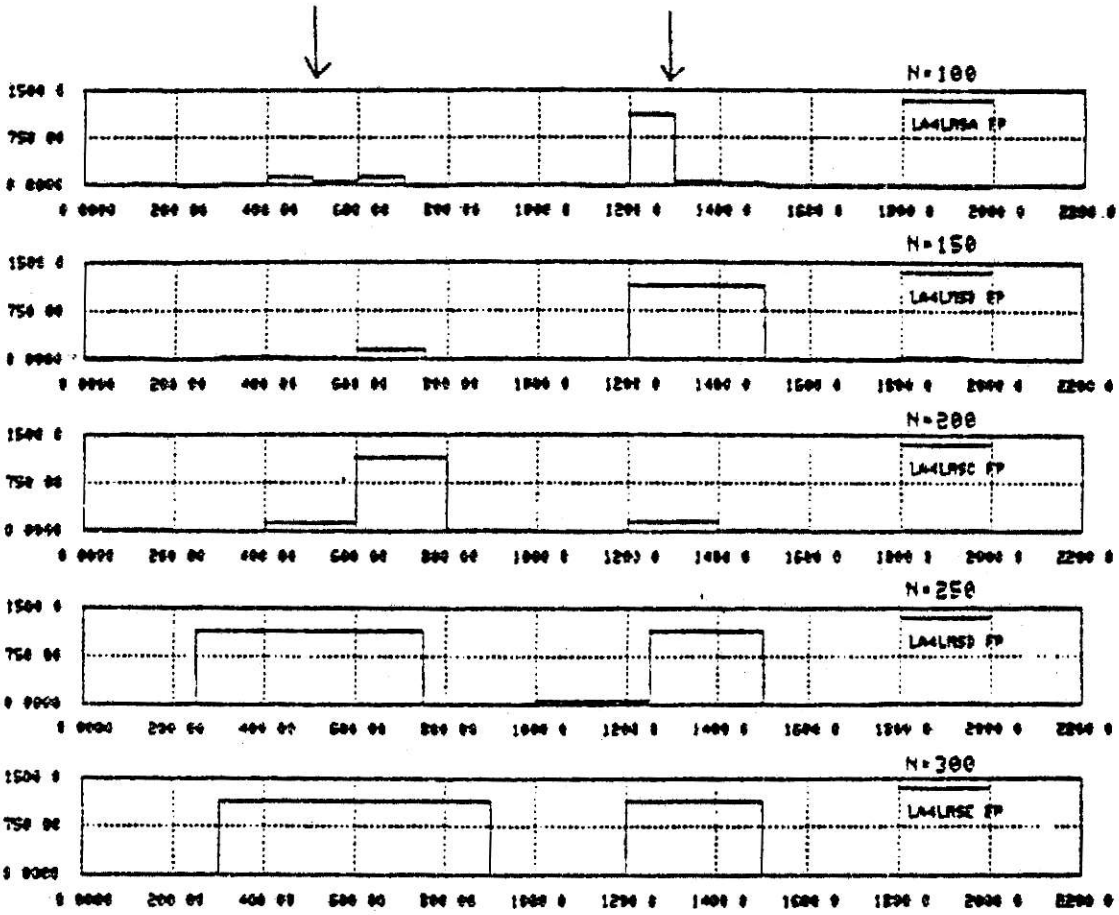


(b) (LTSER) pred.- X^2 test comb. output of 7B09B6. FP

Fig. 4 . 4 Chi-square test response versus block size (LTSER predictor, input file 7B09B6.FP).

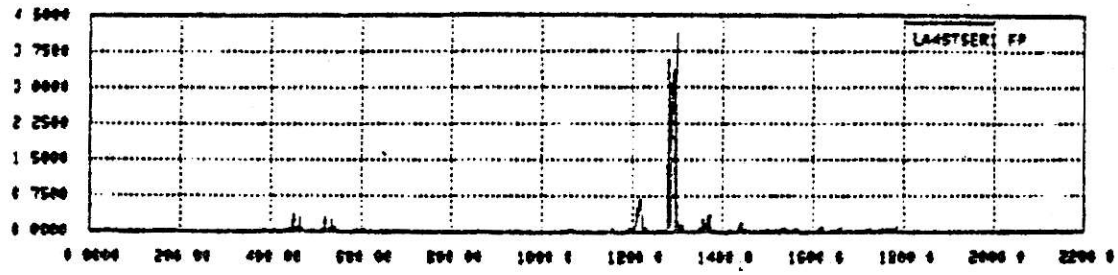


(a) (LMS) pred.- MAF comb. output of 7L01A4. FP

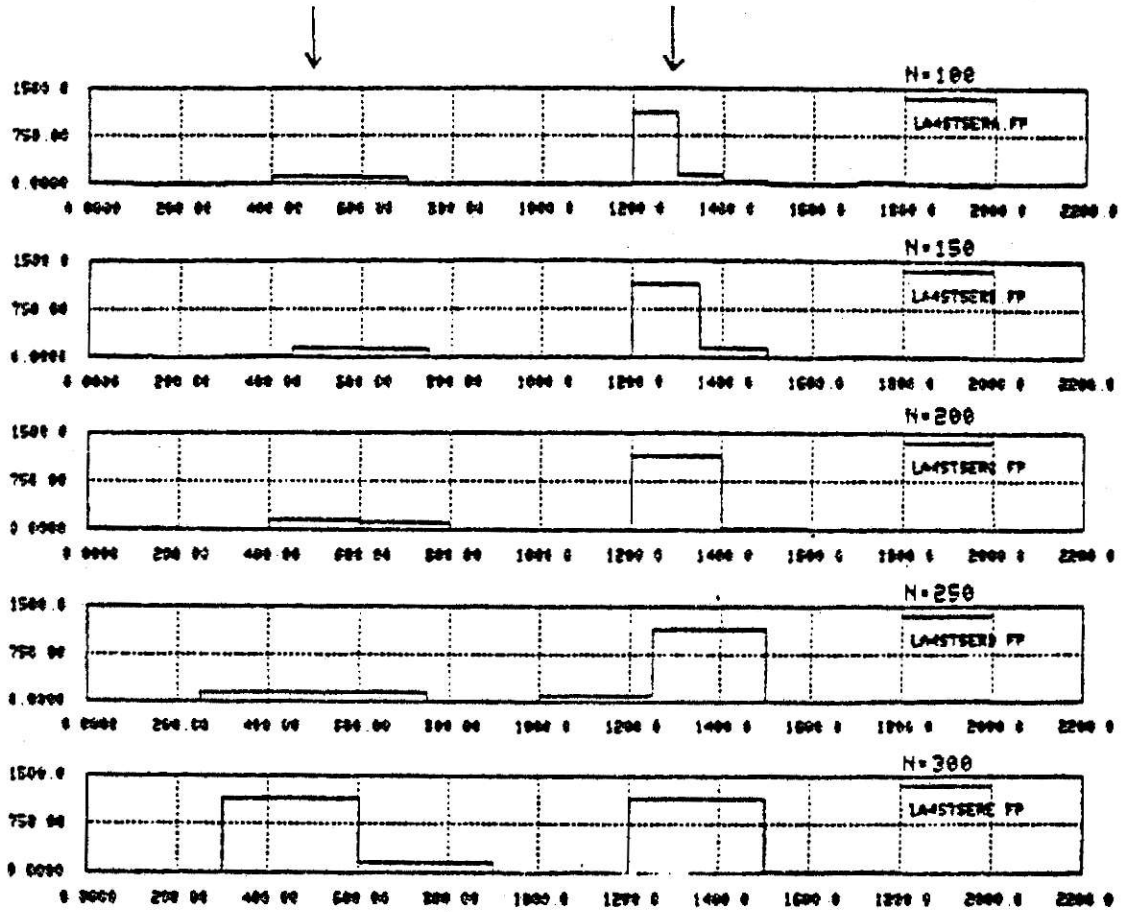


(b) (LMS) pred.- χ^2 test comb. output of 7L01A4.FP

Fig. 4 . 5 Chi-square test response versus block size (LMS predictor, input file 7L01A4.FP).

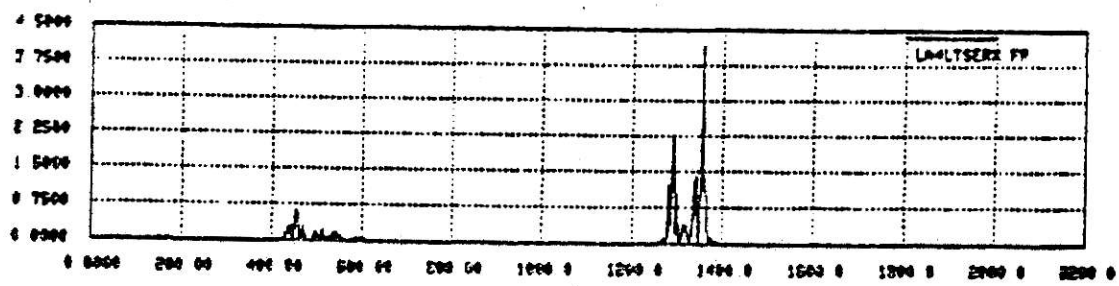


(a) (WSER) pred.- MAF comb. output of 7L01A4. FP

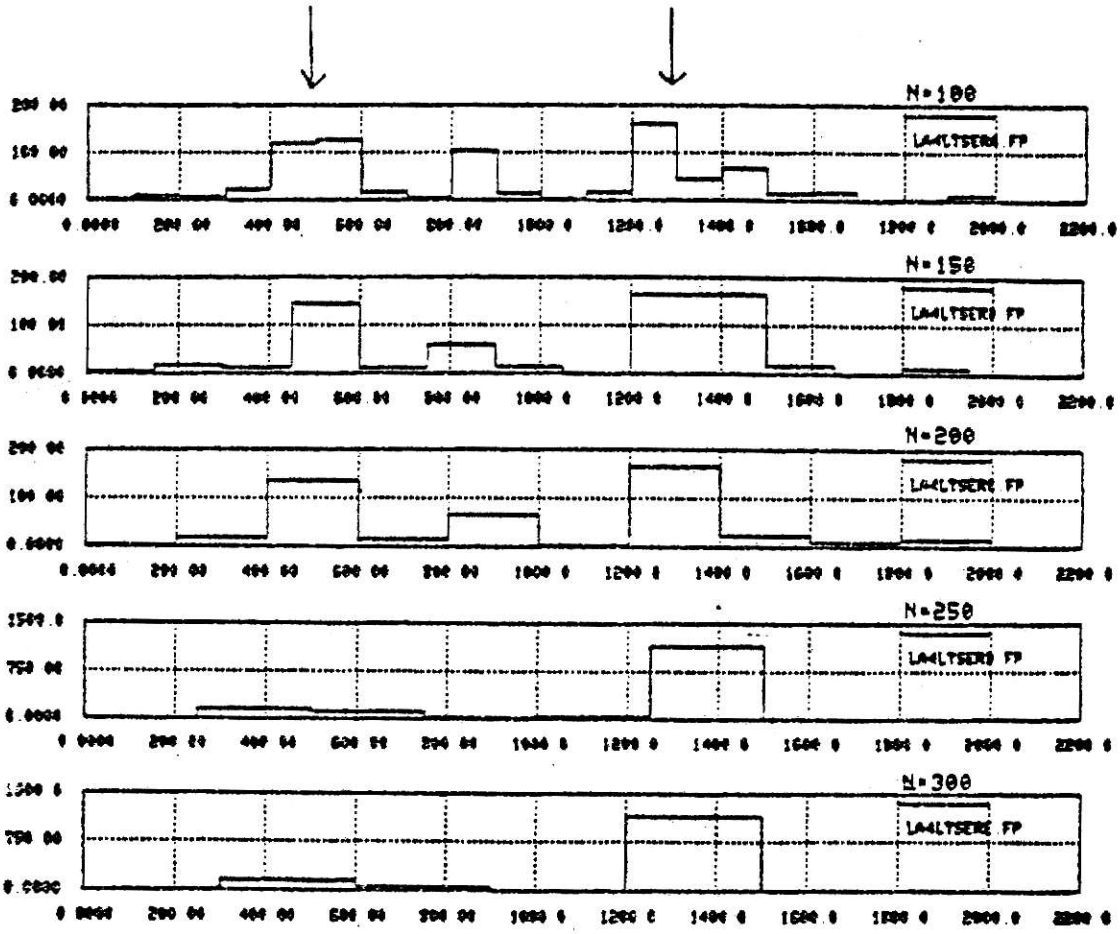


(b) (WSER) pred.- X^2 test comb. output of 7L01A4. FP

Fig. 4 . 6 Chi-square test response versus block size (WSER predictor, input file 7L01A4.FP).



(a) (LTSER) pred.-MAF comb. output of 7101A4. FP



(b) (LTSER) pred.- X^2 test comb. output of 7L01A4. FP

Fig. 4 . 7 Chi-square test response versus block size (LTSER predictor, input file 7L01A4.FP).

value of q is 0.8, and the initial value of the P matrix P_0 is I, where I is the identity matrix. For the LTSER predictor, the values used for the initial diagonal matrix P_0, the number of weights M, and the length of delay Δ are I, 4, and 1, respectively. Note that the values of the number of classes K used for each block are 6, 8, 8, 10, and 12 when the corresponding values of N are equal to 100, 150, 200, 250, and 300, respectively.

Each of the illustrations which follow show Chi-square test output versus block size for a given type of predictor and a given input data file. For comparison purposes the performance of a predictor MAF combination is also shown for each case.

These detection systems are constructed in such a way that a maximum or peak in the output signal is taken as an indication of the presence of an intruder. In practice the signal is compared with a threshold and each positive crossing of the threshold is taken as an indication of an intruder. In all of data files used in this report the location of the intruder is known. Because these signals are often not apparent due to background noise, the location of the intruder is marked on the input file using the symbol "↑". In this way we are able to evaluate performance in a subjective way by looking at the output files from the various tests to see whether the intruder would have been recognized by a single threshold test.

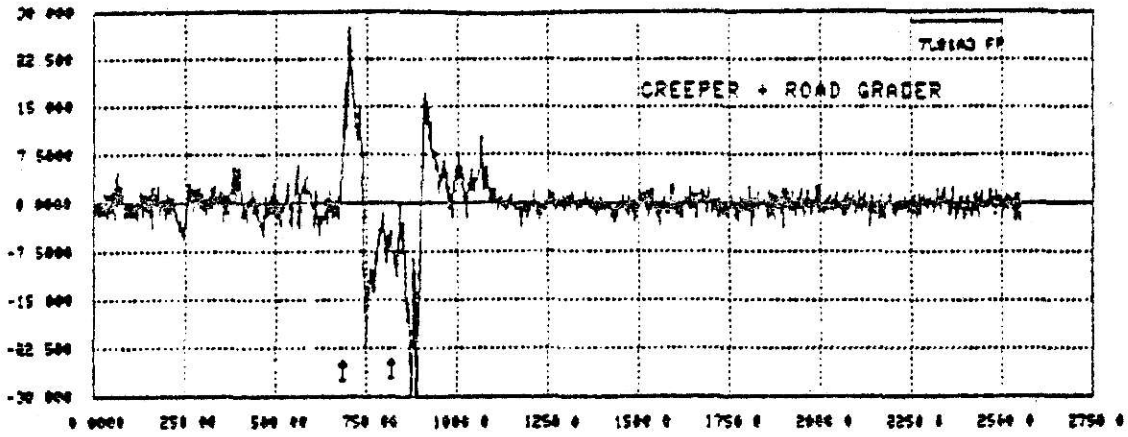
It is clear from the Figures that in most cases a block size of N = 200 points for the predictor Chi-square test combination results in a good detection of the intruder's presence. Hereafter in experiments involving the use of a predictor Chi-square test combination for detecting the presence an intruder, the value of N equal to 200 points will be used.

B. PERFORMANCE OF THE LMS PREDICTOR CHI-SQUARE TEST COMB.

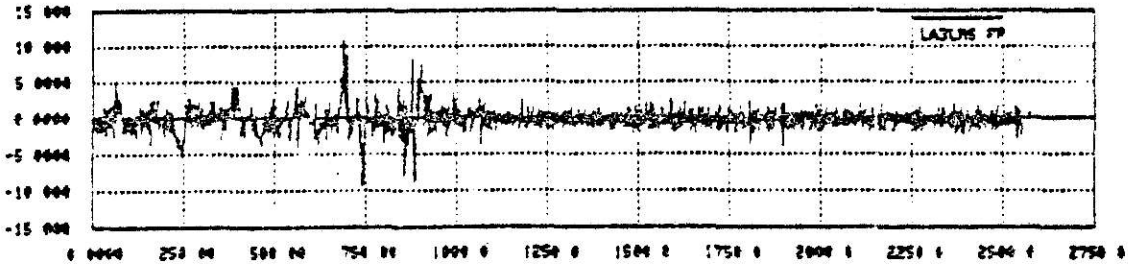
The objective of the experiment described in this section was to evaluate the performance of the Chi-square test approach to intrusion detection when Widrow's LMS algorithm is used for the predictor. The approach taken was to select ten different data files and process these files using the LMS predictor Chi-square test combination.

The results of the tests are shown in Figures 4.8 through 4.17. In each of these illustrations four signals are shown. The first is the input data file. The input files differ with respect to the type of background noise and intruder. In each case the location of the intruder is marked on the input file using the symbol "↑". The other signals shown are the LMS predictor output, the LMS predictor-MAF combination output, and the LMS predictor Chi-square test combination output. The MAF signal was included for comparision with the Chi-square test signal.

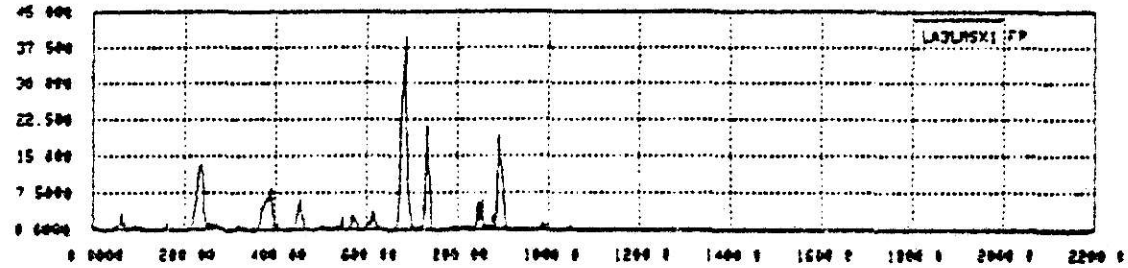
In conducting these tests the values used for the convergence parameter ν, the length of delay Δ, and the number of weights M were 0.0001, 1, and 4, respectively. The number of data points N in each block, and the number of classes in each block were 200, and 8, respectively.



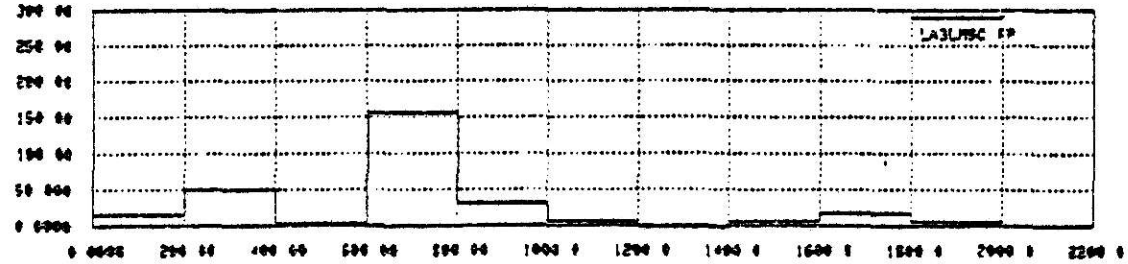
a. input file



b. (LMS) predictor output

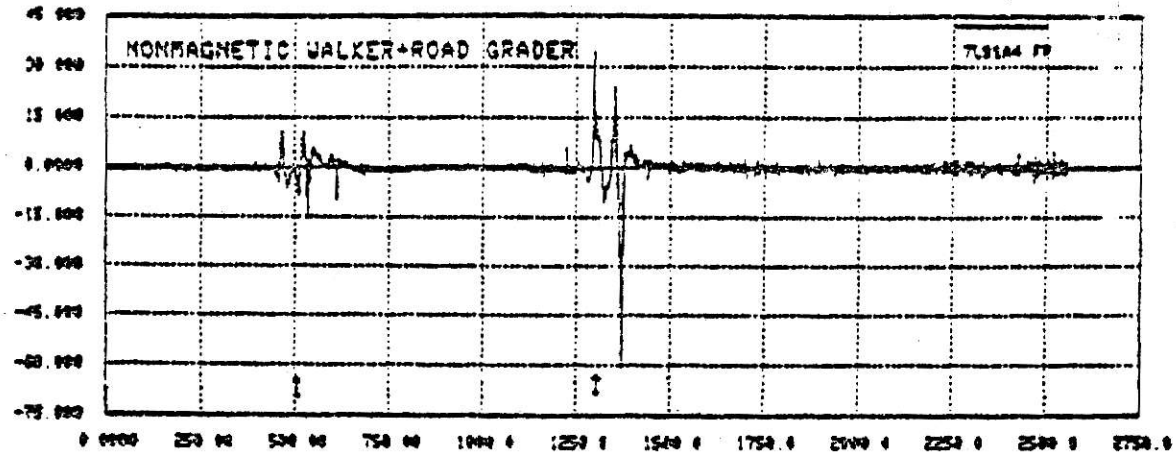


c. (LMS) pred.- MAF comb. output

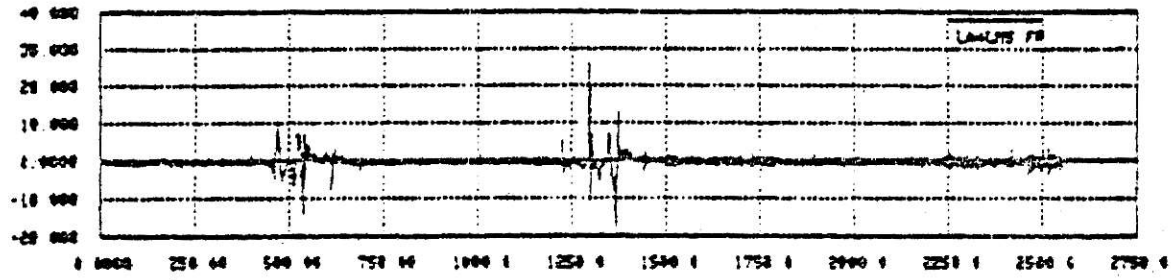


d. (LMS) pred.- X^2 test combination output

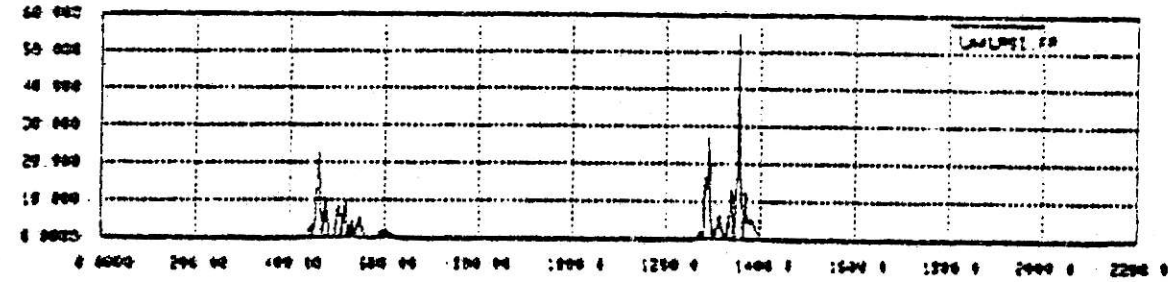
Fig. 4 . 8 Chi-square test performance using LMS pred. (creeper plus road grader input).



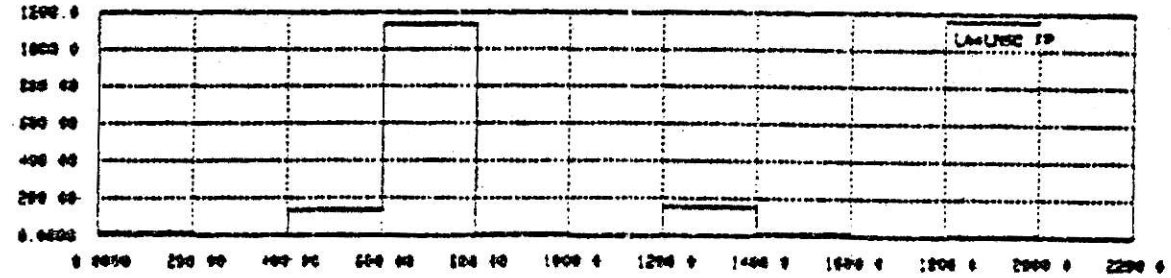
a. input file



b. (LMS) predictor output

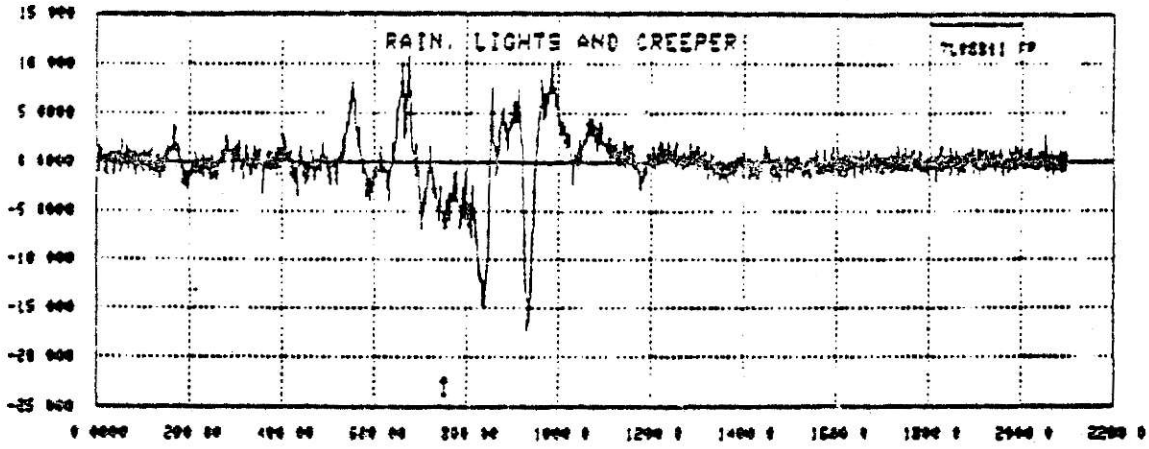


c. (LMS) pred.- MAF comb. output

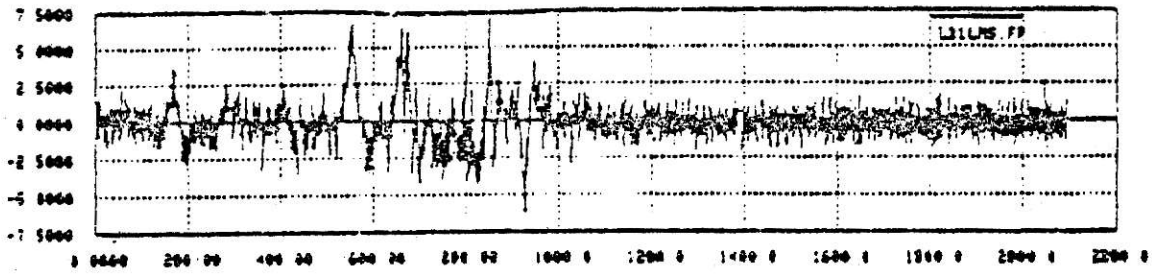


d. (LMS) pred.- Ψ^2 test comb. output

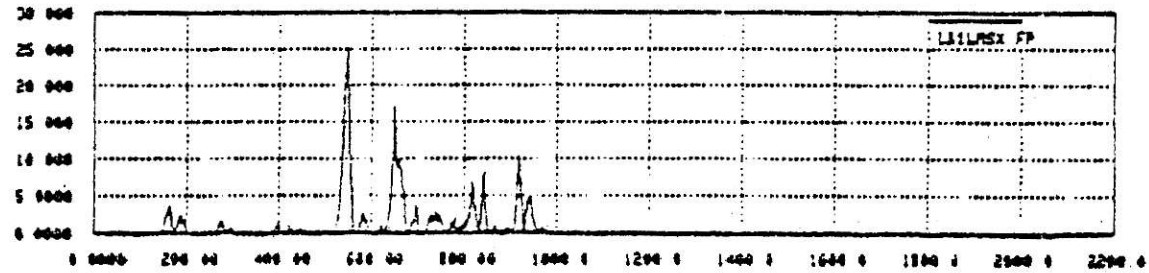
Fig. 4 . 9 Chi-square test performance using LMS pred. (nonmagnetic walker plus road grader input).



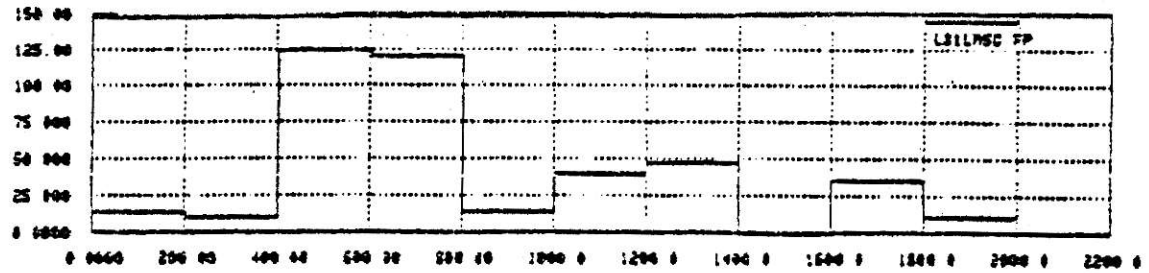
a. Input file



b. (LMS) predictor output

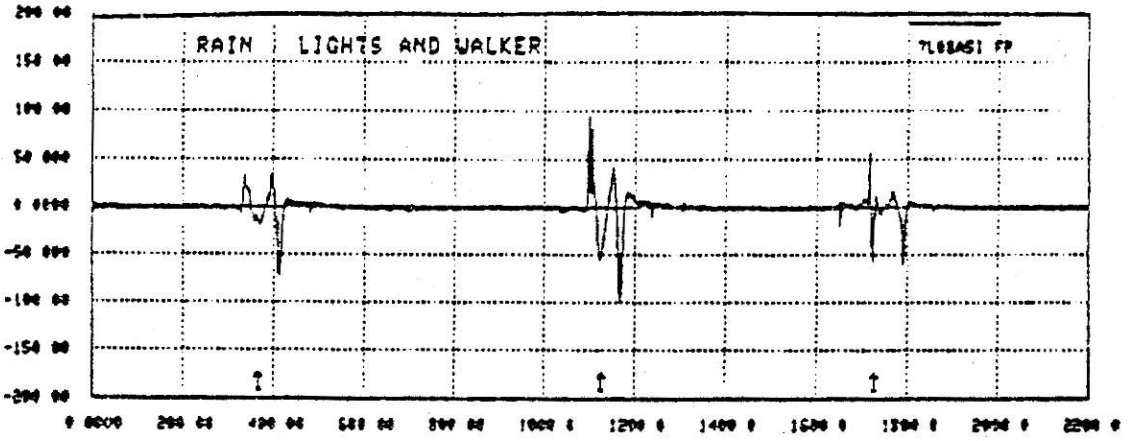


c. (LMS) pred.- MAF comb. output

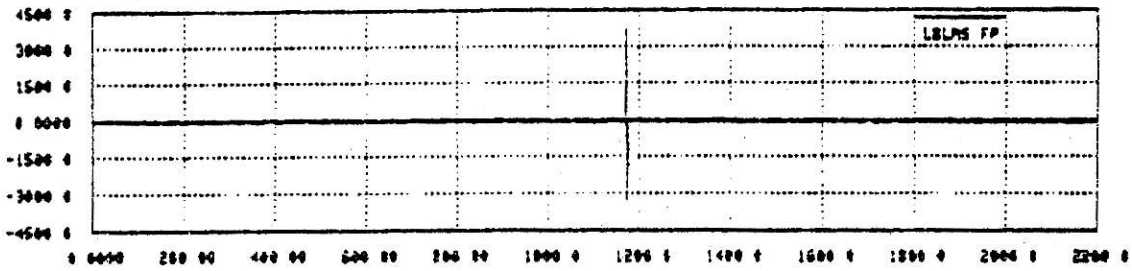


d. (LMS) pred.- X^2 test combination output

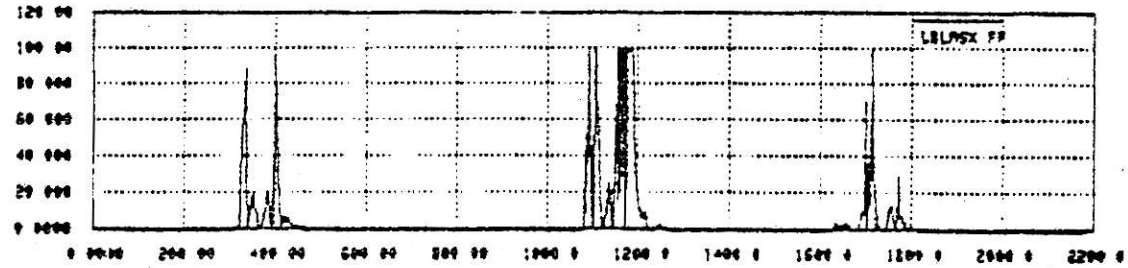
Fig. 4 . 10 Chi-square test performance using LMS pred. (creeper plus rain, lights input).



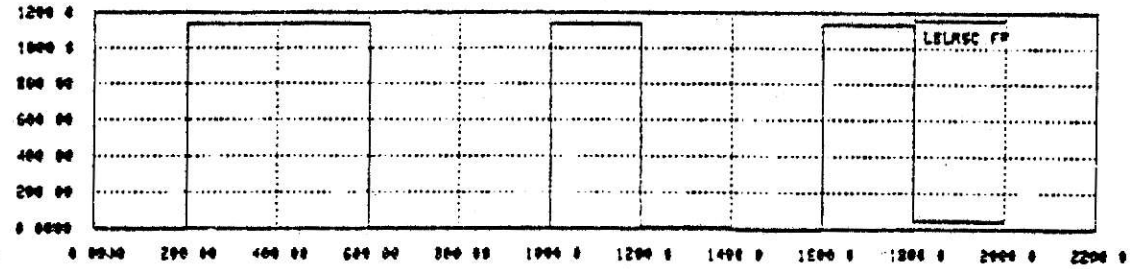
a. input file



b. (LMS) predictor output

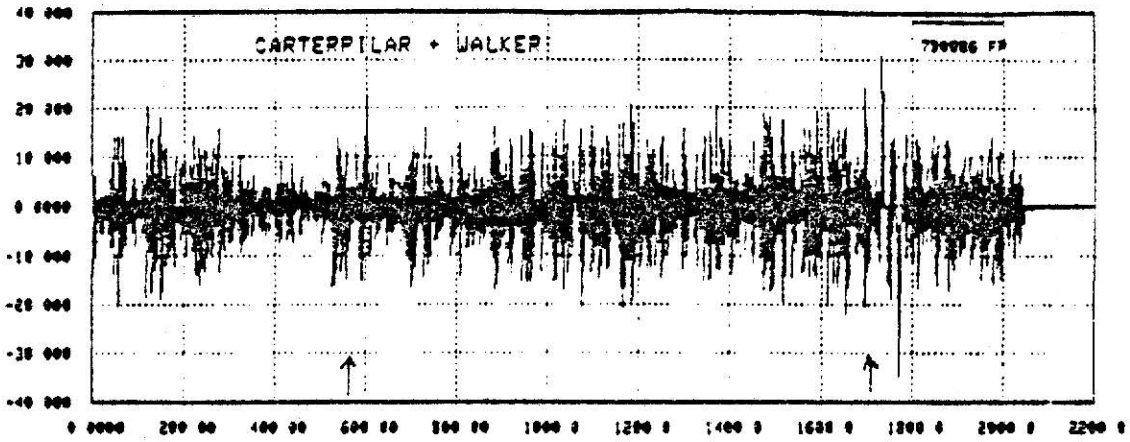


c. (LMS) pred.- MAF comb. output

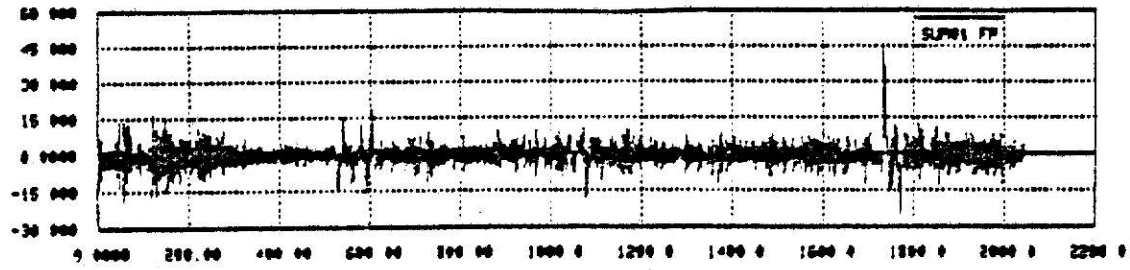


d. (LMS) pred.- X^2 test combination output

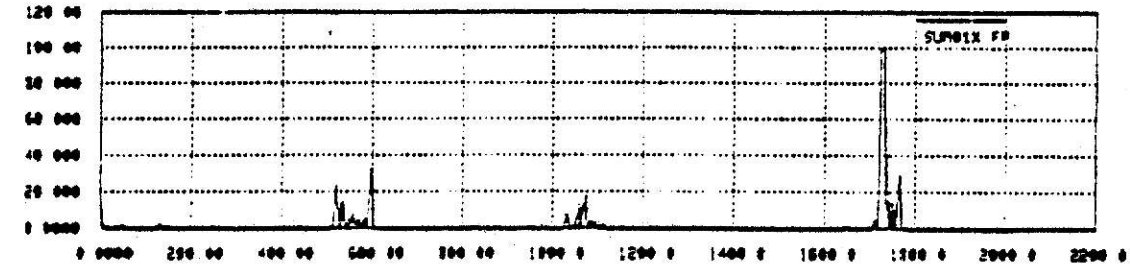
Fig. 4 . 11 Chi-square test performance using LMS pred. (walker plus rain, lights input).



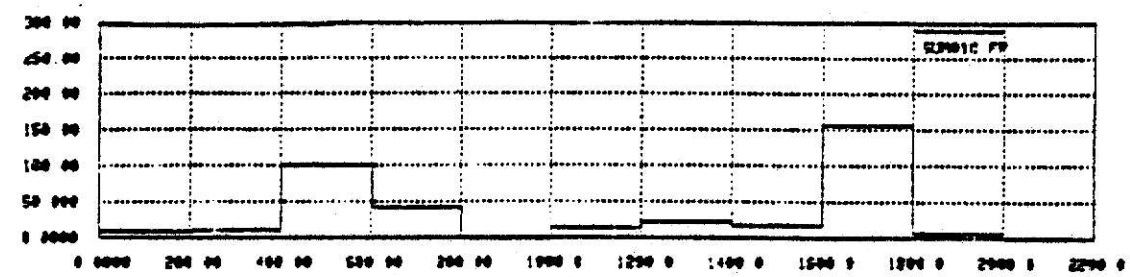
a. input file



b. (LMS) predictor output

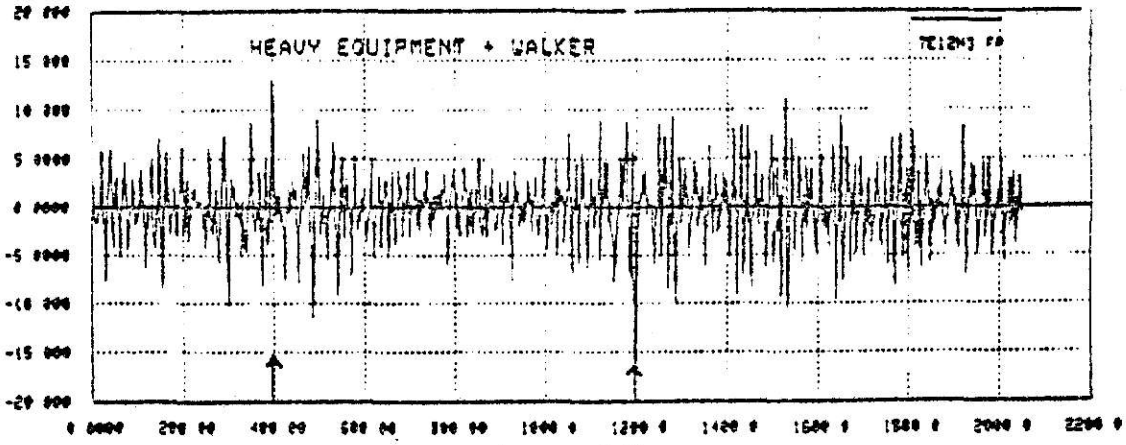


c. (LMS) pred.- MAF comb. output

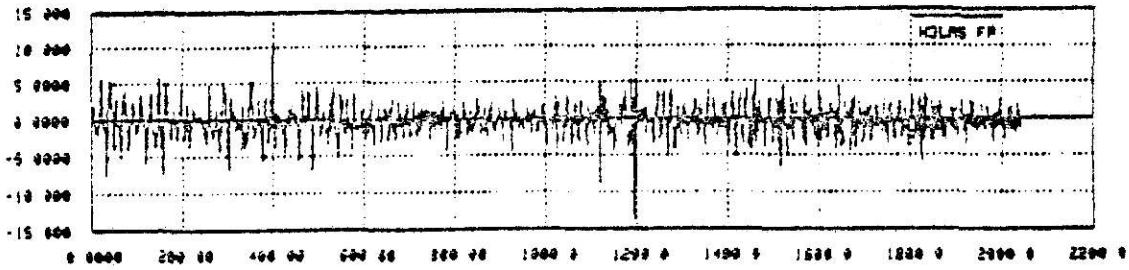


d. (LMS) pred.- X^2 test combination output

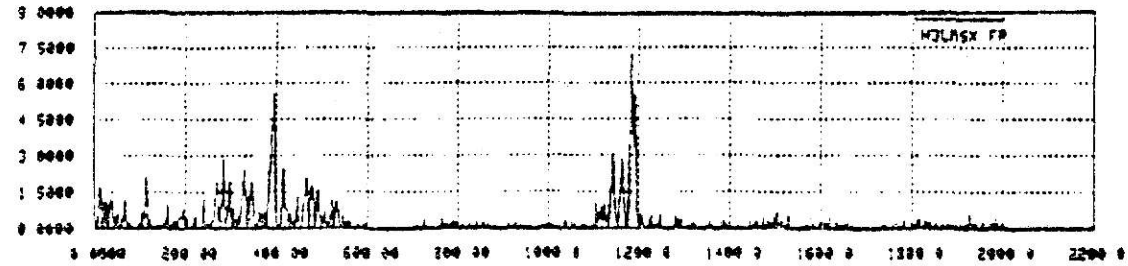
Fig. 4 . 12 Chi-square test performance using WSER pred. (walker plus caterpilar input).



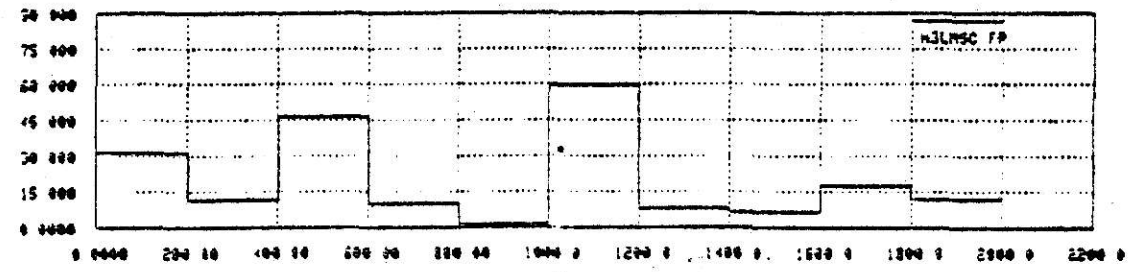
a. input file



b. (LMS) predictor output

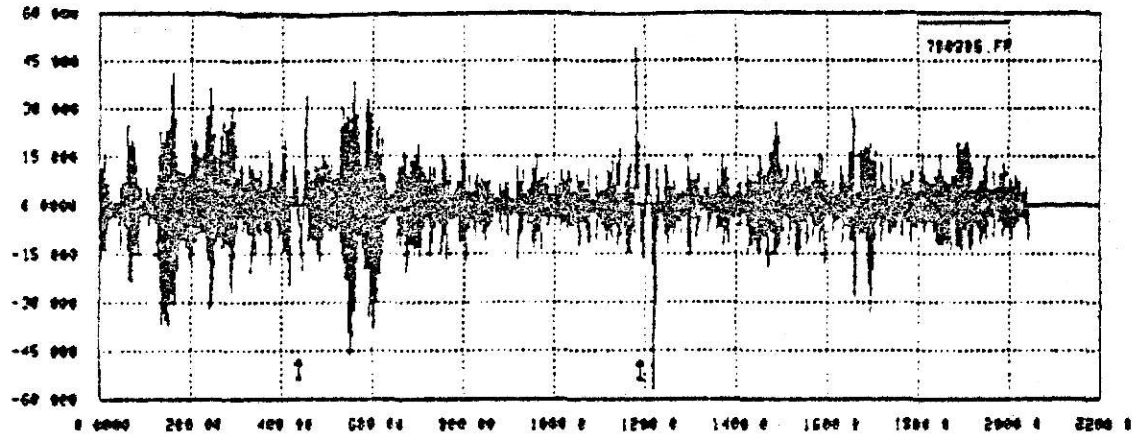


c. (LMS) pred.- MAF comb. output

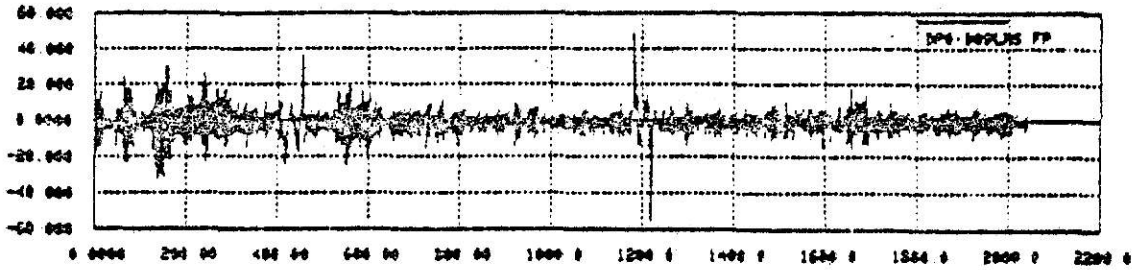


d. (LMS) pred.- X^2 test combination output

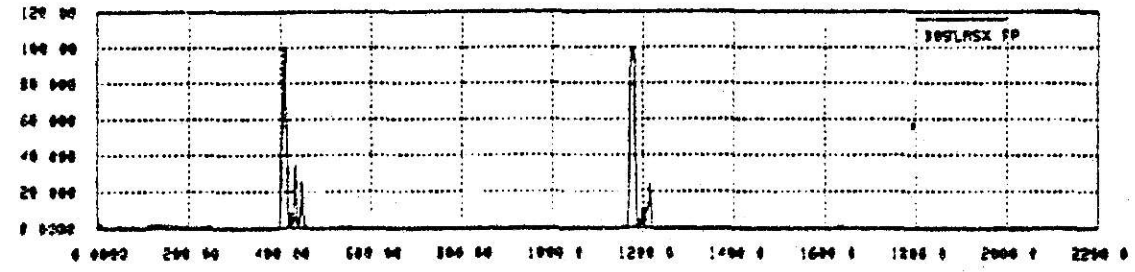
Fig. 4 . 13 Chi-square test performance using LMS pred. (walker plus heavy equipment input)



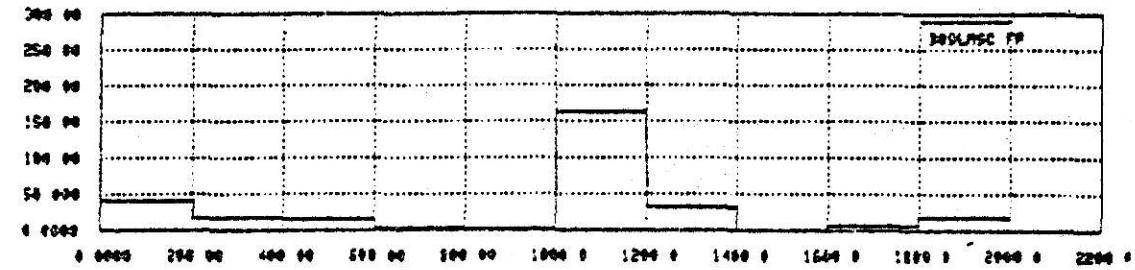
a. input file



b. (LMS) predictor output

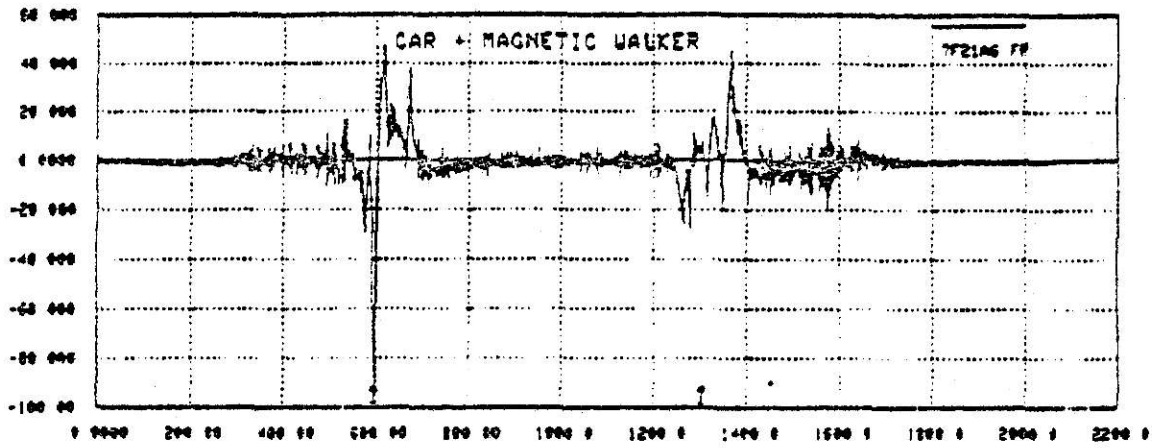


c. (LMS) pred.- MAF comb. output

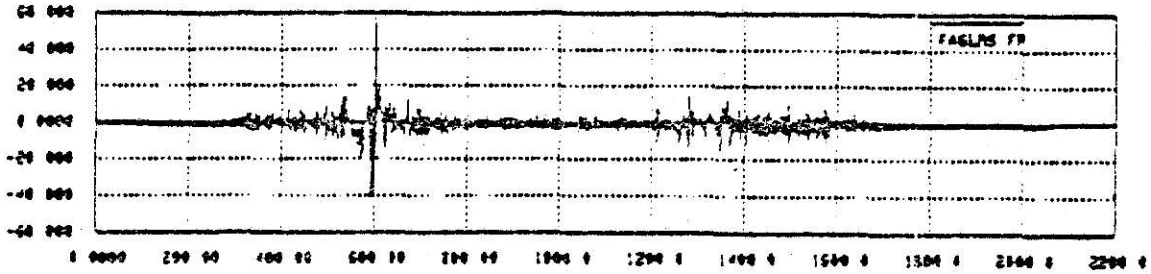


d. (LMS) pred.- X^2 test combination output

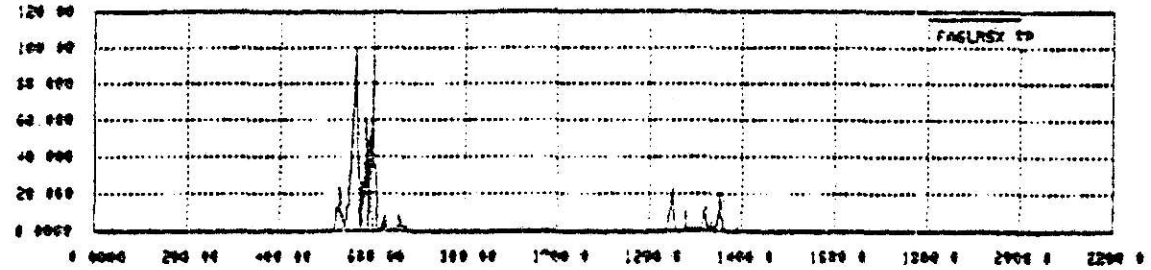
Fig. 4 . 14 Chi-square test performance using LMS pred. (unindentify input).



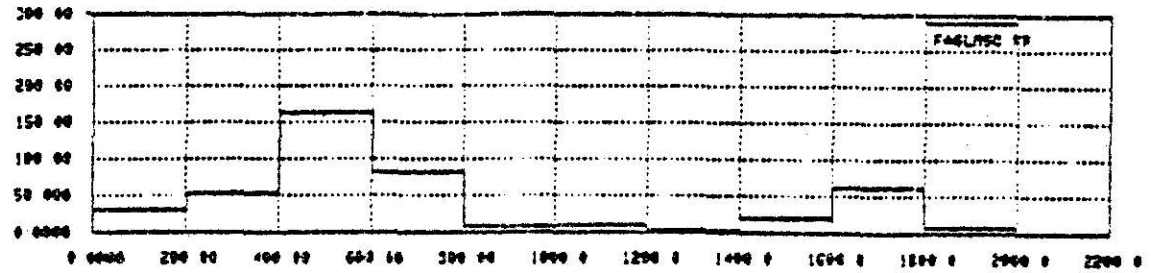
a. input file



b. (LMS) predictor output

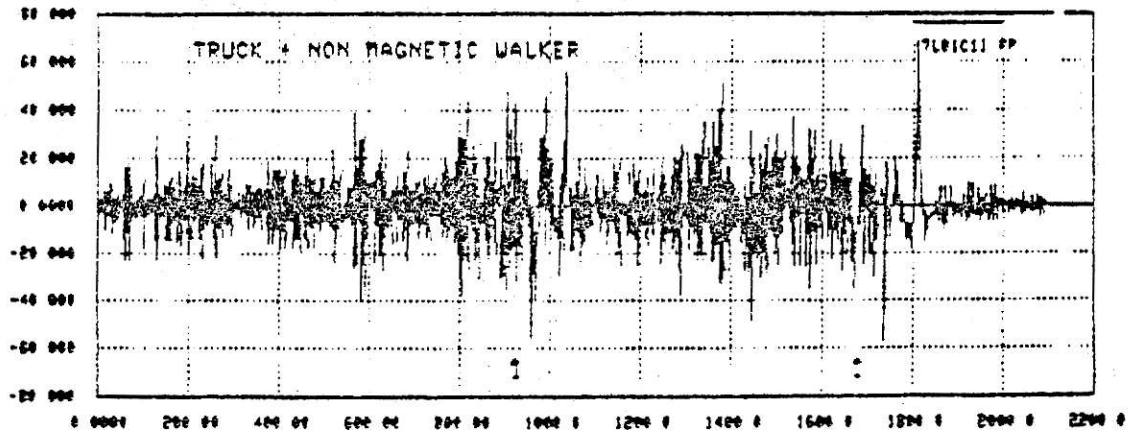


c. (LMS) pred.- MAP comb. output

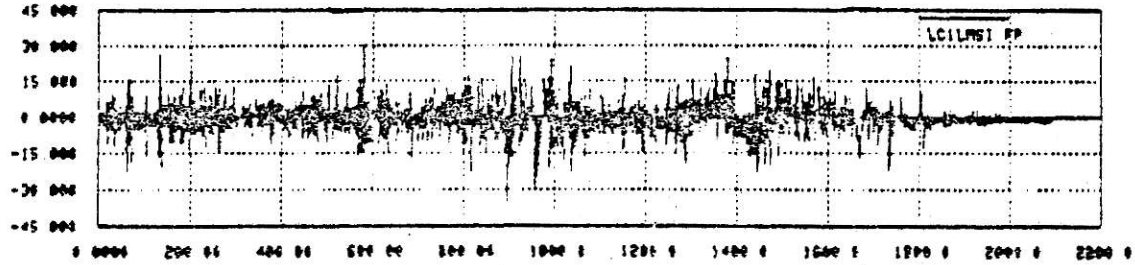


d. (LMS) pred.- I^2 test combination output

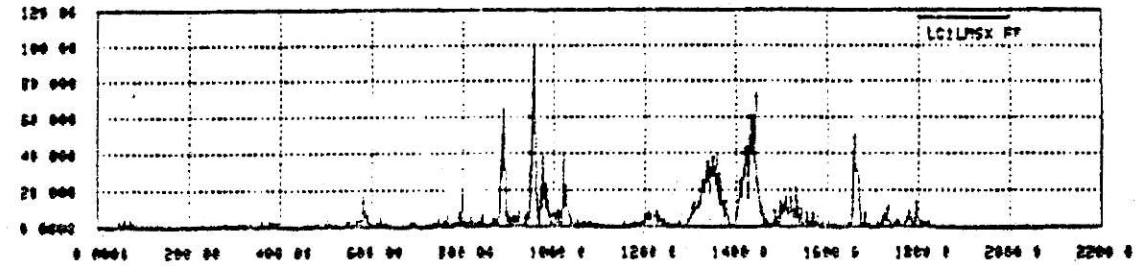
Fig. 4 . 15 Chi-square test performance using LMS pred. (magnetic walker plus car input).



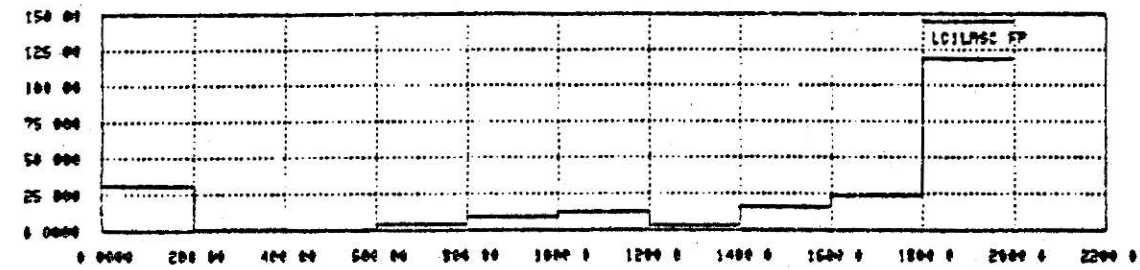
a. input file



b. (LMS) predictor output

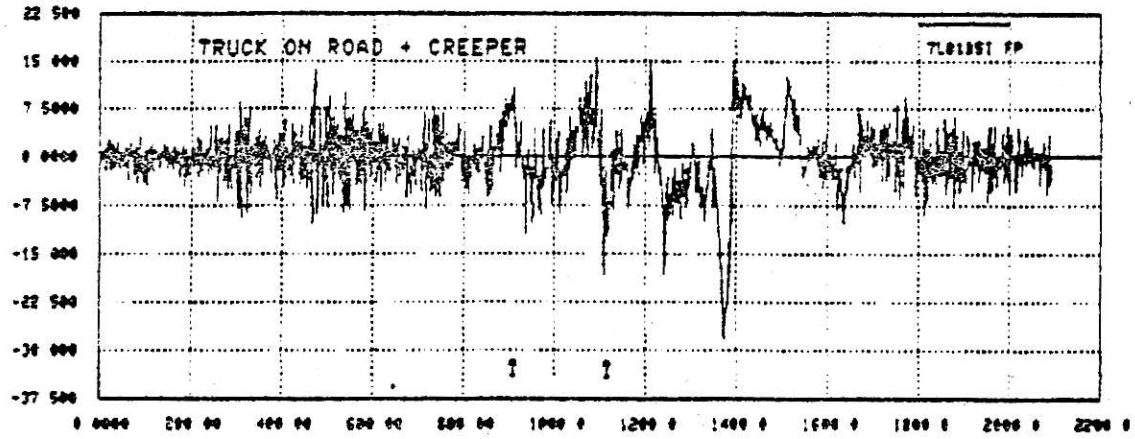


c. (LMS) pred.- MAF comb. output

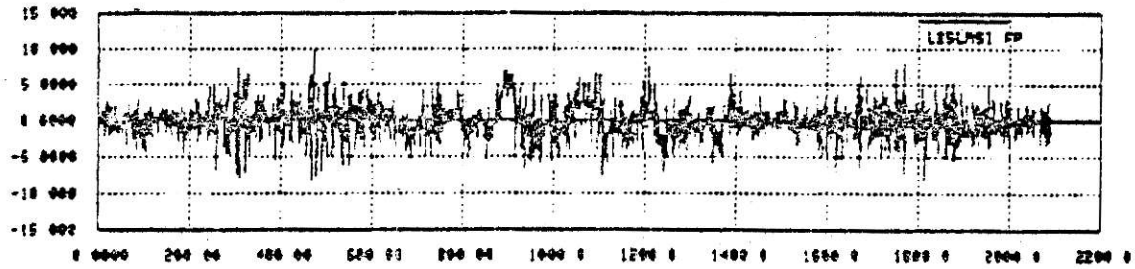


d. (LMS) pred.- X^2 test combination output

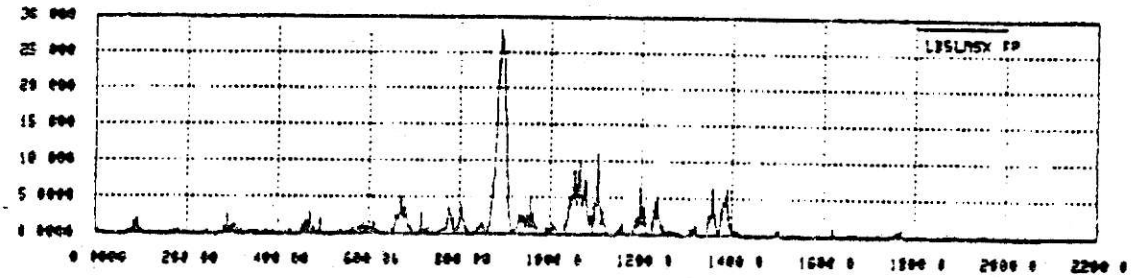
Fig. 4 . 16 Chi-square test performance using LMS pred. (nonmagnetic walker plus truck input).



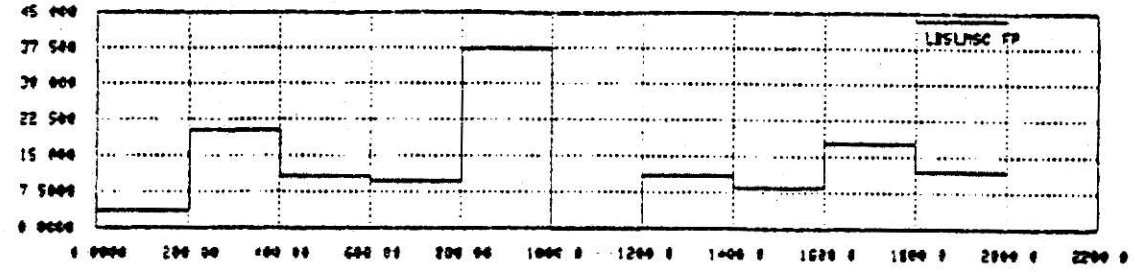
a. input file



b. (LMS) predictor output



c. (LMS) pred.- MAF comb. output



d. (LMS) pred.- X^2 test combination output

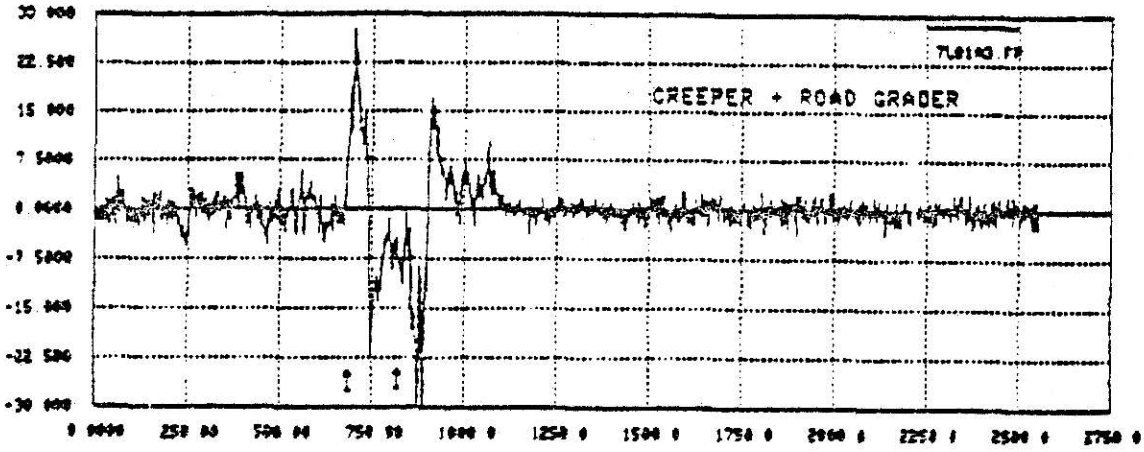
Fig. 4 . 17 Chi-square test performance using LMS pred. (creeper plus truck on road input).

C. PERFORMANCE OF THE WSER PREDICTOR CHI-SQUARE TEST COMB.

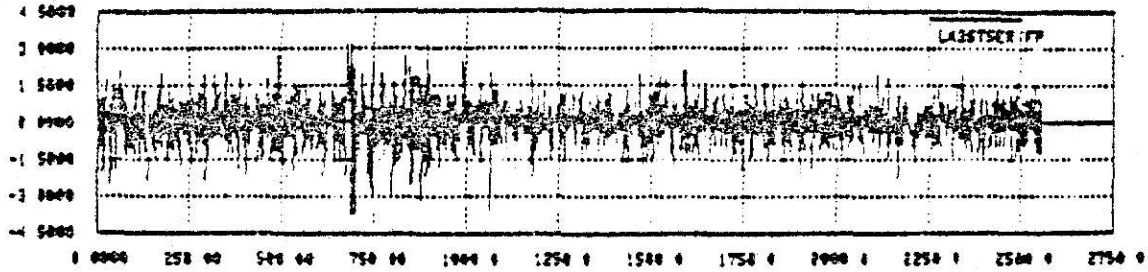
The performance of the WSER predictor Chi-square test combination are reported in this section. In this case the ten data files used in previous experiment were processed using the WSER predictor Chi-square test combination. It was found that for the particular set of parameters used in this experiment the performance as an intrusion detection was not as good as that of the LMS approach of the previous section.

Figures 4.18 through 4.27 show the results of the tests. In each of these illustrations four signals are shown. The first is the input data file. As before, in each case the location of the intruder is marked on the input file using the symbol "*". The other signals shown are the WSER predictor output, the predictor MAF combination output, and the predictor Chi-square test combination output. For the purpose of comparison, the MAF signal was also included.

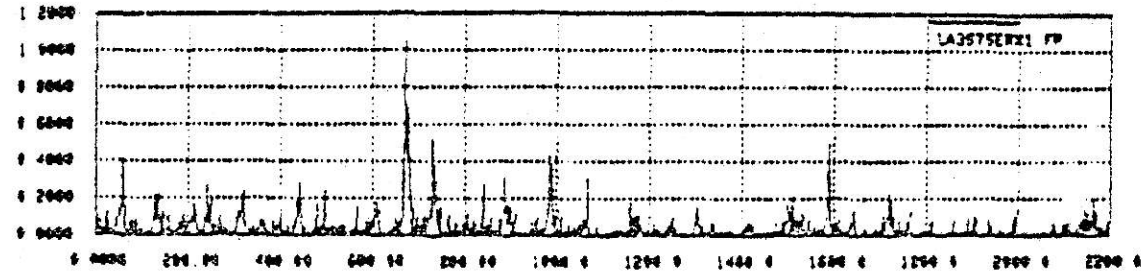
In processing these tests the values used for the parameter u, the parameter q, and the number of weights M were 0.2, 0.8, and 4, respectively. The length of delay was 1 and the initial value of the weight vector A_0 was zero. The initial diagonal term of the matrix P_0 was I. The values used for the number of data points N, and the number of classes K in each block were 200, and 8, respectively.



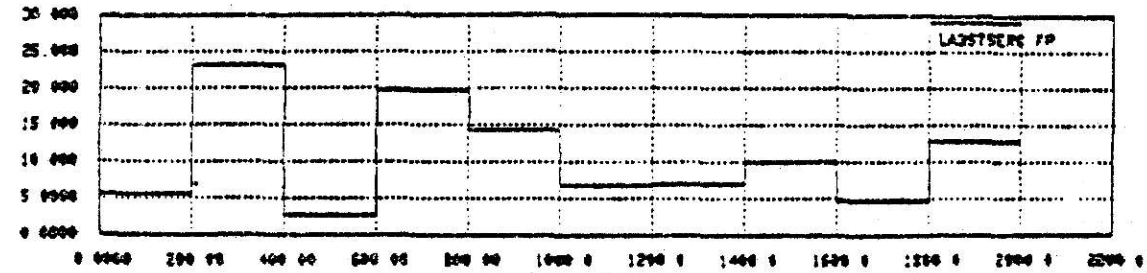
a. input file



b. (WSER) predictor output

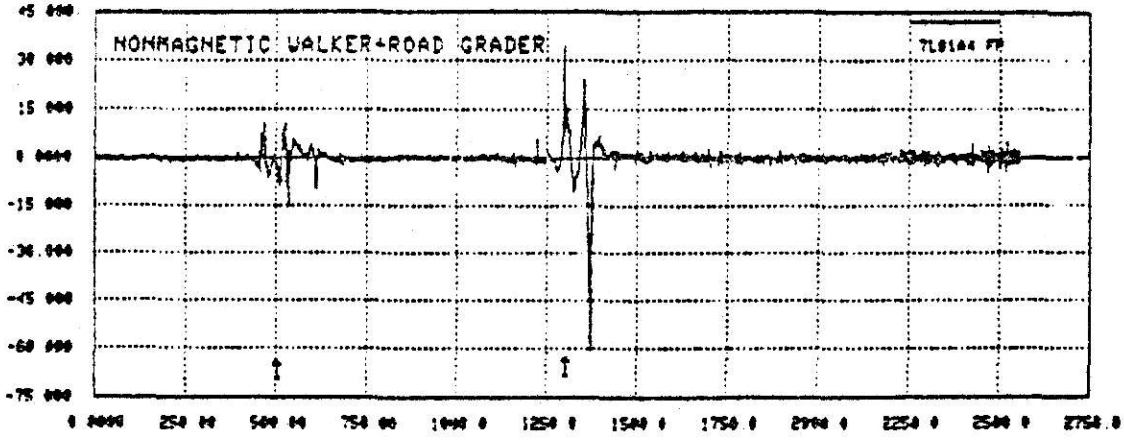


c. (WSER) pred.- MAF comb. output

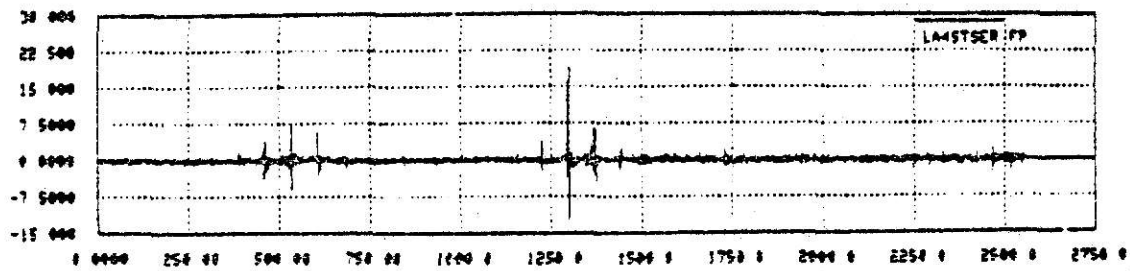


d. (WSER) pred.- X^2 test comb. output

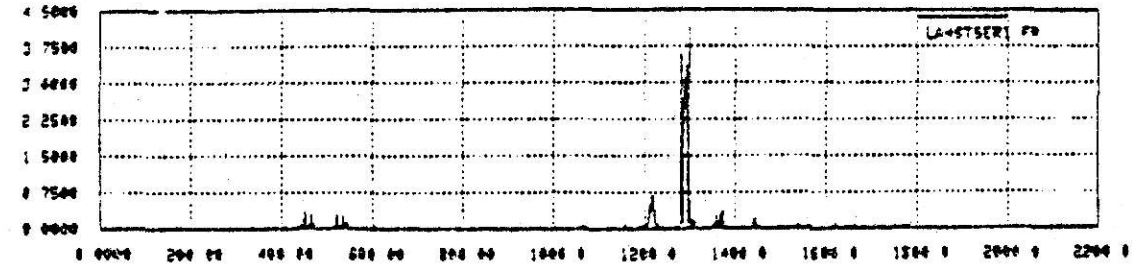
Fig. 4 . 18 Chi-square test performance using WSER pred. (creeper plus road grader input).



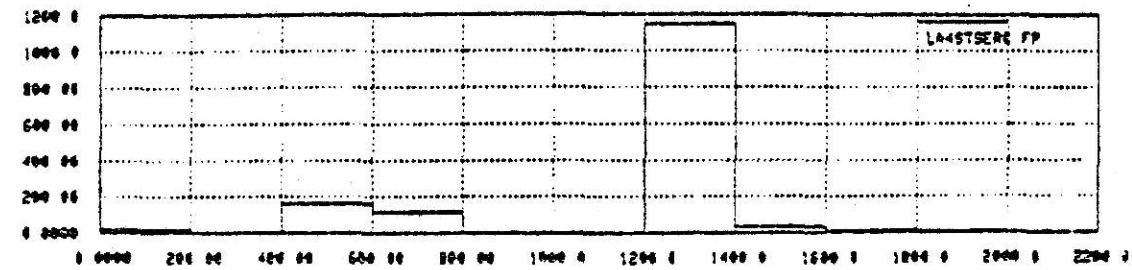
a. input file



b. (WSER) predictor output

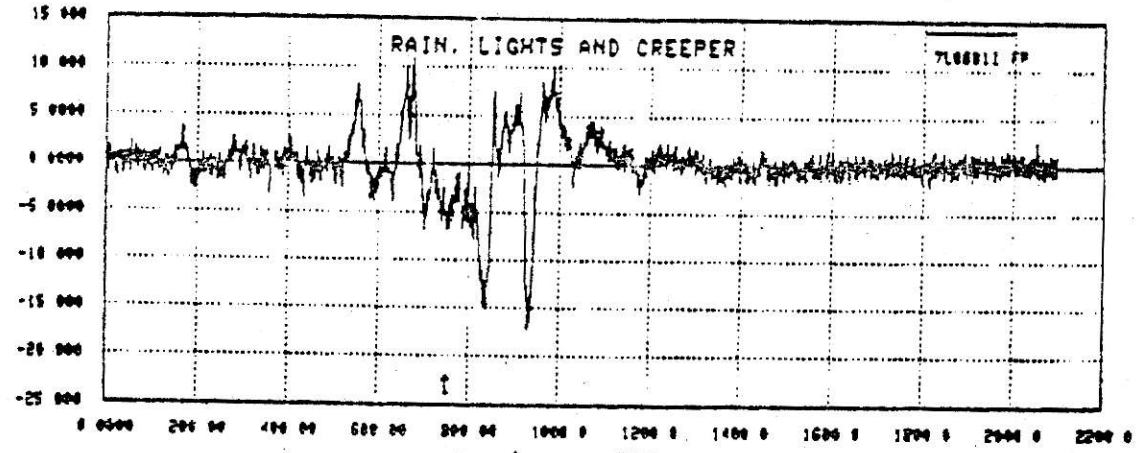


c. (WSER) pred.- MAF comb. output

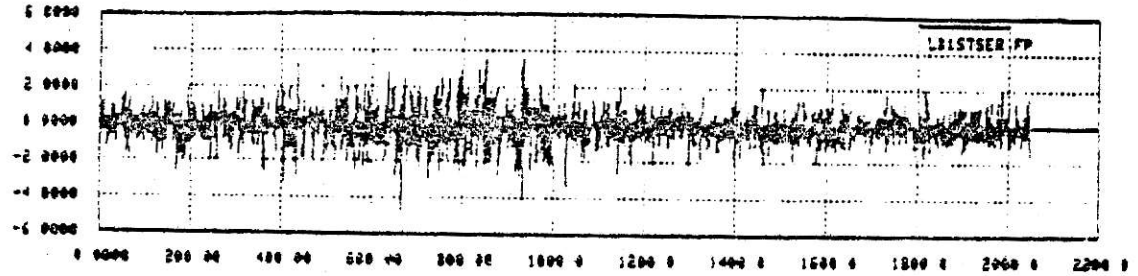


d. (WSER) pred.- X^2 test comb. output

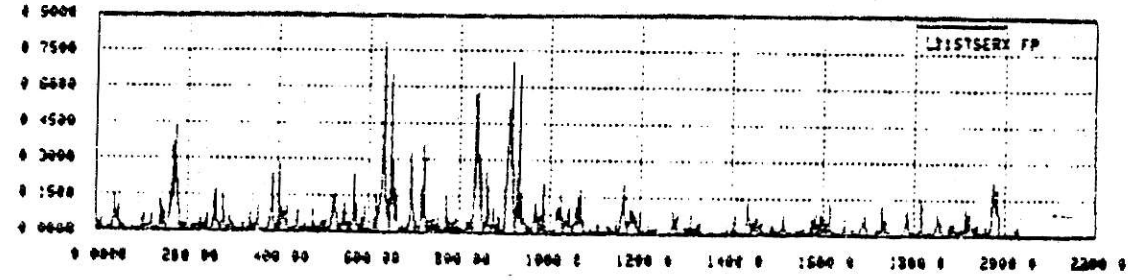
Fig. 4 . 19 Chi-square test performance using WSER pred. (nonmagnetic walker plus road grader).



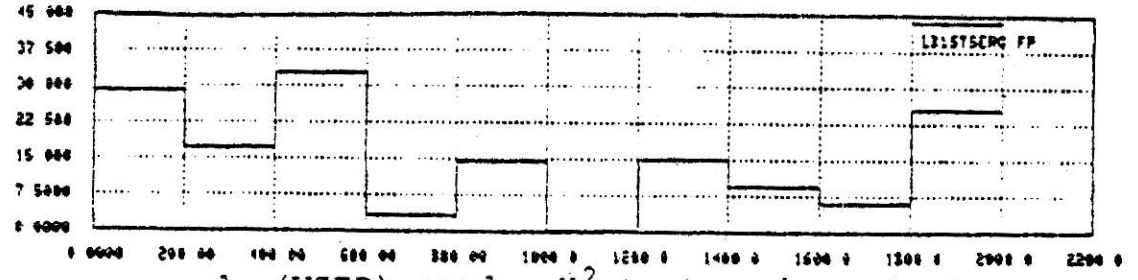
a. input file



b. (WSER) predictor output

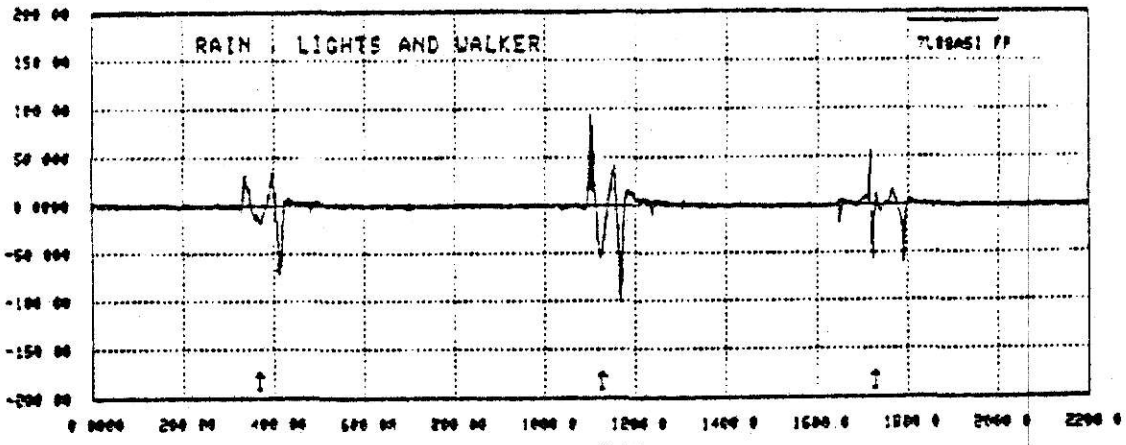


c. (WSER) pred.- MAF comb. output

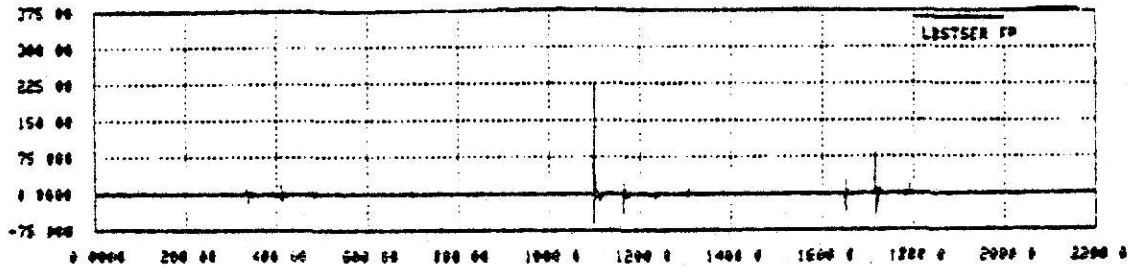


d. (WSER) pred.- X^2 test comb. output

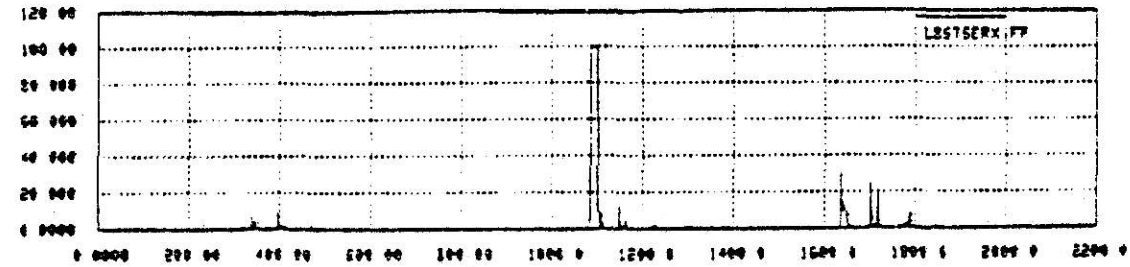
Fig. 4 . 20 Chi-square test performance using WSER pred (creeper plus rain, lights input).



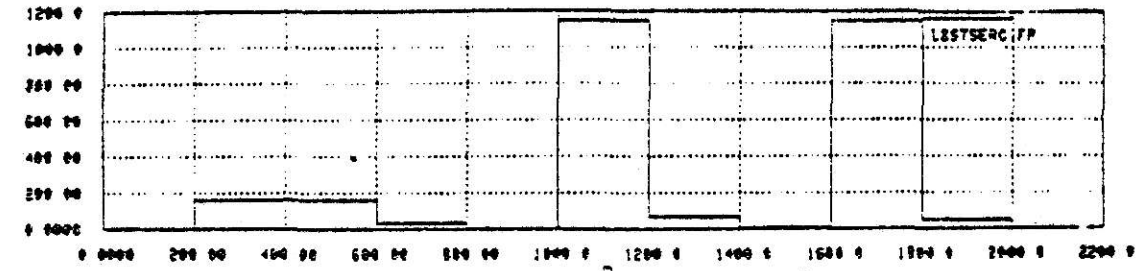
a. input file



b. (WSER) predictor output

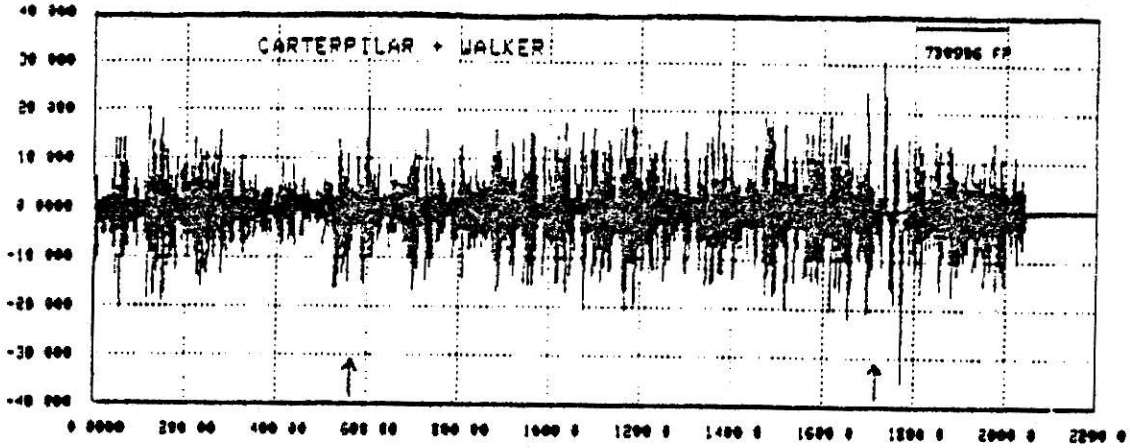


c. (WSER) pred.- MAF comb. output

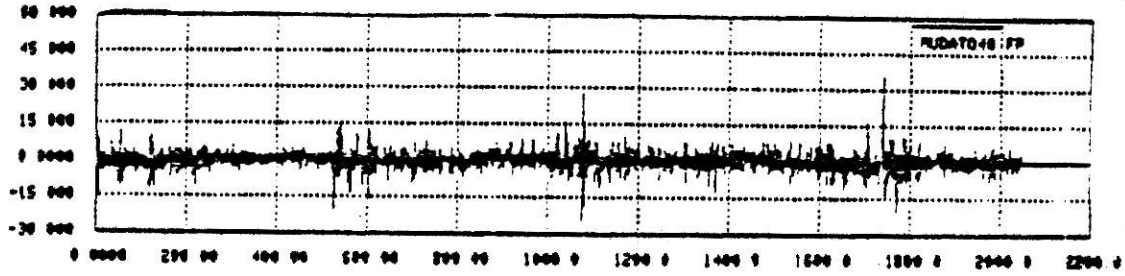


d. (WSER) pred.- X^2 test comb. output

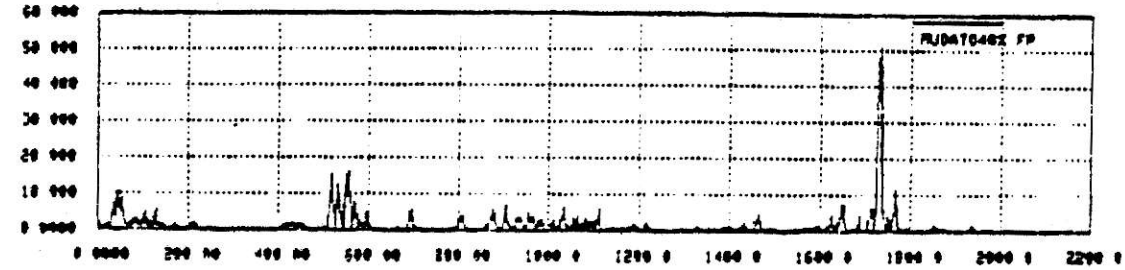
Fig. 4 . 21 Chi-square test performance using WSER pred (walker plus rain, lights input).



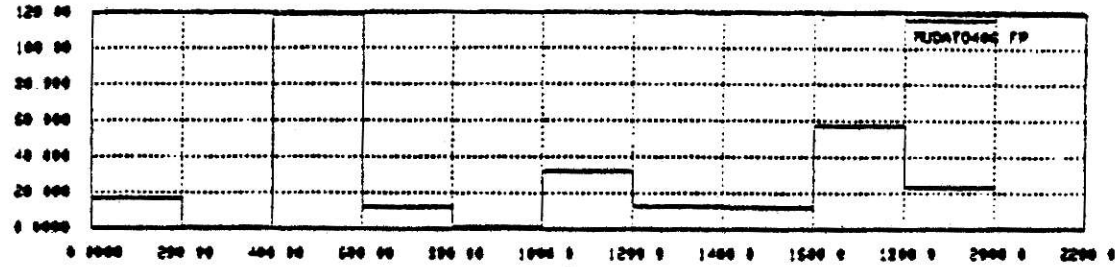
a. input file



b. (WSER) predictor output

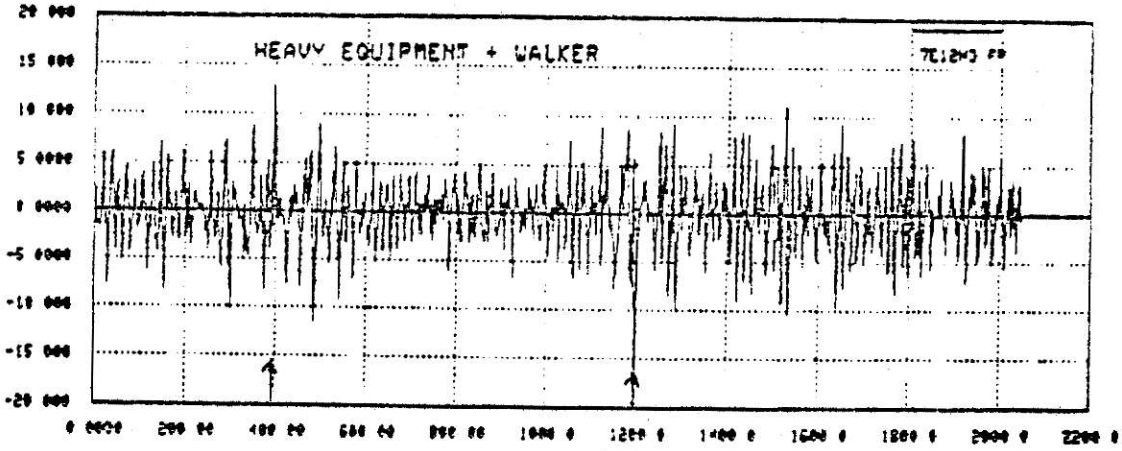


c. (WSER) pred.- MAF comb. output

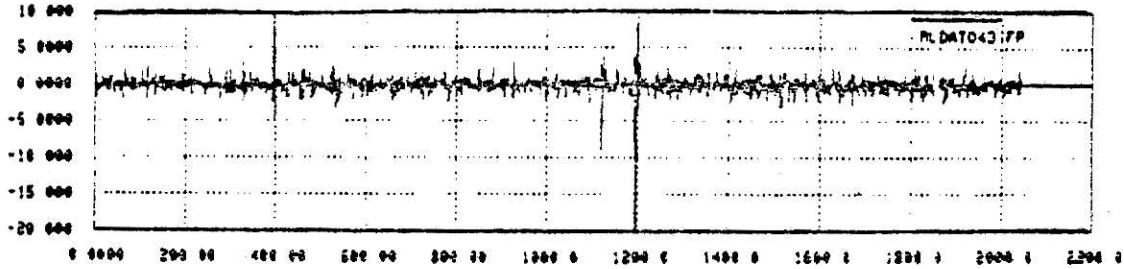


d. (WSER) pred.- X^2 test comb. output

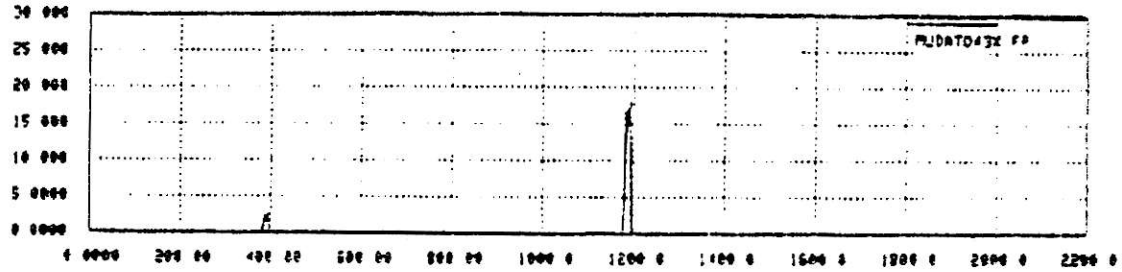
Fig. 4 . 22 Chi-square test performance using LMS pred. (walker plus caterpilar input).



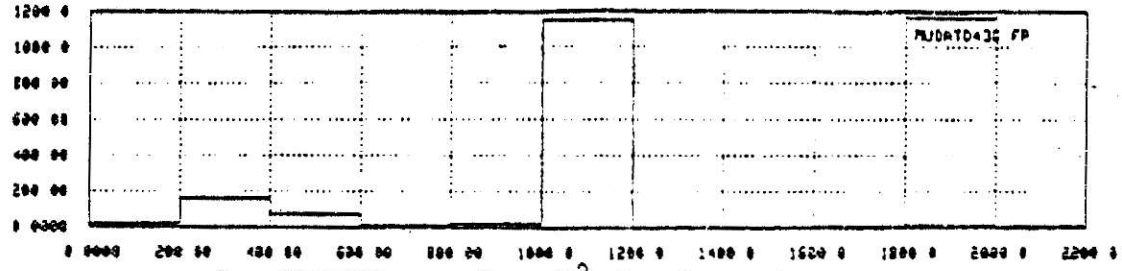
a. input file



b. (WSER) predictor output

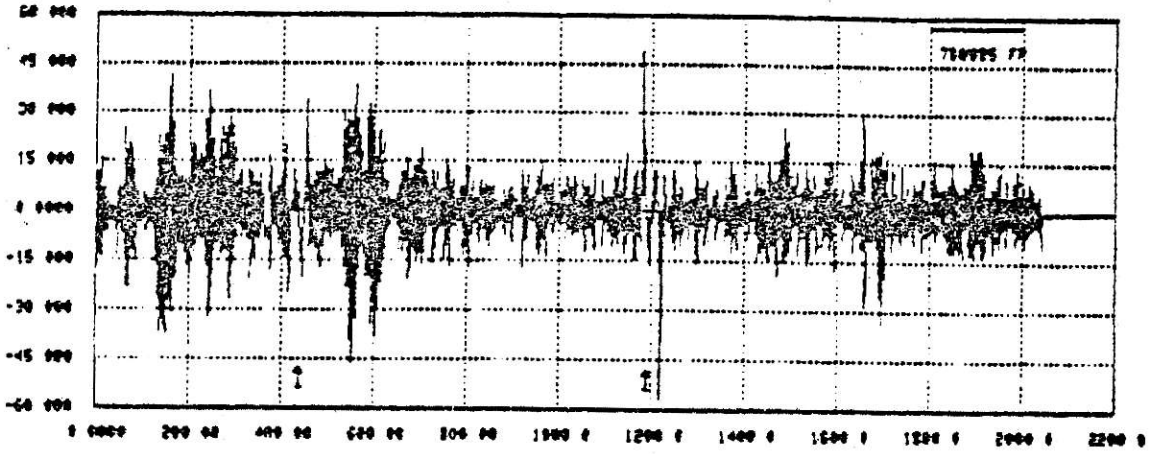


c. (WSER) pred.- MAF comb. output

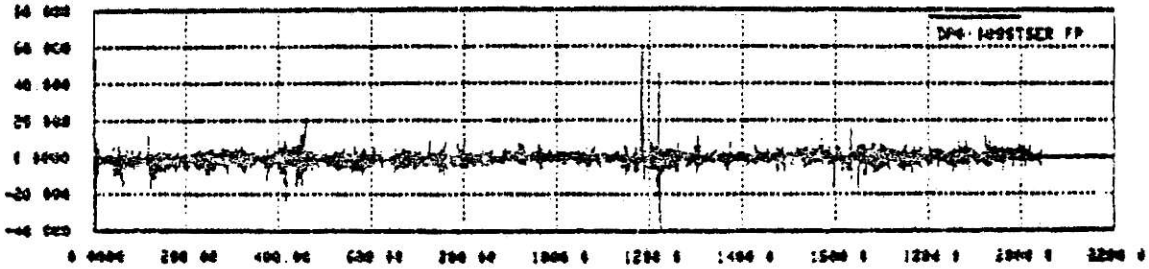


d. (WSER) pred.- X^2 test comb. output

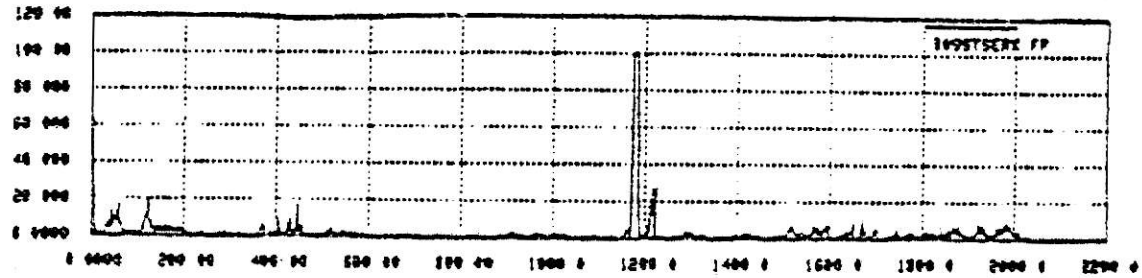
Fig. 4 . 23 Chi-square test performance using WSER pred. (walker plus heavy equipment input).



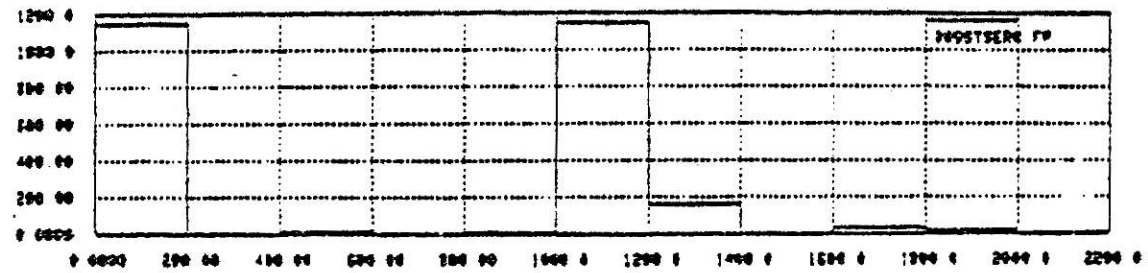
a. input file



b. (WSER) predictor output

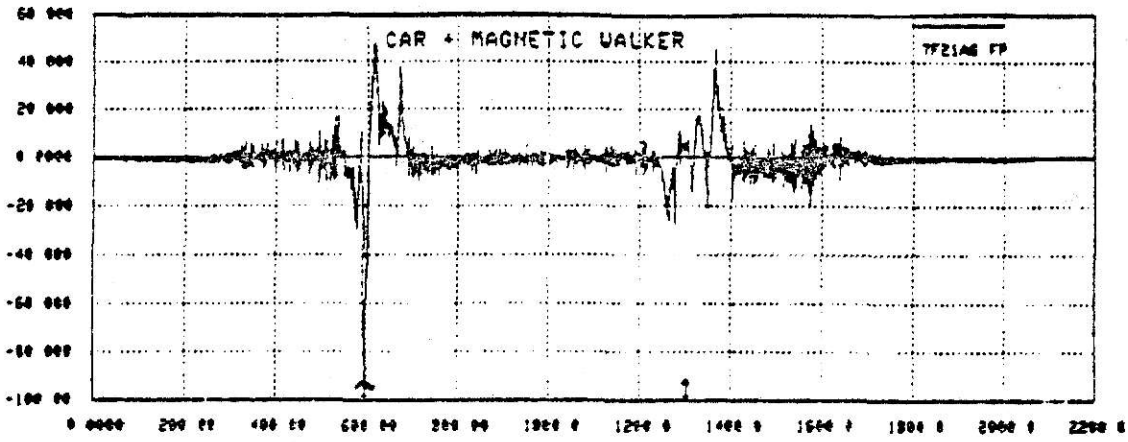


c. (WSER) pred.- MAF comb. output

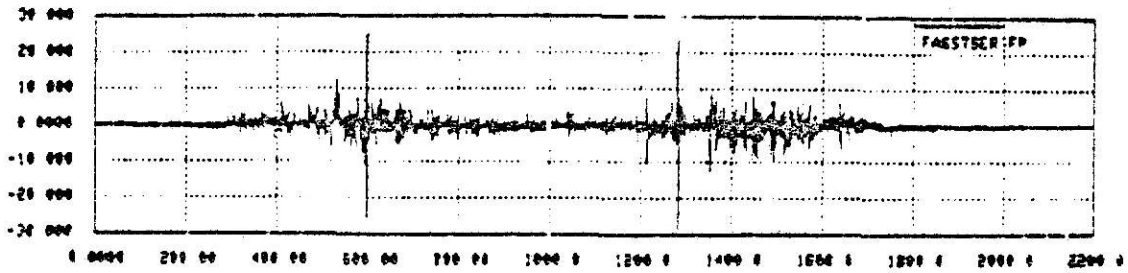


d. (WSER) pred.- X^2 test comb. output

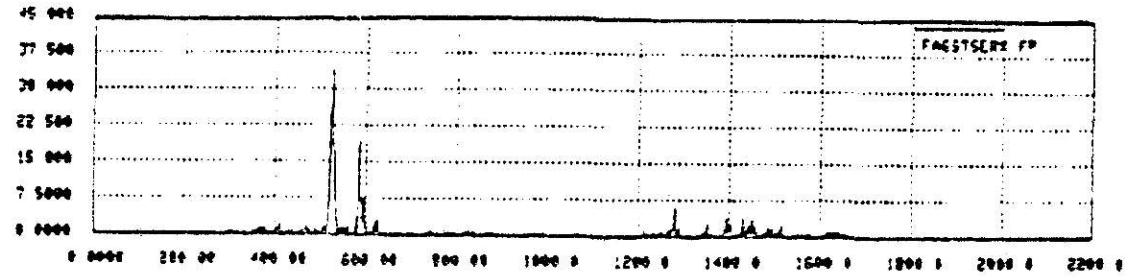
Fig. 4 . 24 Chi-square test performance using WSER pred. (unindentify input).



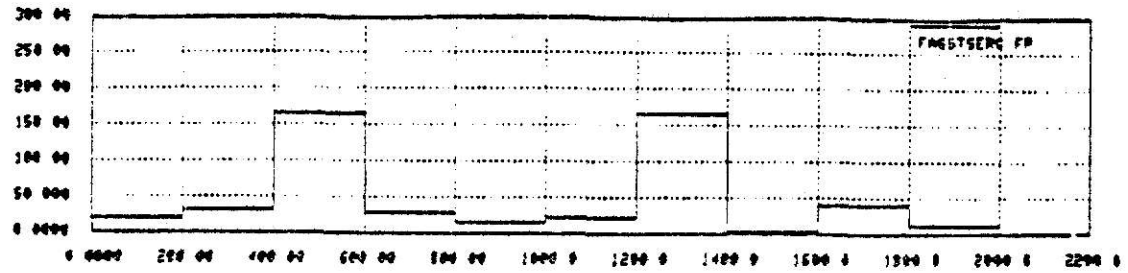
a. input file



b. (WSER) predictor output

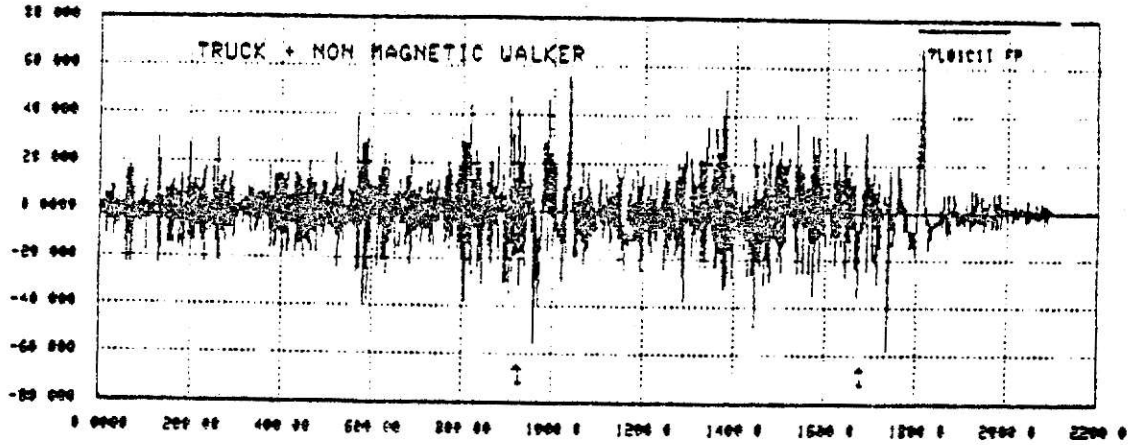


c. (WSER) pred.- MAF comb. output

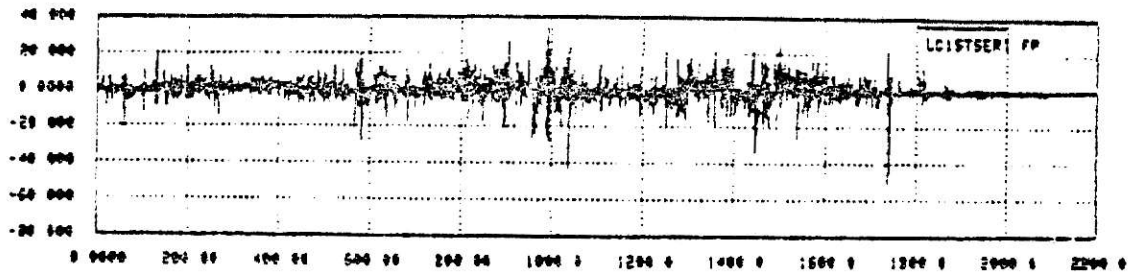


d. (WSER) pred.- X^2 test comb. output

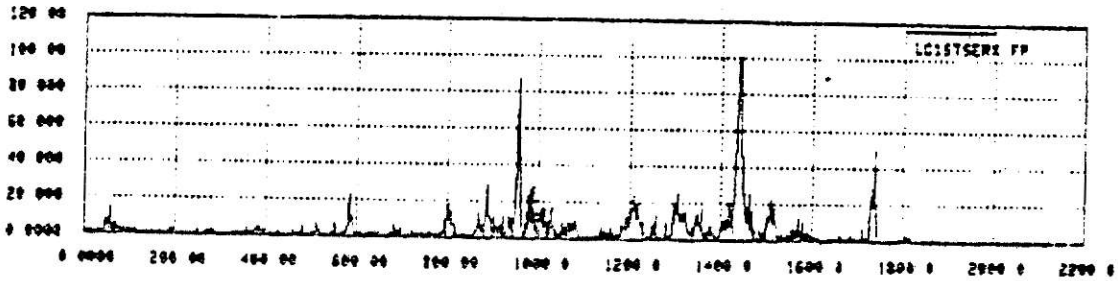
Fig. 4 . 25 Chi-square test performance using WSER pred. (magnetic walker plus car input).



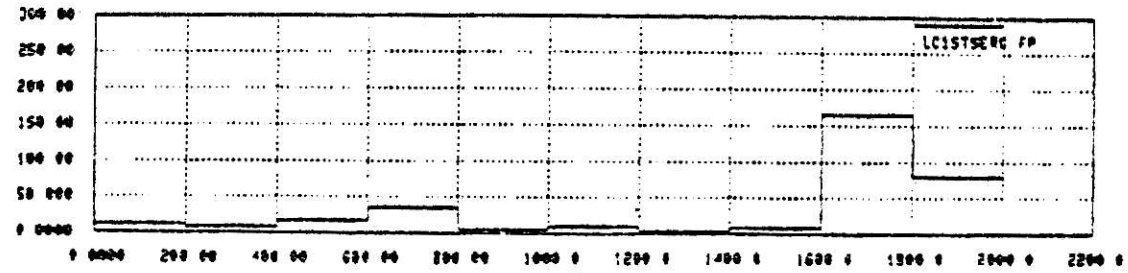
a. input file



b. (WSER) predictor output

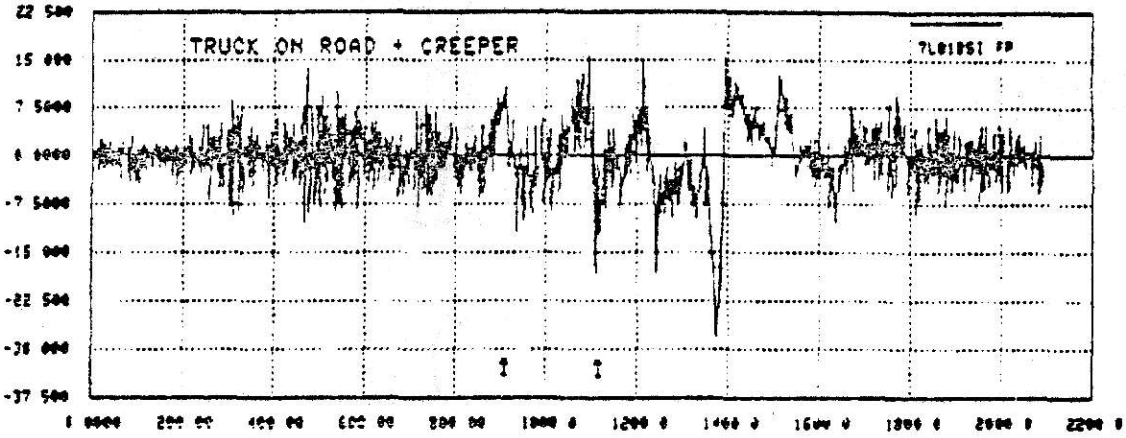


c. (WSER) pred.- MAF comb. output

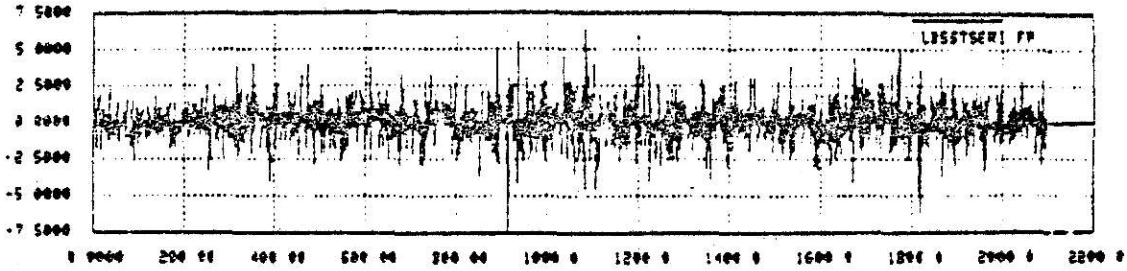


d. (WSER) pred.-X^2 test comb. output

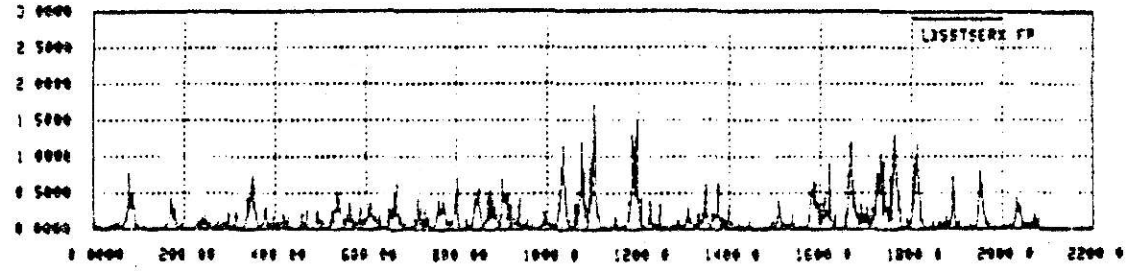
Fig. 4 . 26 Chi-square test performance using WSER pred. (nonmagnetic walker plus truck input).



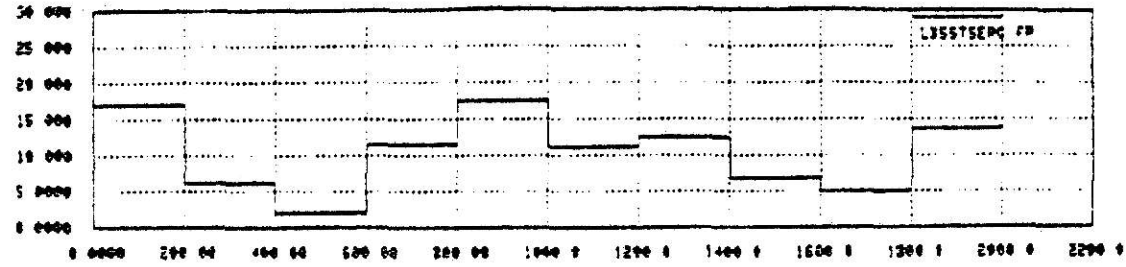
a. input file



b. (WSER) predictor output



c. (WSER) pred.- MAF comb. output



d. (WSER) pred.- X^2 test comb. output

Fig. 4 . 27 Chi-square test performance using WSER pred. (creeper plus truck on road input).

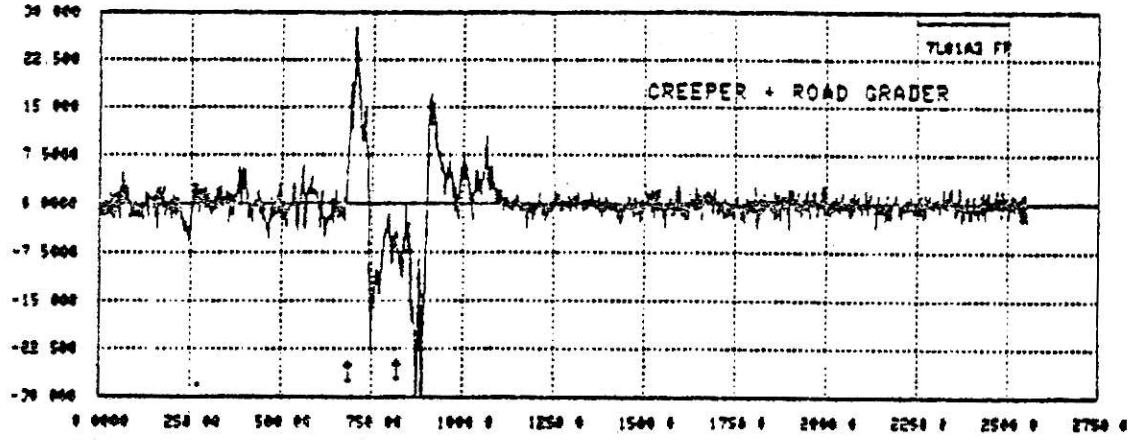
D. PERFORMANCE OF THE LTSER PREDICTOR CHI-SQUARE TEST COMB.

The objective of the experiment described in this section was to evaluate the performance of the Chi-square test approach to intrusion detection when the LTSER algorithm is used for the predictor. The approach taken was to use the ten data files of the two previous tests and to process these files using the LTSER predictor Chi-square test combination.

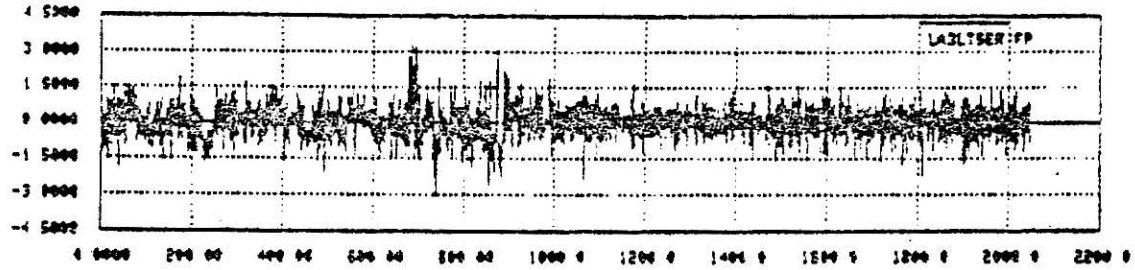
In conducting these tests the values used for the initial diagonal matrix P_0, length of delay Δ, and the number of weights M were I, 1, 4, respectively. The number of data points N in each block, and the number of classes in each block were 200, and 8, respectively.

The results of the tests are shown in Figures 4.28 through 4.37. In each of these illustrations four signals are shown in the same format as was used in the two previous sections.

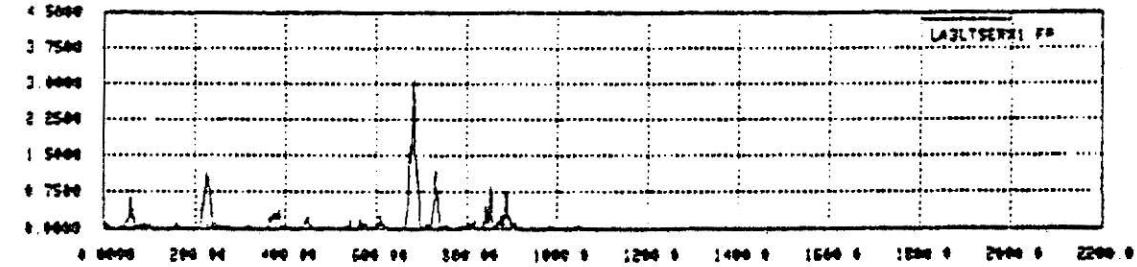
A comparison with the results of the two previous sections indicates that the performance of using the LTSER algorithm is superior to that obtained with the LMS and WSER algorithms. It can also be argued that it is competitive when compared to the MAF approach. In making the comparison there are some additional factors to consider. The LTSER Chi-square approach offers the advantage of being responsive primarily to statistics of the input data rather than its relative amplitude. This is offsetby the fact that the complexity of the approach is several times that of the MAF system.



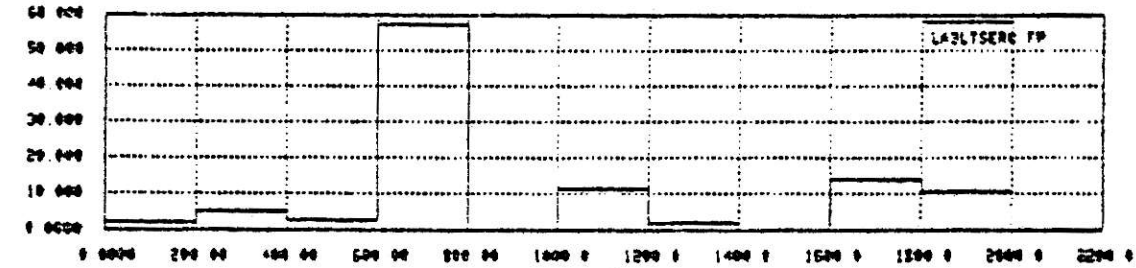
a. input file



b. (LTSER) predictor output

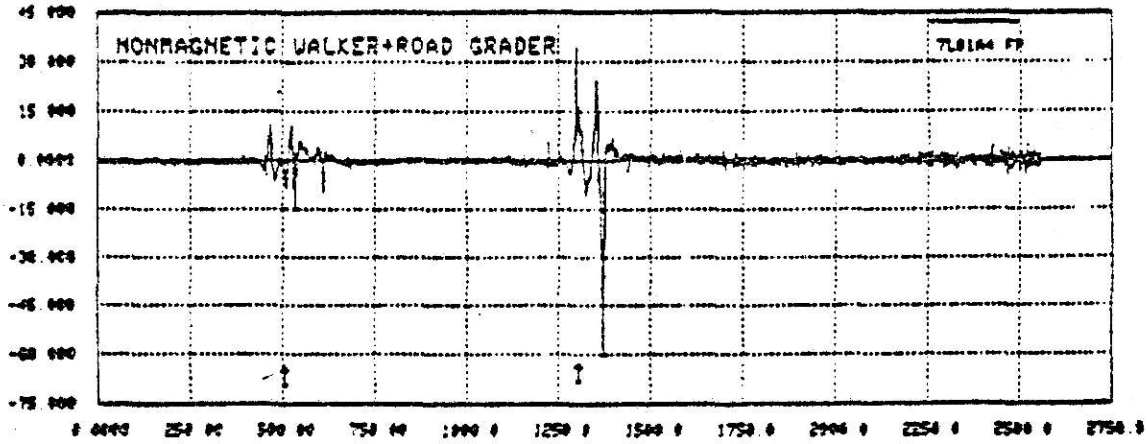


c. (LTSER) pred.- MAF comb. output

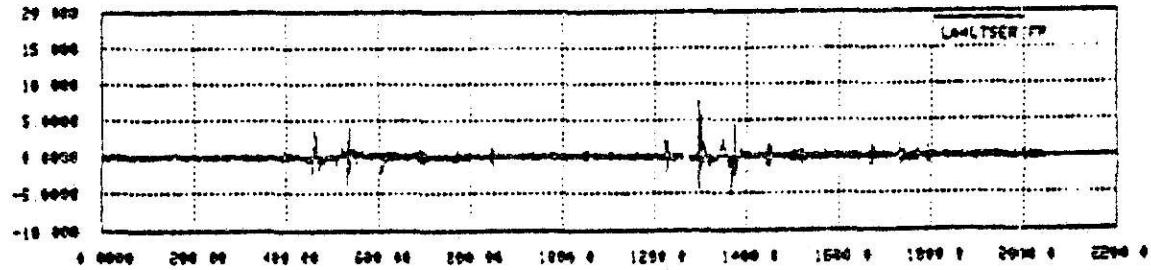


d. (LTSER) pred.- X^2 test comb. output

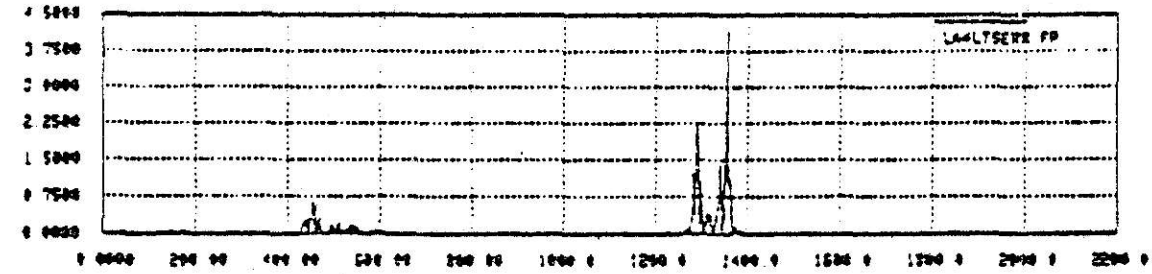
Fig. 4 . 28 Chi-square test performance using LTSER pred. (creeper plus road grader input).



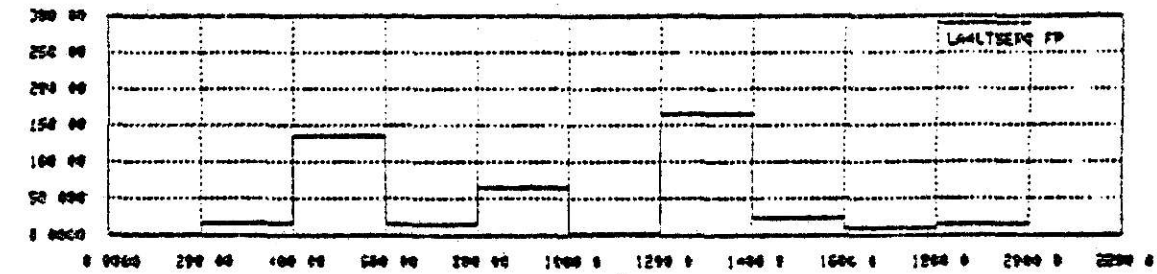
a. Input file



b. (LTSER) predictor output

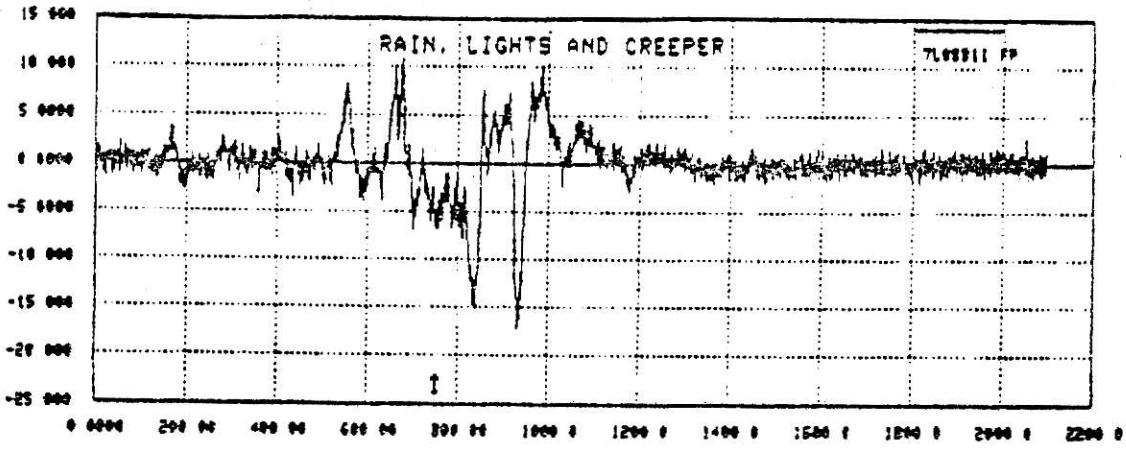


c. (LTSER) pred.- MAF comb. output

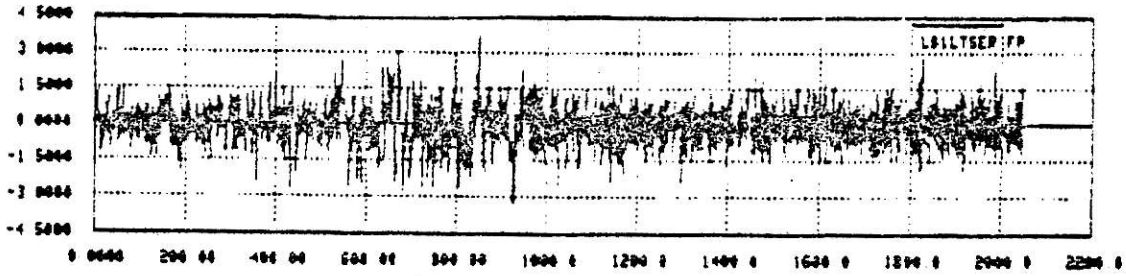


d. (LTSER) pred.- X^2 test comb. output

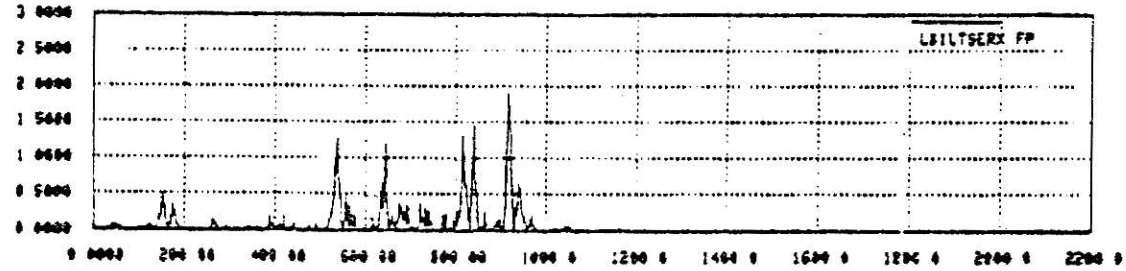
Fig. 4 . 29 Chi-square test performance using LTSER pred. (nonmagnetic walker plus road grader input).



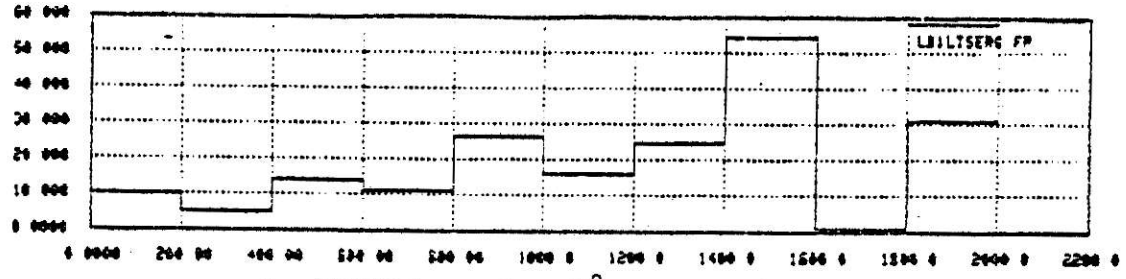
a. input file



b. (LTSER) predictor output

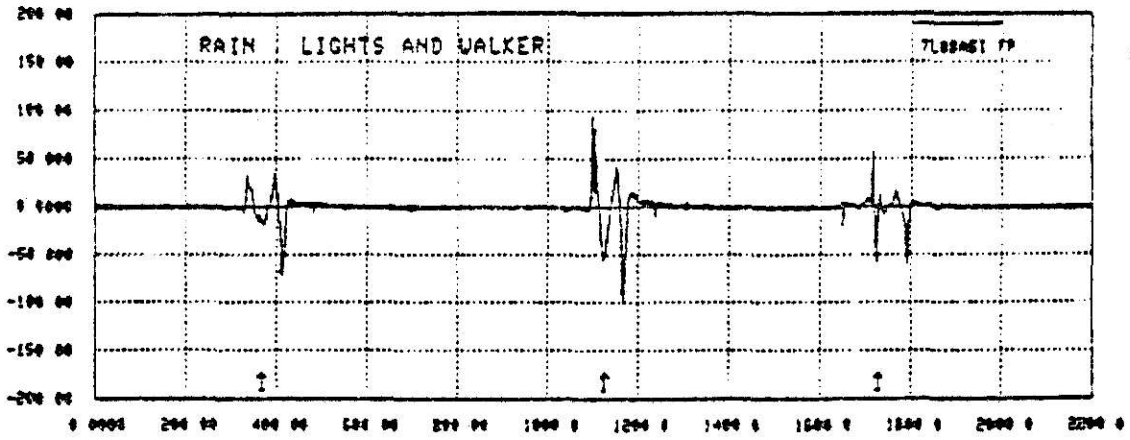


c. (LTSER) pred.- MAF comb. output

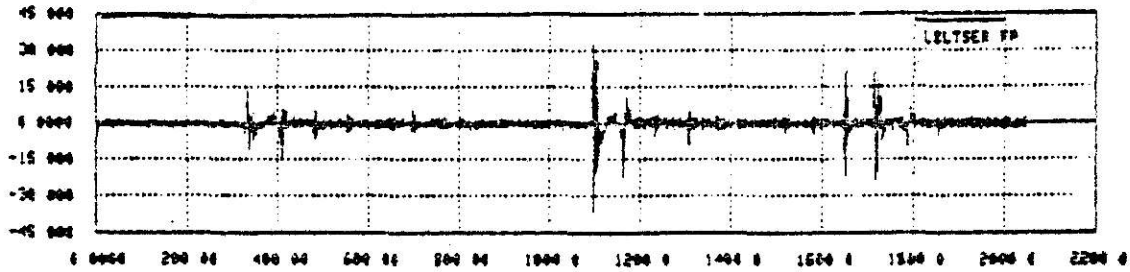


d. (LTSER) pred.- X^2 test comb. output

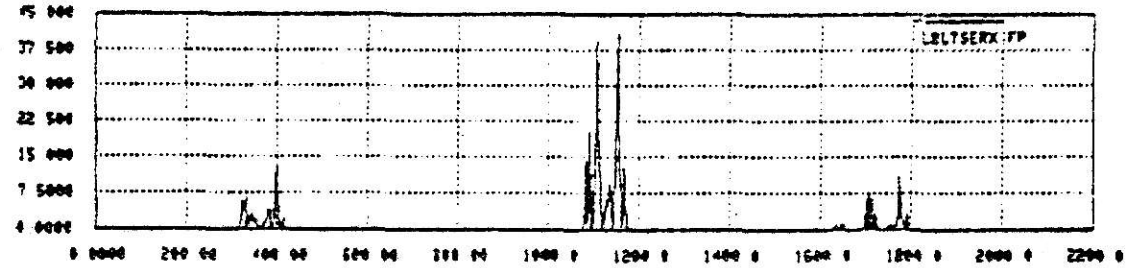
Fig. 4 . 30 Chi-square test performance using LTSER pred. (creeper plus rain, lights input).



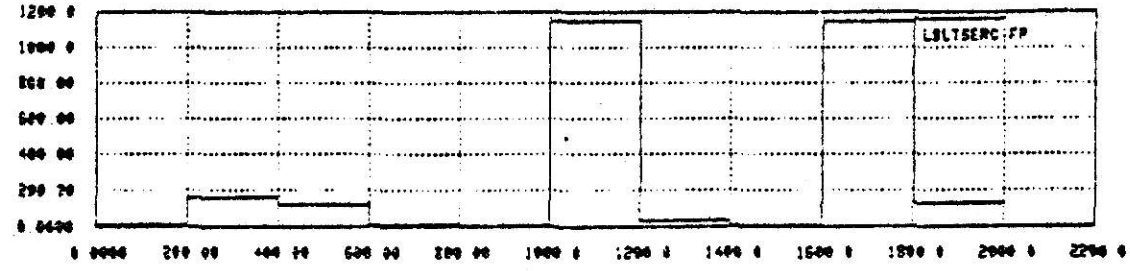
a. input file



b. (LTSER) predictor output

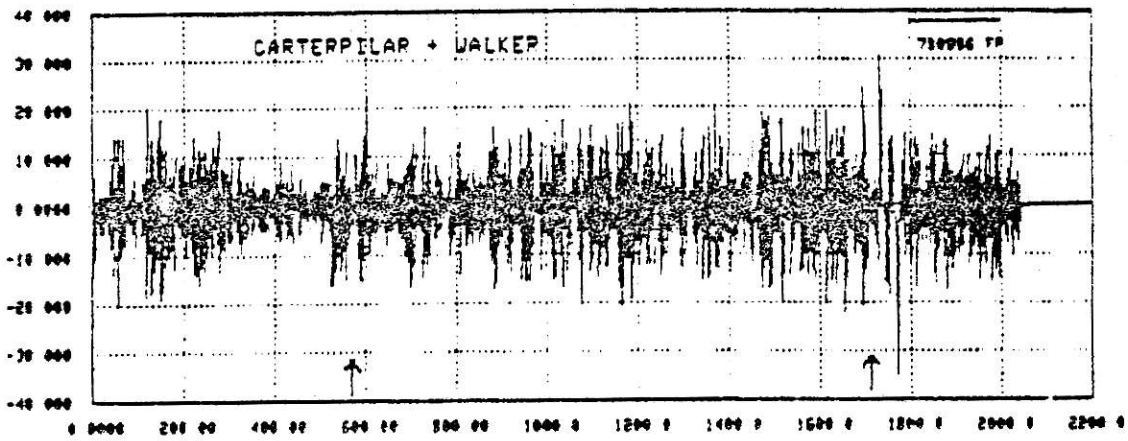


c. (LTSER) pred.- MAF comb. output

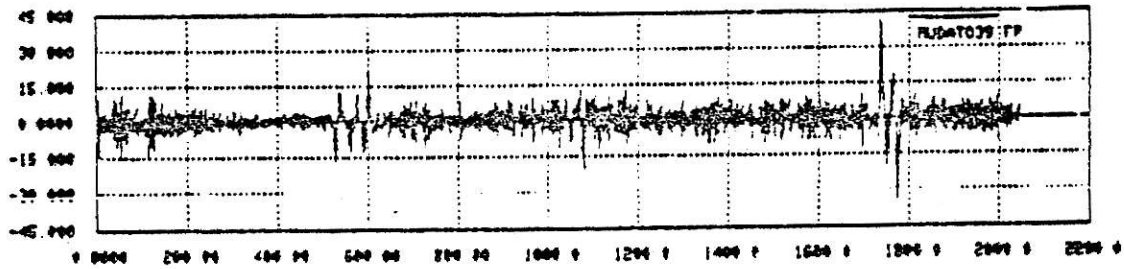


d. (LTSER) pred.- X^2 test comb. output

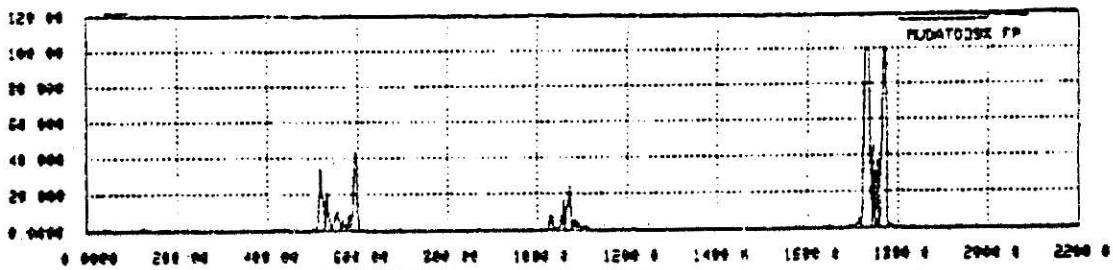
Fig. 4 . 31 Chi-square test performance using LTSER pred. (walker plus rain, lights input).



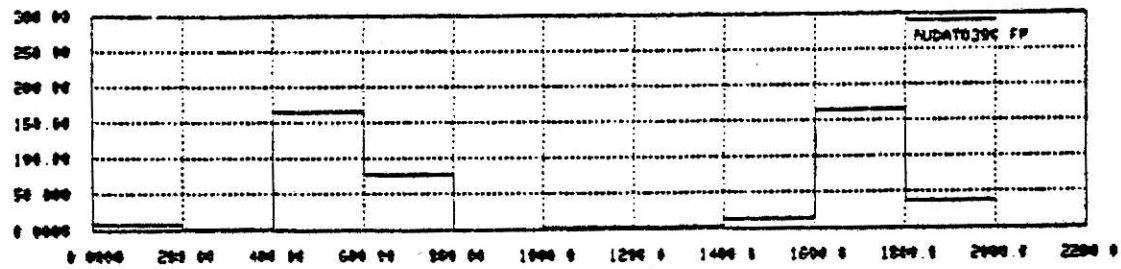
a. input file



b. (LTSER) predictor output

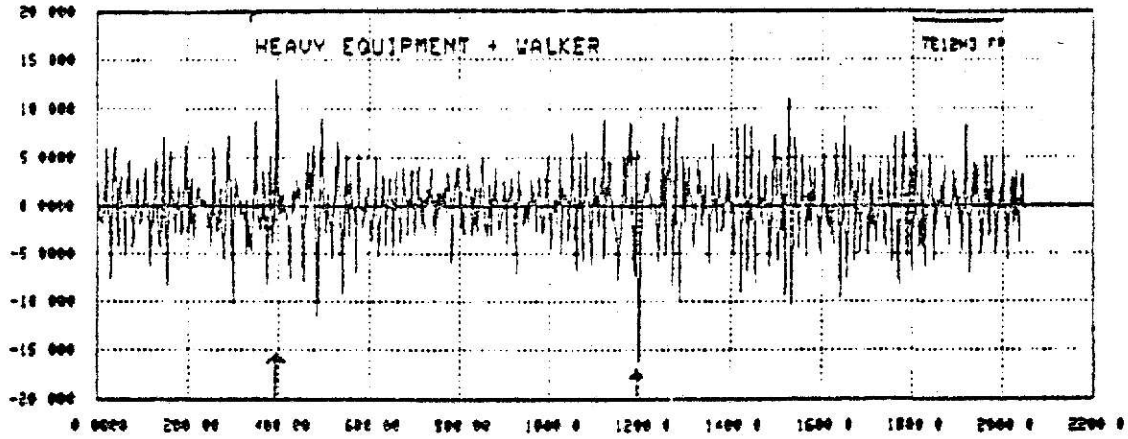


c. (LTSER) pred.-MAF comb. output

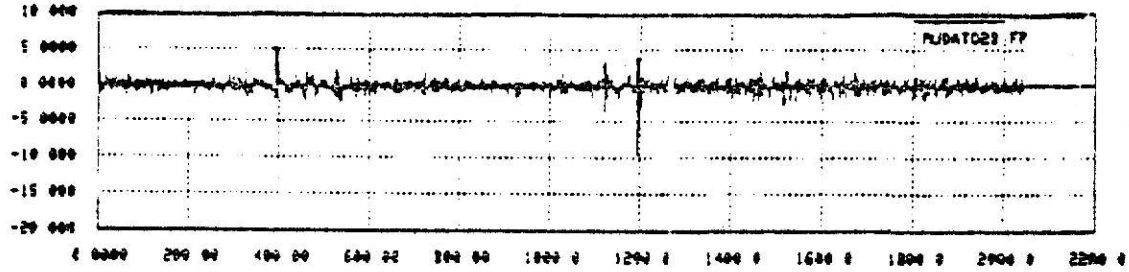


d. (LTSER) pred.- X^2 test output

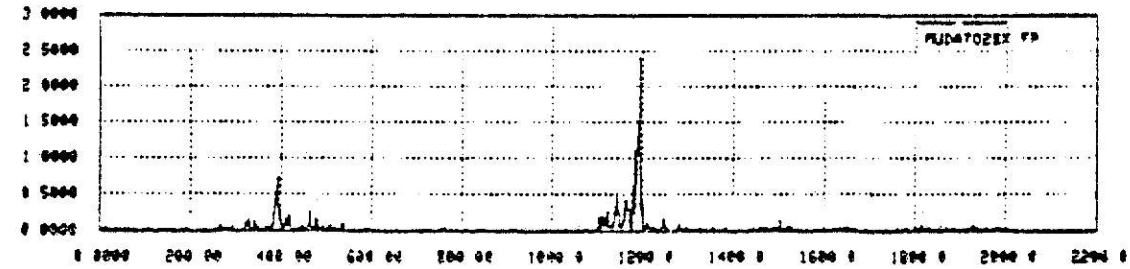
Fig. 4 . 32 Chi-square test performance using LTSER pred (walker plus caterpilar input).



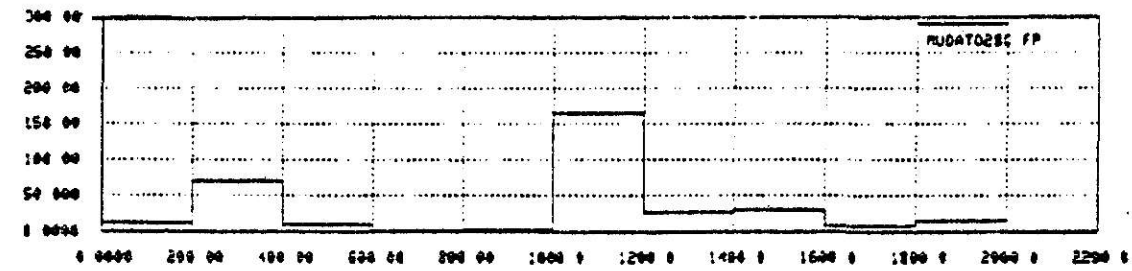
a. input file



b. (LTSER) predictor output

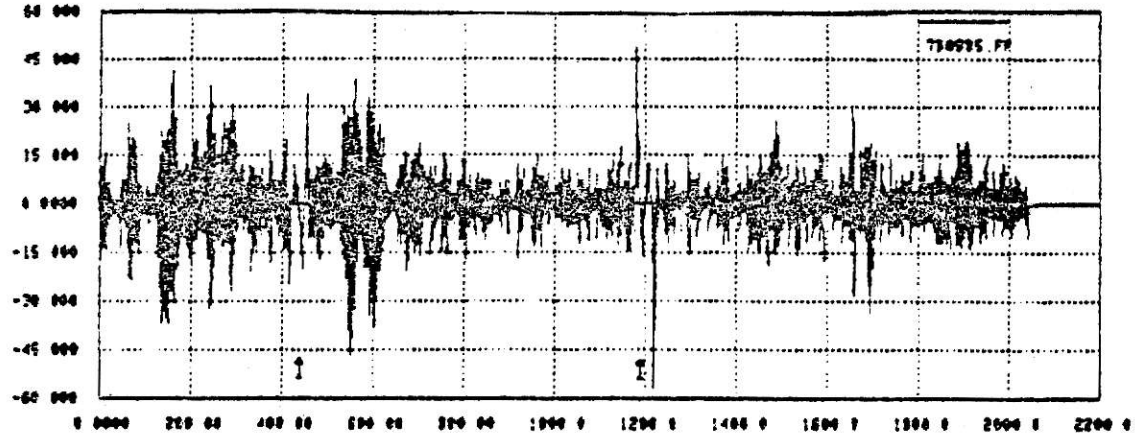


c. (LTSER) pred.- MAF comb. output

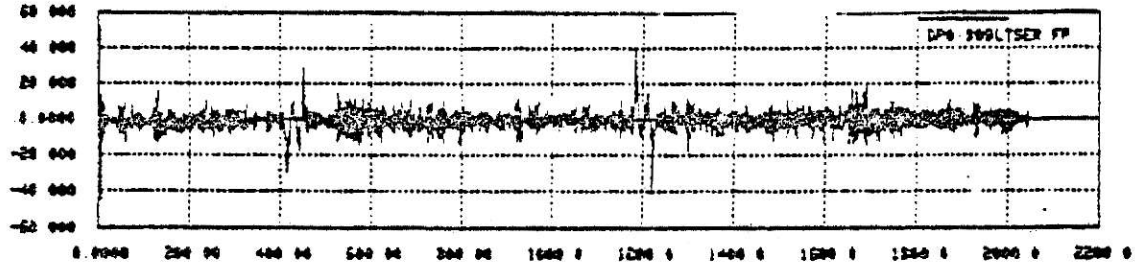


d. (LTSER) pred.- X^2 test comb. output

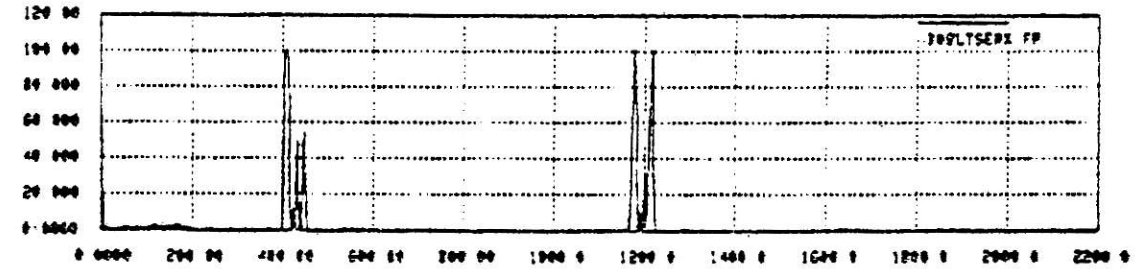
Fig. 4 . 33 Chi-square test performance using LTSER pred (walker plus heavy equipment input).



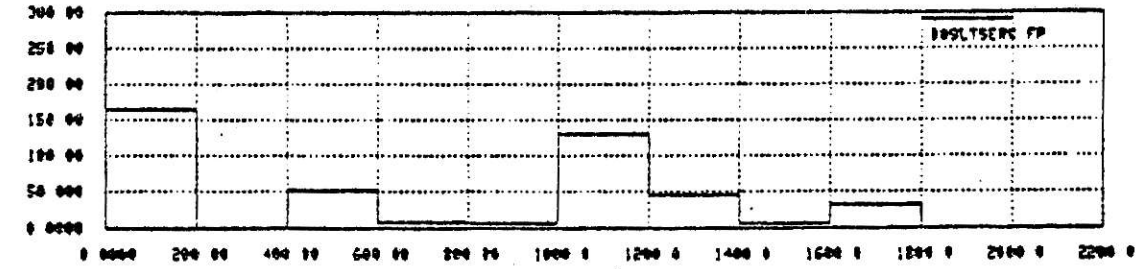
a. input file



b. (LTSER) predictor output

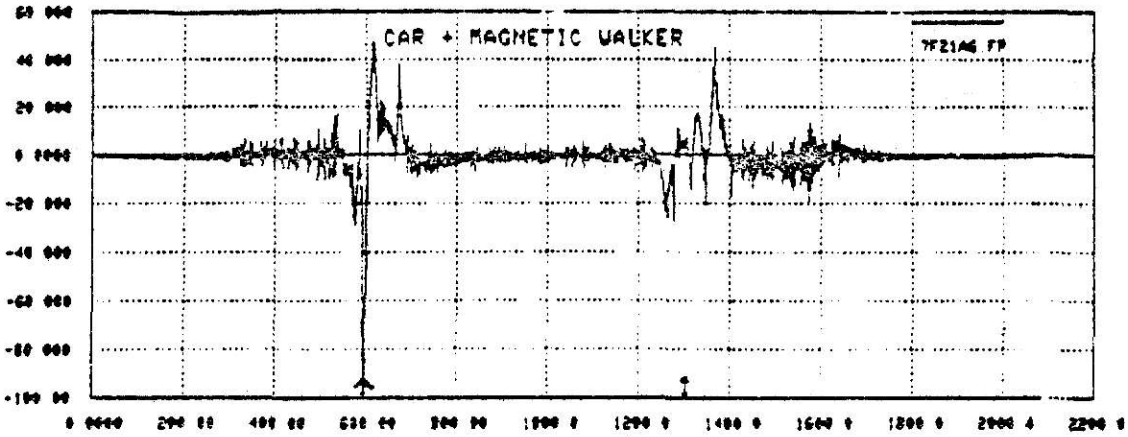


c. (LTSER) pred.- MAF comb. output

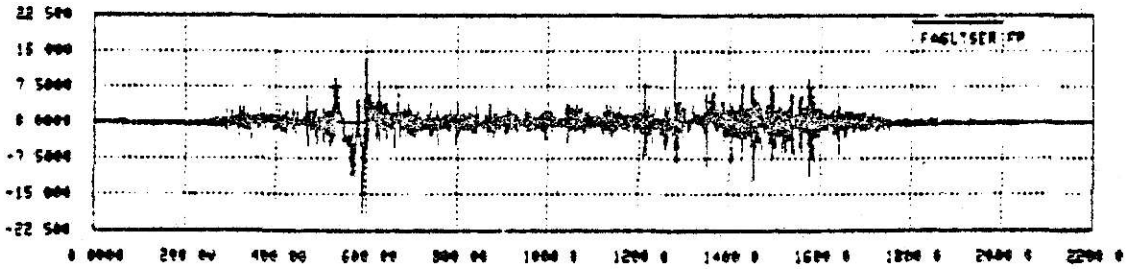


d. (LTSER) pred.- X^2 test comb. output

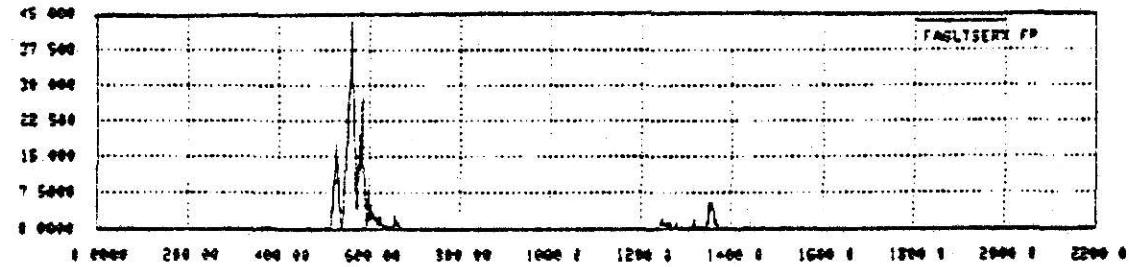
Fig. 4 . 34 Chi-square test performance using LTSER pred (unindentify input).



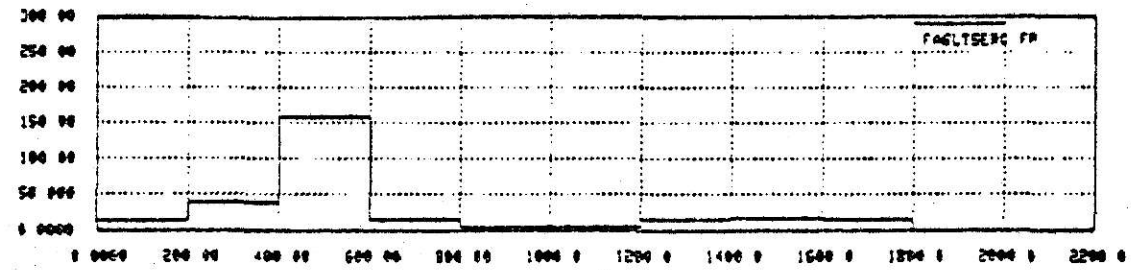
a. input file



b. (LTSER) predictor output

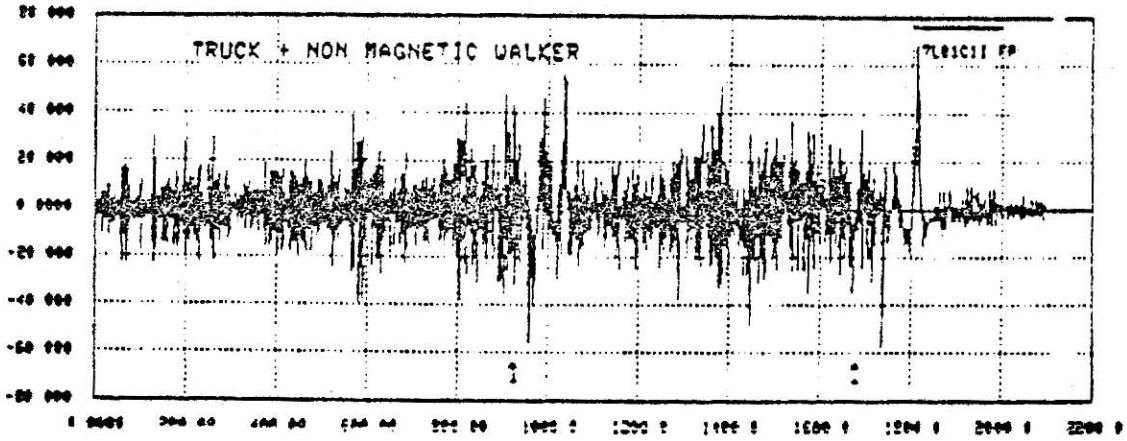


c. (LTSER) pred.- MAF comb. output

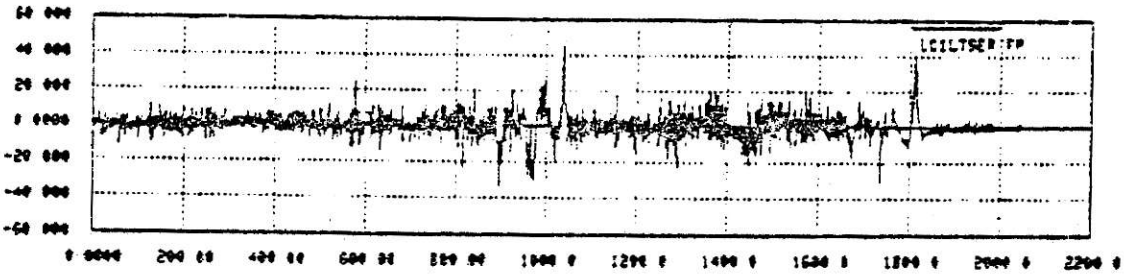


d. (LTSER) pred.- X^2 test comb. output

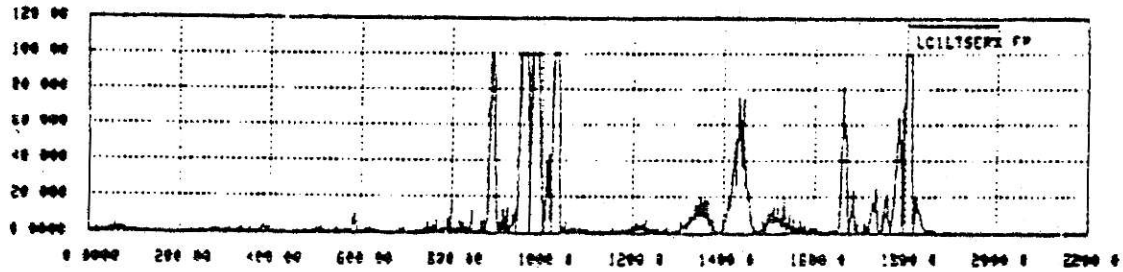
Fig. 4 . 35 Chi-square test performance using LTSER pred (magnetic walker plus car input).



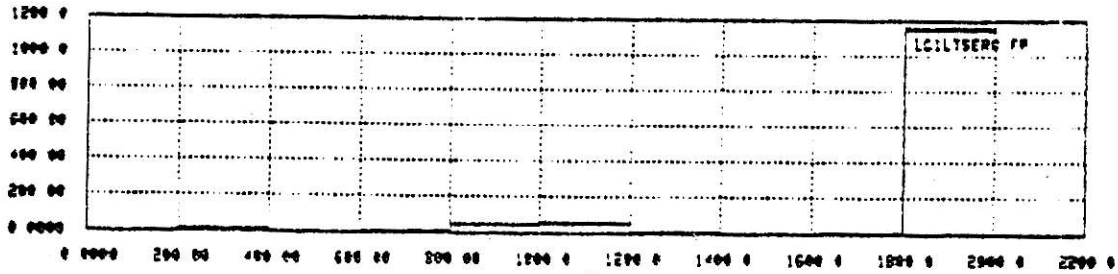
a. input file



b. (LTSER) predictor output

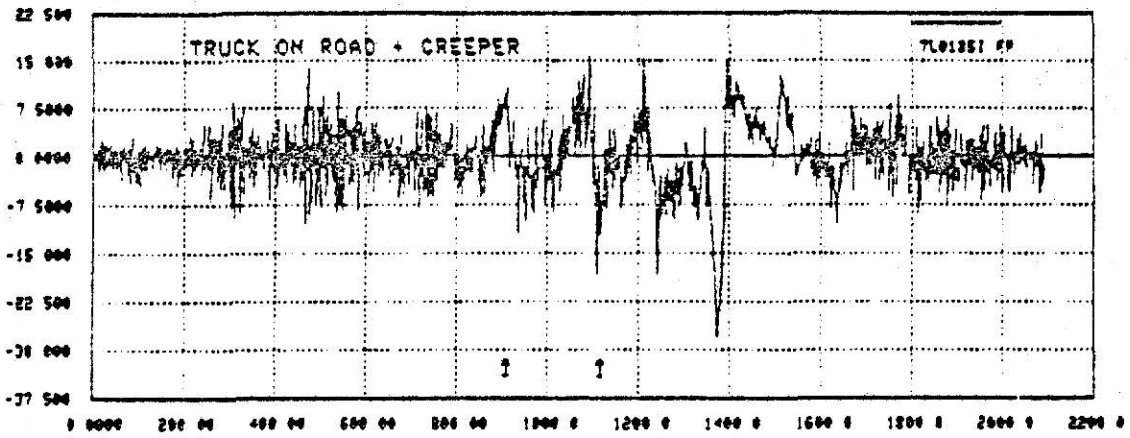


c. (LTSER) pred.- MAF comb. output

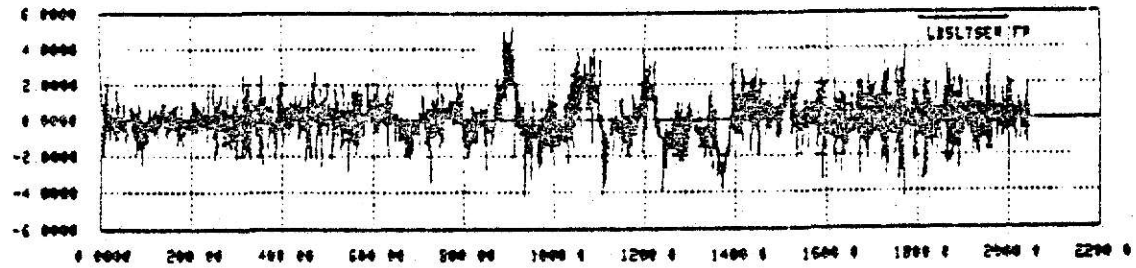


d. (LTSER) pred.- X^2 test comb. output

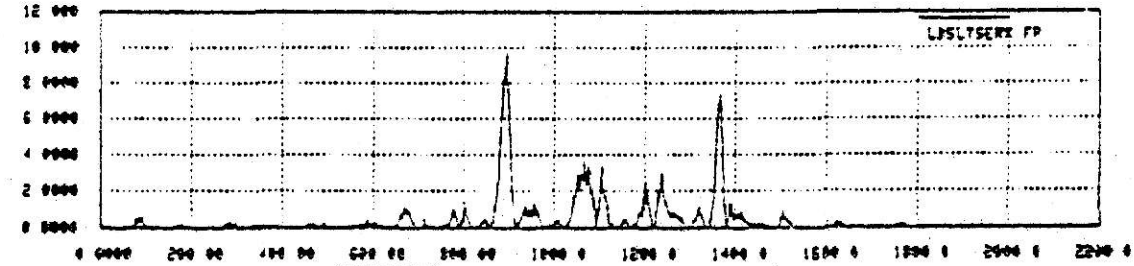
Fig. 4 . 36 Chi-square test performance using LTSER pred. (nonmagnetic walker plus truck input).



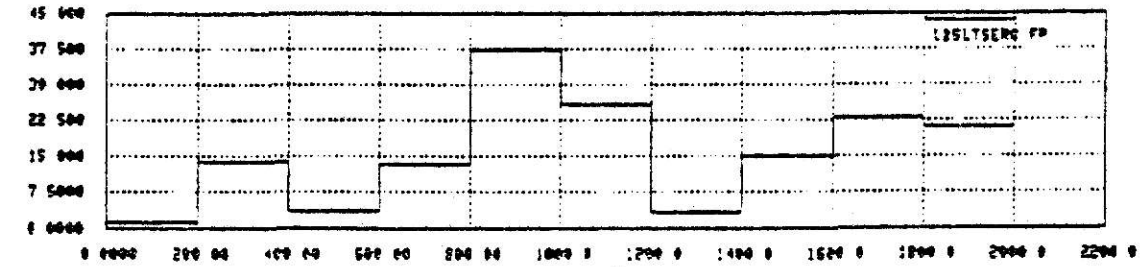
a. input file



b. (LTSER) predictor output



c. (LTSER) pred.- MAF comb. output



d. (LTSER) pred.- X^2 test comb. output

Fig. 4 . 37 Chi-square test performance using LTSER pred (creeper plus truck on road input).

Chapter 5

CONCLUSIONS

We have shown that the concept of using the predictor Chi-square test combination for intrusion detection purposes is promising. It is shown that the number of data points N per block equal to 200 points is a moderate number to detect an intuder in this problem. The LMS, and the LTSER predictor Chi-square test combinations gave more promising results in detecting an intruder than the WSER predictor Chi-square test combination.

If we hold to the Chi-square test definition (the confidence level must be at least 0.05 to accept the hypothesis) then for relatively short sequences of the output of the predictor the sample are Gaussian in an intruder's absence, and non-Gaussian in intruder's presence.

There is a restriction in using the predictor Chi-square test combination that the first block of the combination output should not be counted, because the first block of the predictor output tends to be non-Gaussian most of the time. This is due to the transient that occurs when the predictor initially adapts to the data.

The interesting feature of the predictor Chi-square test combination is that the Chi-square test detects the presence or the absence of an intruder by using the statistical characteristics of the predictor output.

APPENDIX

```
C     ***************************************
C
C
C                    CHITEST
C
C
C     PURPOSE :
C     TO DETECT CREEPER BY USING CHI-SQUARE DISTRIBUTION
C     AND USING THE APPROXIMATION TO THE NORMAL
C     DISTRIBUTION FUNCTION
C
C                P=.5+.5(1.- EXP(-2X**2/PAI ))
C
C     AS THE EXPECTED VALUES OF THE SYSTEM
C
C          BY  R.D. SUMANTRI              FALL '78
C
C        KANSAS STATE UNIVERSITY
C
C
C     ***************************************
      DIMENSION Y(2200),F(50),Z1(100),CO(2200)
      DIMENSION E(50),PZ(100),X1(100)
1     WRITE (10,11)
11    FORMAT ("0 INPUT,# OF SAMPLE PER BLOCK ? ",Z)
      READ (11)N
      WRITE (10,12)
12    FORMAT ("0 INPUT, # OF CLASSES PER BLOCK(EVEN#) ? ",Z)
      READ (11)K
      WRITE (10,13)
13    FORMAT ("0 INPUT, # OF BLOCKS PER FILE  ?   ",Z)
      READ (11)NN
      MM=N
      JJ=1
      CALL IOPEN (0,1,2,1,0,'INPUT FILE?')
      CALL IOPEN (1,3,2,1,0,'OUTPUT FILE ? ')
      DO 10 J=1,NN
      G1=0.
      G2=0.
      DO 20 I=JJ,MM
      READ BINARY (0,ERR=9,END=999)Y(I)
20    G1=G1+Y(I)
      U=G1/FLOAT(N)
      DO 30 I=JJ,MM
      V1=ABS(Y(I)-U)
30    G2=G2+V1**2.
      V2=G2/FLOAT(N)
      V2=SQRT(V2)
      L=K/2-1
      K1=K/2
      Z1(K1)=U
      DO 40 I=1,L
```

```
      Z1(K1-I)=-(.50*I*(ABS(V2)))+U
40    Z1(K1+I)=.50*I*(ABS(V2))+U
      L1=K-1
      F(1)=0.
      F(K)=0.
      DO 50 I=2,L1
      F(I)=0.
      DO 60 M=JJ,MM
      IF (Y(M).GE.Z1(I))GO TO 60
      IF (Y(M).LT.Z1(I-1))GO TO 60
      F(I)=F(I)+1.
60    CONTINUE
50    CONTINUE
      DO 67 I=JJ,MM
      IF(Y(I).GE.Z1(1)) GO TO  65
      F(1)=F(1)+1.
      GO TO 67
65    IF (Y(I).LT.Z1(L1))GO TO 67
      F(K)=F(K)+1.
67    CONTINUE
      LL=K+1
      KK=K/2+1
      DO 70 I=K1,L1
      WW=(Z1(I)-U)/V2
      AX=ABS(WW)
      AY=(-2.*(AX)**2.)/3.14159265
      AZ=EXP(AY)
      PZ1=.5+.5*SQRT(1.-AZ)
      PZ(I)=ABS(PZ1)
70    CONTINUE
      DO 80 I=KK,L1
      A=ABS(PZ(I)-PZ(I-1))
      E(I)=A*(FLOAT(N))
      E(LL-I)=E(I)
      E(K)=ABS(1.-PZ(L1))*FLOAT(N)
      E(1)=E(K)
      IF (E(1).LT.5) GO TO 99
80    CONTINUE
      X2=0.
      DO 90 I=1,K
      X1(I)=(ABS(F(I)-E(I))**2.)/E(I)
      X2=X2+X1(I)
90    CONTINUE
      L2=L1-1
      X=X2
      G=FLOAT(K)-3.
      CALL CDTR(X,G,P,D,IER)
      SIG=1.-P
      IF (SIG.GT.0.0) GO TO 101
      SIG=1.E-50
101   IF (SIG.LT.0.999999999) GO TO 102
      SIG=1.0
102   DO 110 I=JJ,MM
      CO(I)=-10.*ALOG(SIG)
110   WRITE BINARY (1)CO(I)
      MM=MM+N
      JJ=JJ+N
10    CONTINUE
      CALL CLOSE (0,IERR)
      CALL CLOSE (1,IERR)
      CALL QUERY (' RE-PROGRAM ? ',I)
      IF (I.EQ.1) GO TO 1
      STOP
99    TYPE' K IS TO BIG '
      STOP
9     TYPE*ERROR IN READ'
999   CONTINUE
      STOP
      END
```

References

[1] N. Ahmed, G.R. Elliott, N.A. Bourgeois, R.J. Fogler, D.L. Soldan, "An Intrusion Detection Approach Via Adaptive Prediction", Kansas State University.

[2] Henry L. Alder, E.B. Roessler, "Introduction to Probability and Statistics", W.H. Freeman and Company, London, 1962.

[3] Meyer Dwass, "Probability and Statistics", W.A. Benjamin Inc. New York, 1970.

[4] M. Abramowitz, L.A. Stegun, "Handbook of Mathematical Functions", U.S. Department of Commerce, National Bureau of Standards, June 1964.

[5] IBM Applications Program, "System/360 Scientific Subroutine Package", Version III, program manual, IBM.

[6] B. Widrow, et. al., "Adaptive Noise Cancelling: Principles and Applications", Proc. IEEE, Vol. 63, No. 12, P. 162, 1975.

[7] N. Ahmed, D.R. Hummels, M. Uhl, "A Weighted Sequential Regression Algorithm", Department of Electrical Engineering, Kansas State University.

[8] D. Soldan, N. Ahmed, "On a Sequential Regression Predictor", Electronics Letters, 16th February 1978.

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank all those who assisted me during my study. I would especially like to thank Dr. D.R. Hummels for his encouragement and guidance. My appreciation is given to Dr. E. Haft, and Dr. K. Stromberg for serving on my committee.

Several persons deserve thanks for answering my questions about the NOVA, particularly Dave Soldan, Dakshesh Parikh, and Dave Hein.

The financial support of the Electrical Engineering Department of Kansas State University has made possible my study and is appreciated.

INTRUSION DETECTION VIA AN ADAPTIVE DIGITAL PREDICTOR CHI-SQUARE TEST COMBINATION

by

RADEN DJAFAR SUMANTRI

B. S., Tokai University, 1976

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Electrical Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1979

ABSTRACT

A study of a new algorithm for intrusion detection has been made. The method uses an adaptive digital predictor Chi-square test combination. The inputs used in this experiment were obtained via a sensor which is in the form of a buried cable. The basic idea of the method is to use the predictor as a noise decorrelator. The resulting detection problem is then reduced to that of determining the presence or the absence of an intruder in a noise background which essentially is white. There are three different predictors used in the work reported here, (i) Widrow's least mean square, (ii) the weighted sequential regression, and (iii) the long term sequential regression predictor. For each of these three predictors, some predictor χ^2 test combination performance results are included.