

AD-A067 771

TEXAS A AND M UNIV COLLEGE STATION INST OF STATISTICS
TIME SERIES ANALYSIS OF LH HORMONE LEVELS.(U)
FEB 79 H J NEWTON
TR-N-8

F/G 6/1

UNCLASSIFIED

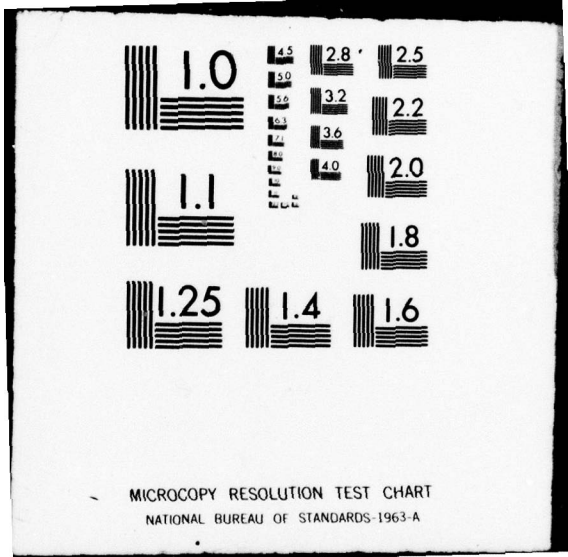
N00014-78-C-0599
NL

| OF |

AD
A067 771



END
DATE
FILMED
6 -79
DDC



LEVEL

TEXAS A&M UNIVERSITY
COLLEGE STATION, TEXAS 77843

INSTITUTE OF STATISTICS
Form 713-343-3141

ADA067271

TIME SERIES ANALYSIS OF LH HORMONE LEVELS

By H. Joseph Newton
Institute of Statistics, Texas A&M University

Technical Report No. H-5
February 1979

Texas A & M Research Foundation
Project No. 3838

"Multiple Time Series Modeling and Time
Series Theoretic Statistical Methods"

Sponsored by the Office of Naval Research

Professor Emanuel Parzen, Principal Investigator

Approved for public release; distribution unlimited.

DDC
APR 20 1979
RECEIVED
C

Unclassified SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)		REPORT DOCUMENTATION PAGE	
1. REPORT NUMBER Technical Report No. H-5	2. GOVT ACCESSION NO.	3. REPORT NUMBER COVERED Technical Report	4. REPORT NUMBER COVERED Technical Report
5. TITLE (and Subtitle) Time Series Analysis of LH Hormone Levels,		6. PERFORMING ORGANIZATION NAME(S) AND ADDRESS H. Joseph Newton Texas A&M University Institute of Statistics College Station, TX 77843	
7. AUTHOR(S)		8. CONTRACT OR GRANT NUMBER(S) N00014-78-C-0599	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Texas A&M University Institute of Statistics College Station, TX 77843		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Code 436 Arlington, VA 22217		12. MONITORING AGENCY NAME(S) AND ADDRESS(ES) (If different from Controlling Office)	
13. MONITORING AGENCY REPORT NUMBER 1233P		14. SECURITY CLASS. (of this report) Unclassified	
15. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.			
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)			
18. SECURITY CLASS. (of this abstract) Unclassified			
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Empirical time series analysis			
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) This report illustrates the methods of time series analysis as applied to data from the Physiology of Animal Reproduction. The presentation is intended to be nontechnical in nature so that researchers in a wide variety of fields can infer the applicability of the methods. The output of a computer program written by the author is described.			

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 68 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

9 04 17 136

347380

46

DDC FILE COPY

TIME SERIES ANALYSIS OF LH HORMONE LEVELS

H. Joseph Newton
Institute of Statistics, Texas A&M University

Abstract

This report illustrates the methods of time series analysis as applied to data from the Physiology of Animal Reproduction. The presentation is intended to be nontechnical in nature so that researchers in a wide variety of fields can infer the applicability of the methods. The output of a computer program written by the author is described.

1. Introduction

This report describes a set of data consisting of levels of the hormone LH measured at ten minute intervals for 24 hours in four cows each on three different days. Thus there are twelve time series each having 144 observations. Included in the appendix are plots of the data sets. Plots are included on an overall scale (to compare level of the series) and on an individual scale (to compare types of variability in the series).

Section 2 describes some questions of interest concerning the time series and outlines the statistical methodology used. In Section 3 initial results are presented.

2. Aims of the Analysis

These time series present many interesting questions. In this report the following particular aspects are investigated:

- i) What is the periodic behavior of the various series; i.e., can the variability in a series be explained by sinusoidal behavior of various periods?
- ii) Are mean levels of LH significantly different for the different series?
- iii) Can one of the standard time series models be used to fit the data and predict future values of the series?

To answer these questions a brief description of the basic methodology of time series analysis is necessary. What distinguishes time series data $X(1), \dots, X(T)$ from the type of data traditionally considered in statistics is that individual observations are not independent of the others. We

ACCESSION for	<input checked="" type="checkbox"/> White Section	<input type="checkbox"/>
NTIS	<input type="checkbox"/> Buff Section	<input type="checkbox"/>
BDC		
UNANNOUNCED		
JUSTIFICATION		
DISTRIBUTION AND AVAILABILITY CODES		
SPECIAL		

A

assume that observations v time units apart all have covariance $R(v)$ and that as v increases, $R(v)$ approaches zero, i.e., observations far apart in time are almost independent. Thus one description of a time series is the correlogram $\rho(v)$ vs. v where $\rho(v) = R(v)/R(0)$ measures the correlation of data points v apart.

Another useful tool is to consider the discrete Fourier transform of the sequence of $R(v)$. This is the function $f(w)$, $w \in [0, 1]$ which is called the spectral density function of the time series. This function is used roughly in the same way an electrical engineer uses harmonic analysis. Thus the plot of a time series is similar to a waveform and the spectral density function assesses the amount of variability in the series due to cosines and sines in the various frequency ranges.

Time series methods based on the correlogram are called time domain methods while methods based on the spectral density are called frequency domain methods. Many analysts use only time domain methods while we prefer to use the insights gained by using both types.

The computer program used in analyzing a time series essentially consists of estimating the correlogram by its sample analog and finding three different estimators of the spectral density: (i) periodogram $f_T(w)$, (ii) smoothed periodogram $f_{T,M}(w)$, and (iii) autoregressive estimator $f_p(w)$. The autoregressive estimator also gives a linear statistical model for describing the data. An observation $X(t)$ is assumed to be a linear combination of the previous p observations in time plus a random error. The program estimates a \hat{p} which is best in some sense and the weights $\hat{\alpha}(1), \dots, \hat{\alpha}(\hat{p})$ of the linear combinations.

3. Results

Computer Output

For each data set there are twelve plots. The series is identified by the caption above the plot while the plot is identified (except for the last two) by the label below the horizontal axis. The plots are described as follows:

Identifier	Plot
DATA	Data set
FT	Periodogram estimator
CORR	Correlogram estimator
FH	Smoothed estimator $f_{T,M}$ using the Parzen window for 3 orders M of smoothing
PART	Partial autocorrelations
BRV	Biased Residual Variance Estimator
URV	Unbiased Residual Variance Estimator
CAT	Parzen's CAT criterion
ARSB	Autoregressive Estimator, Best Order
AMS2	Autoregressive Estimator, 2nd Best Order

The last two plots show the data (solid line) and values fitted by the autoregressive model (x 's). The plots PART, BRV, URV, and CAT help identify the autoregressive order \hat{p} .

Periodic Behavior

The CORR plot is used to suggest important periodicities although it is important to realize that such periodicities can be misleading. If

the correlogram is high at a certain value, it could be that a periodicity of that value exists.

Peaks in the FT, FH, ABSB, and ABS2 plots at a frequency suggest a periodicity of period the reciprocal of the frequency. The ABSB and ABS2 plots actually have the periods for the peaks written out.

Thus for series Animal 1, day 1, (identified as NCIA 11) the correlogram is high at lag 7 while the peaks in the spectra occur at frequencies around 1/7.

If one suspects that a periodicity of 1/w exists, this suspicion can be tested by rejecting the hypothesis of no periodicity at the $\alpha = .05$

level if $\frac{f_{1,M}(w)}{f_1(w)} < \frac{1}{3} (1 + \frac{2}{2} \frac{M}{T}) = \frac{2}{12}$. Here M is the order of smoothing and T is the length of the series; we used $M = \frac{1}{6}$. Thus for series (1, 1) at frequencies in the region close to 1/7 the hypothesis of no periodicity is clearly rejected.

The results of the investigation into periodic behavior are summarized:

Periodic Structure of LH Level Series

Series	CORR	SPEC
1, 1	7	Significant peak around 7
1, 2	Decays quickly	High low frequency and possibly significant peak in 20-25 range
1, 3	Decays quickly	High low frequency and possibly significant peak in 20-25 range
2, 1	4	Significant peak around 4
2, 2	Similar to 1, 2	Similar to 1, 2
2, 3	All behavior dominated by low frequency.	Maybe peak around 6.
3, 1	5	Significant peak around 5
3, 2	Similar to 1, 2	Similar to 1, 2

Periodic Structure of LH Level Series (cont.)

Series	CORR	SPEC
3, 3	Same as 2, 3 except maybe peak around 8 or 9.	
4, 1	5	Significant peak around 5. High frequency peak might be significant.
4, 2	Similar to 1, 2	Similar to 1, 2
4, 3	14	Maybe significant peak around 14.

Mean Level of Series

To make inferences about a population mean μ from a sample mean \bar{x} , one needs σ^2 , the variance of \bar{x} . For independent, identically distributed (iid) random variables (i.e., $p = 0$), $\sigma^2_{\bar{x}} = \frac{\sigma^2}{T}$ where σ^2 is the variance of the data. Correlated data is modeled as an autoregression of order p giving approximately $\sigma^2_{\bar{x}} = \frac{\sigma^2 - C}{T - C}$, where

$$C = \frac{1}{p \sum_{j=1}^p \alpha(j)^2}$$

and σ^2 is the variance of the random error in the model. Thus for large samples we conclude that the means of series j and series k are significantly different if

$$|\bar{x}_j - \bar{x}_k| > D_{jk}$$

where $D_{jk} = 2(\sqrt{\hat{\sigma}^2_j + \hat{\sigma}^2_k}) \sqrt{\frac{1}{x_j} + \frac{1}{x_k}}$ where $\hat{\sigma}^2_j$ is the same as $\sigma^2_{\bar{x}_j}$ with estimators $\hat{p}, \hat{\alpha}(1), \dots, \hat{\alpha}(p), \hat{\sigma}^2$ replacing the true unknown $p, \alpha(1), \dots, \alpha(p), \sigma^2$.

It is important to realize that this table is similar to a multiple comparisons procedure in analysis of variance; i.e., the probability of all of these tests leading to correct decisions is very small.

Models

From inspection of the final plot for each series it appears that the autoregressive models fit the data well and provide evidence discriminating between series.

Conclusion

It is obvious that further investigation is necessary to lead to more precise conclusions. However, it is clear from this initial work that these data exhibit some periodic behavior. Further, conclusions can be reached about mean levels of the series and that the autoregressive models can assist in understanding the statistical and physical mechanisms generating the observations.

We summarize these below for the first and second series for each animal. The third series are not included because the approximations to the $\sigma^2_{\bar{x}}$ are not valid.

Mean Levels for LH Level Series

Series	\bar{x}	σ^2	\hat{p}	$\hat{\sigma}^2_{\bar{x}}$
1, 1	17.16	6.30	13	.489
1, 2	15.50	24.75	2	1.189
2, 1	12.72	4.31	18	1.061
2, 2	11.05	18.46	2	1.096
3, 1	15.35	5.65	5	.196
3, 2	12.28	29.74	1	1.695
4, 1	22.67	11.53	7	.126
4, 2	22.20	144.76	2	9.584

Differences and Critical Values for Comparing Means

1, 1	1, 2	2, 1	2, 2	3, 1	3, 2	4, 1
1, 2	-1.66					
	2.59					
2, 1	-4.44	-2.78				
	2.49	3.00				
2, 2	-6.11	-4.45	-1.67			
	2.52	3.02	2.94			
3, 1	-1.81	-0.15	2.63	4.30		
	1.66	2.35	2.24	2.27		
3, 2	-4.88	-3.22	-0.44	1.23	-3.07	
	2.96	3.40	3.32	3.34	2.75	
4, 1	5.51	7.17	9.95	11.62	7.32	10.39
	1.57	2.29	2.18	2.21	1.13	2.70
4, 2	5.04	6.70	9.48	11.15	6.85	9.92
	6.35	6.56	6.52	6.54	6.25	6.72
						-0.47
						6.23

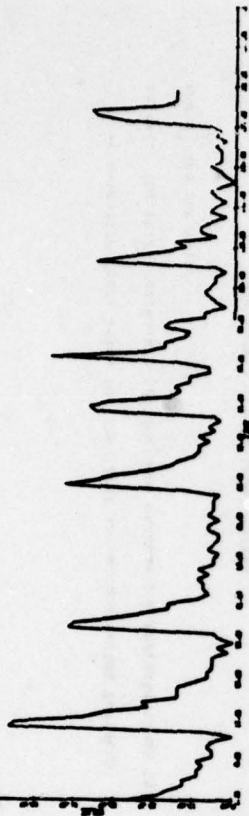
Acknowledgement:

The author wishes to thank Professor P. Harms of the Department of Animal Science at Texas A&M University for providing the data analyzed in this report.

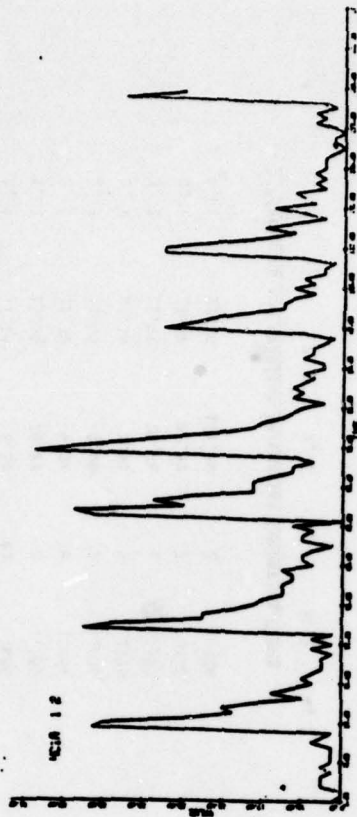
Reference:

For a more technical statistical description of the methods in this report, see "Time Series Modelling, Spectral Analysis, and Forecasting" by E. Parzen, Report M-7, March 1979.

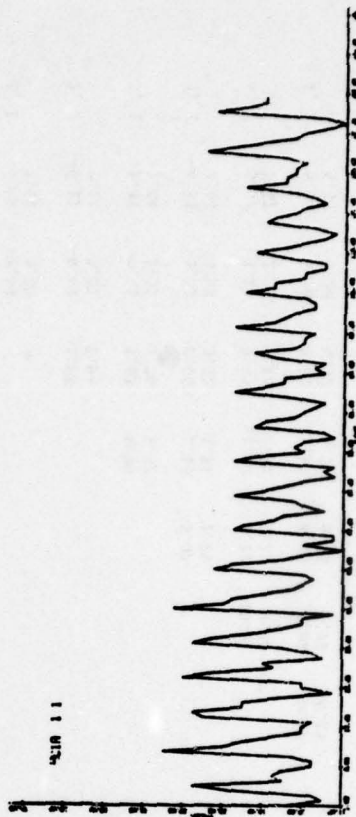
NCIR 13



NCIR 12



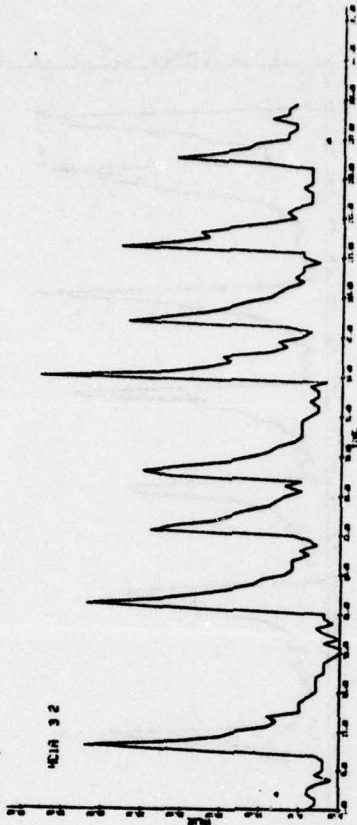
NCIR 11



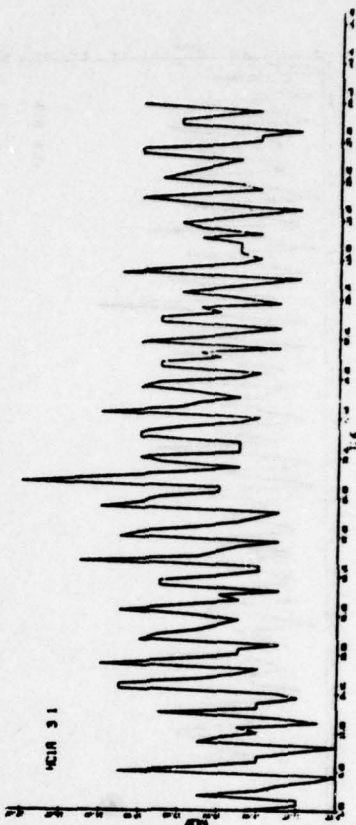
MCIR 33



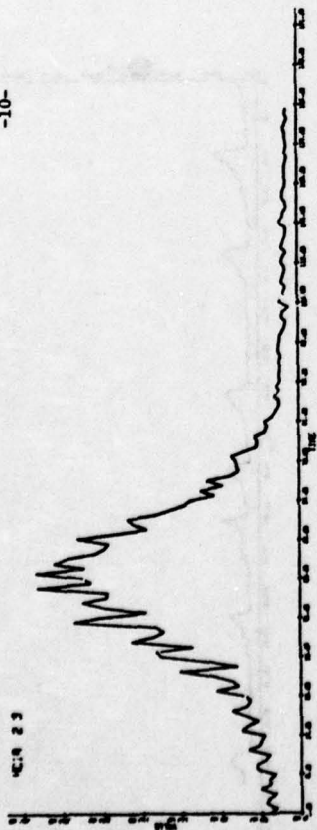
MCIR 32



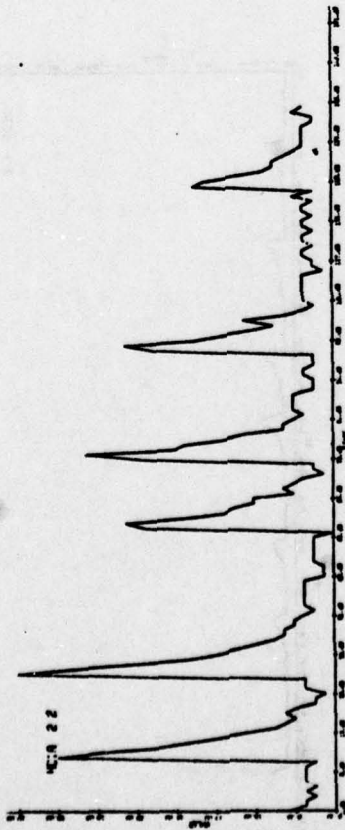
MCIR 31



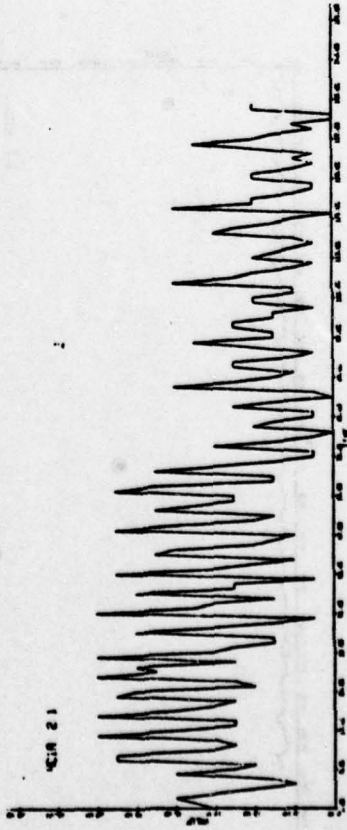
MCIR 23



MCIR 22

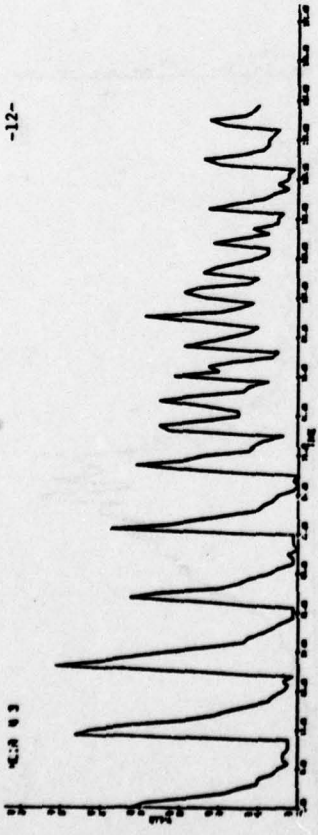


MCIR 21



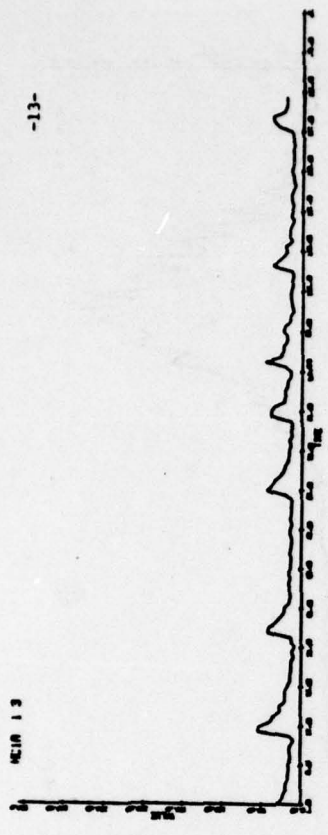
-12-

NCIR 83

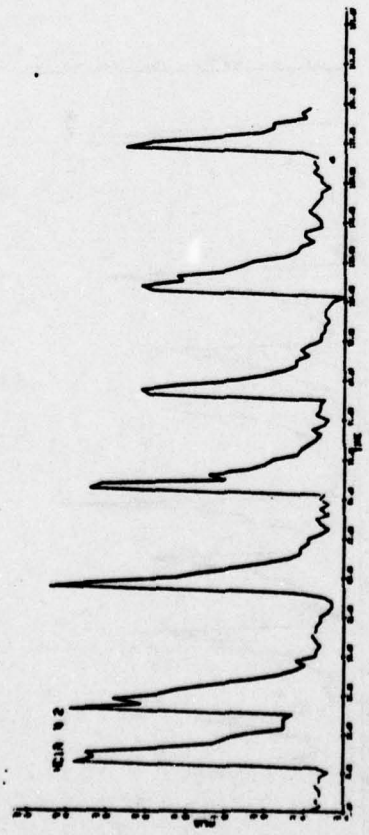


NCIR 13

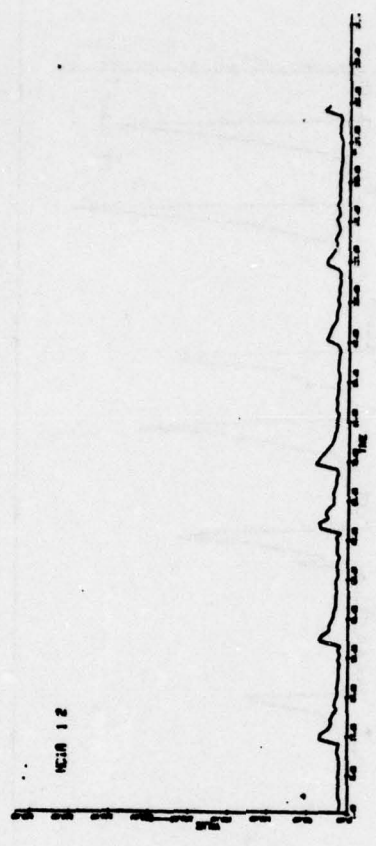
-13-



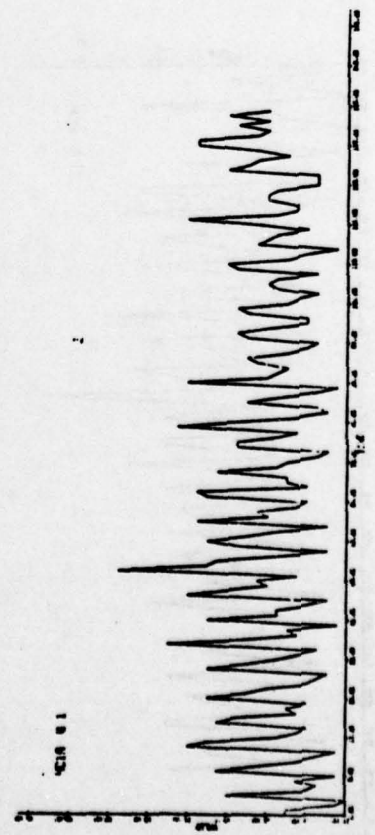
NCIR 82



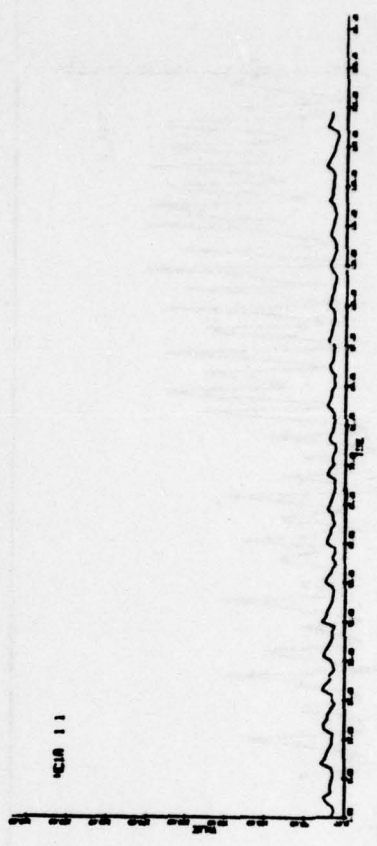
NCIR 12

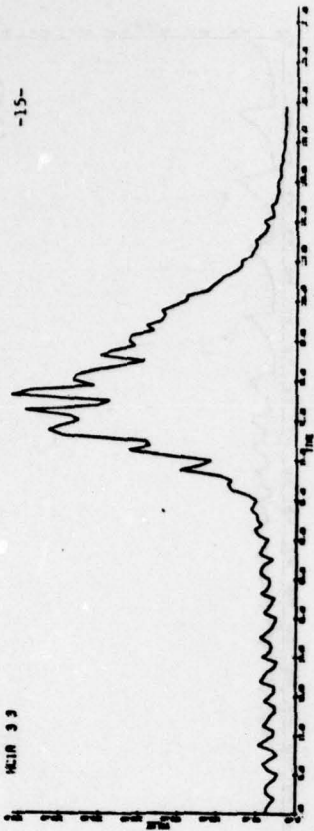


NCIR 81

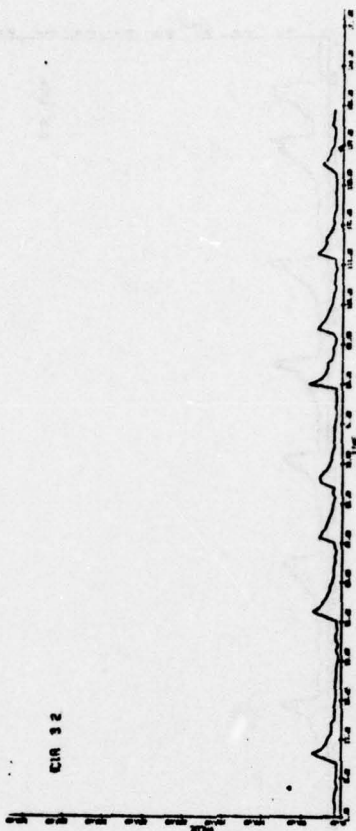


NCIR 11

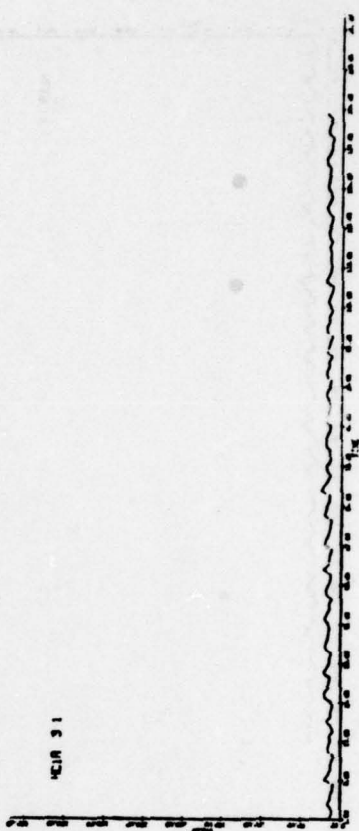




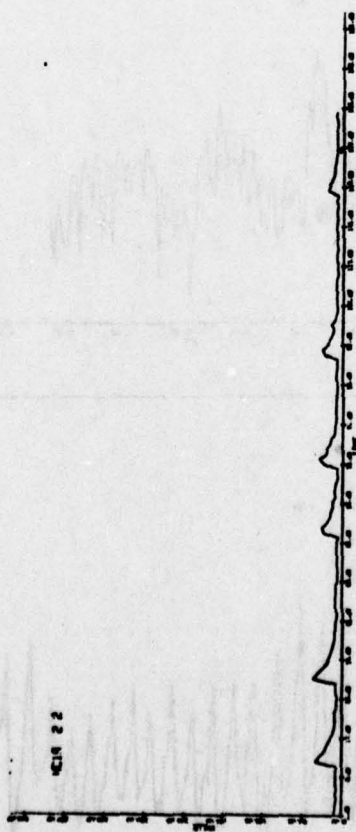
ECIR 32



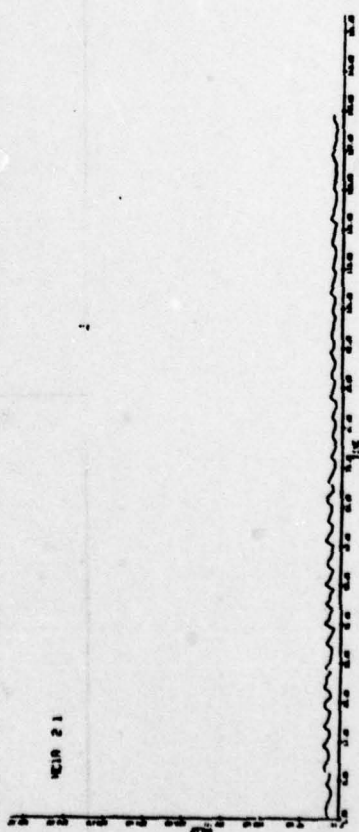
ECIR 31

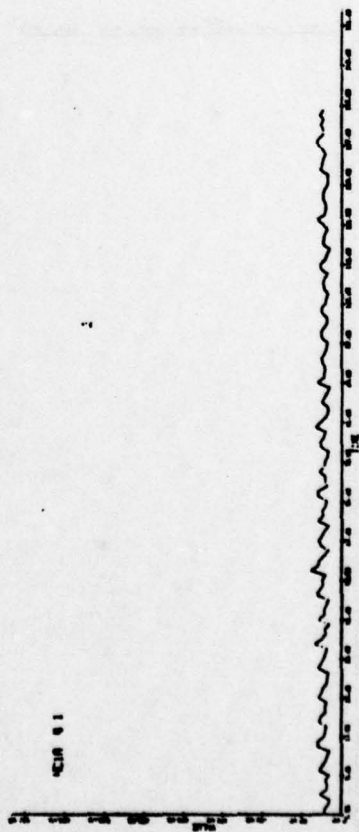
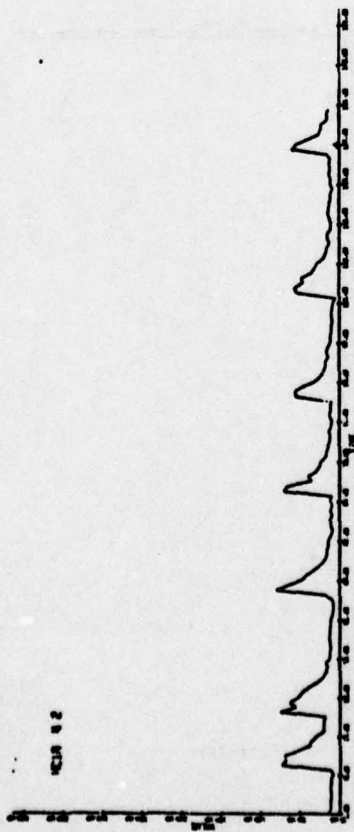
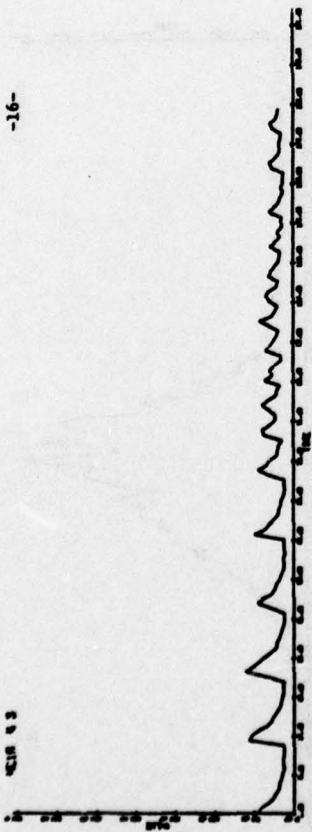
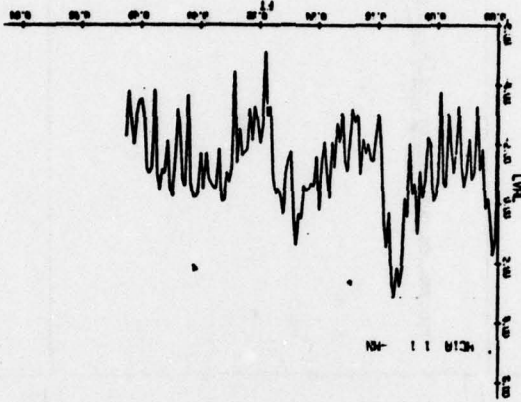
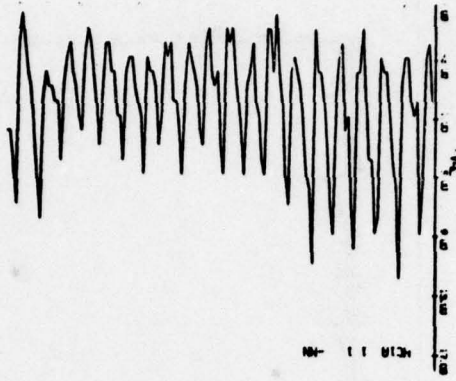


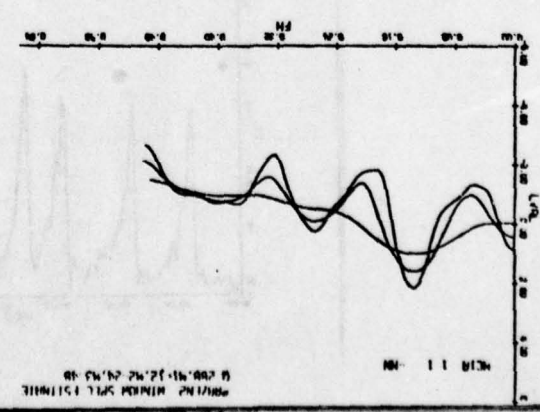
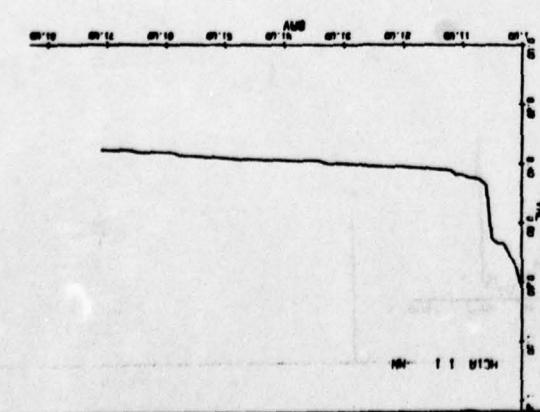
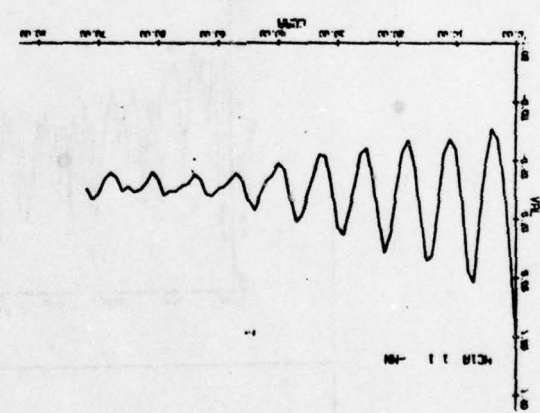
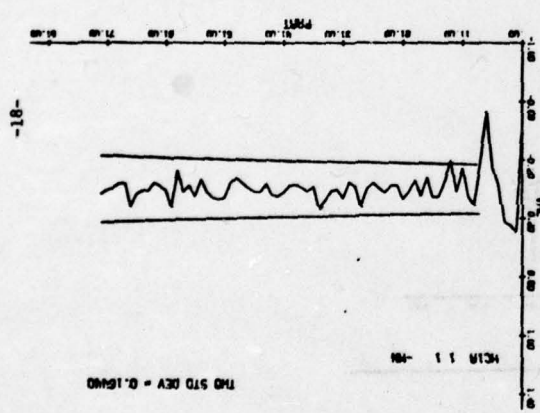
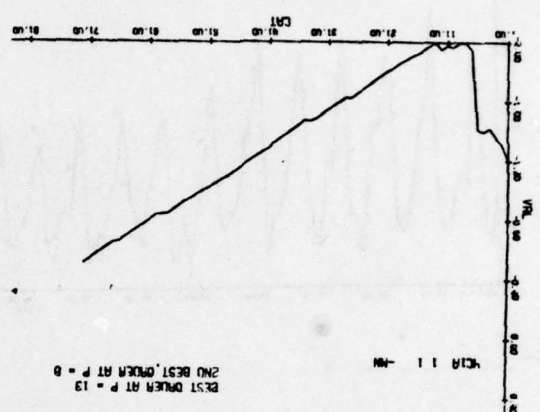
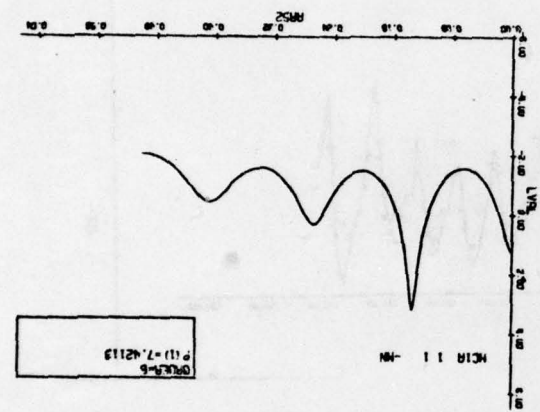
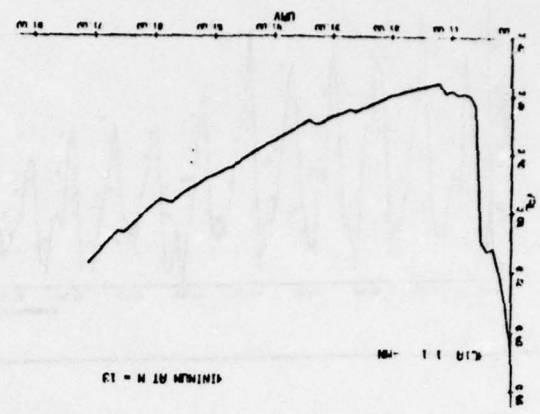
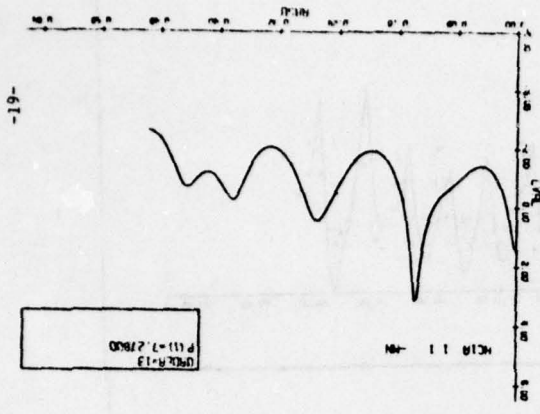
ECIR 22

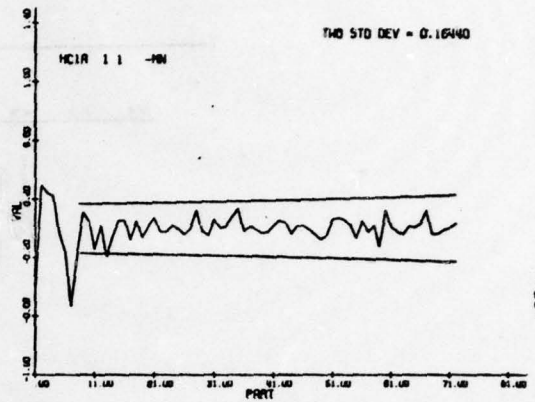
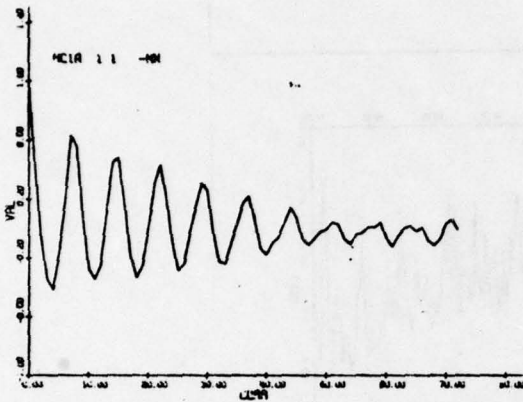
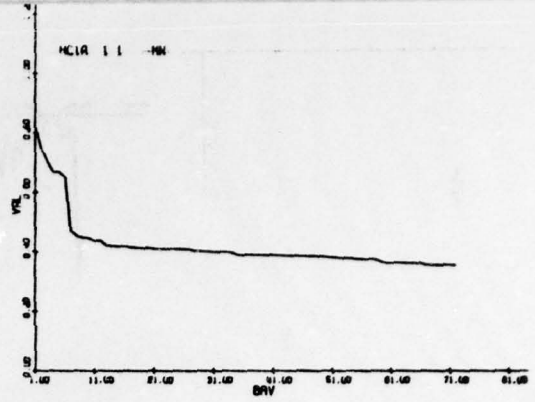
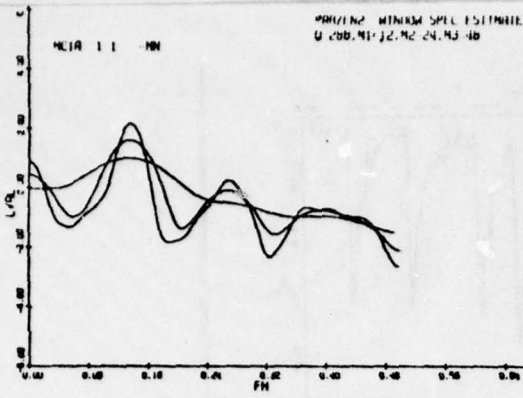


ECIR 21

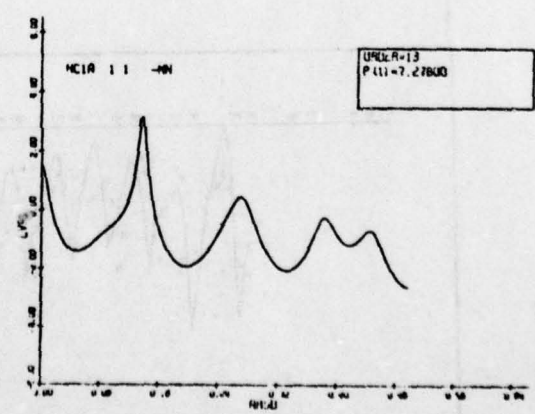
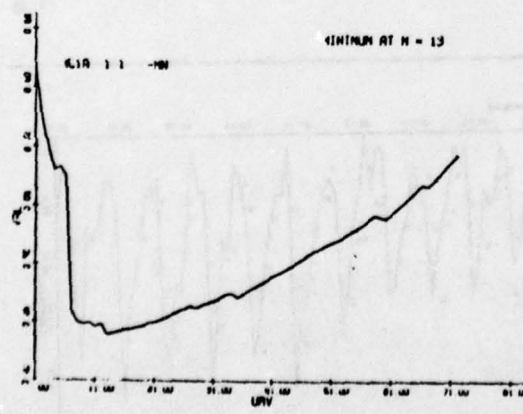
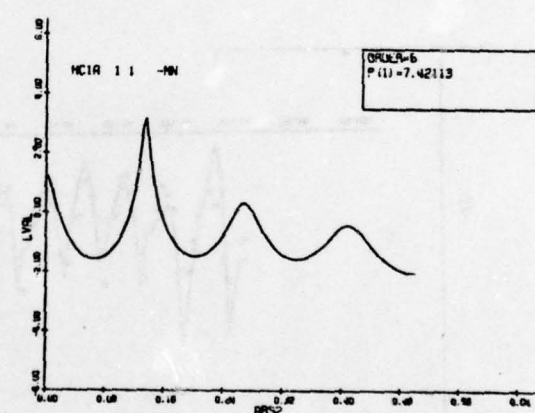
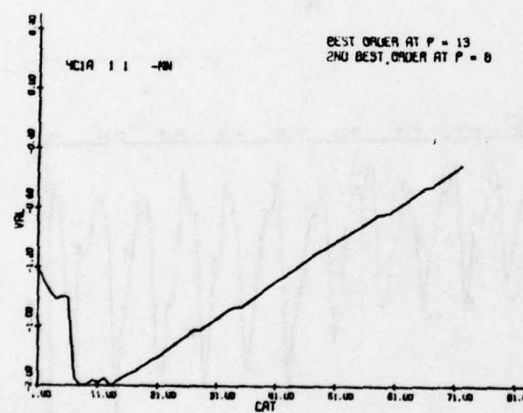




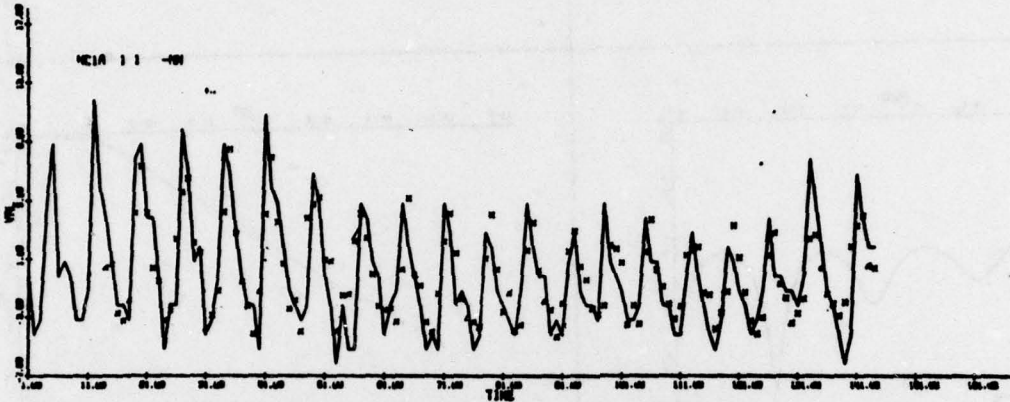
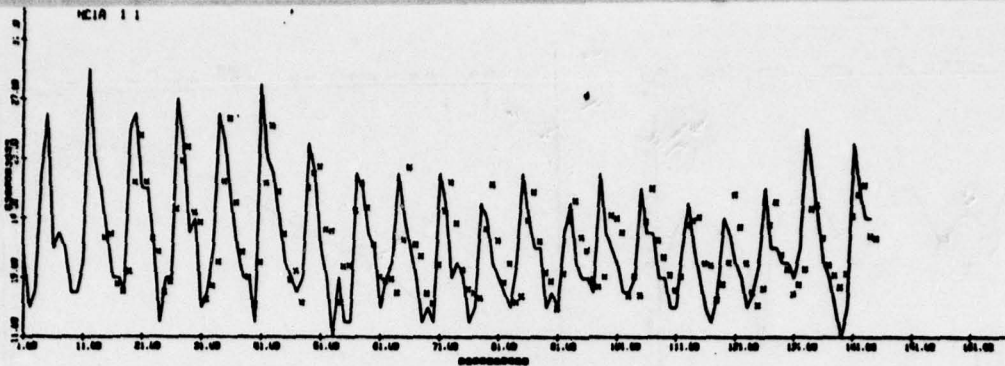




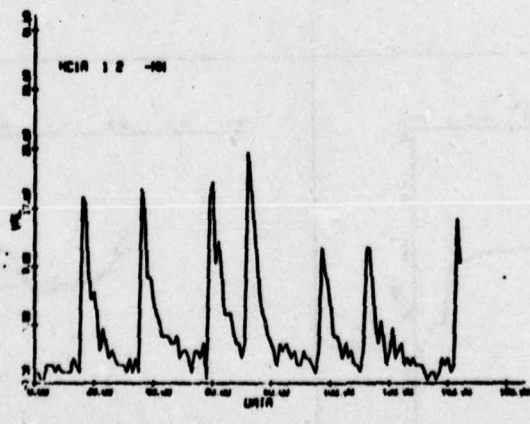
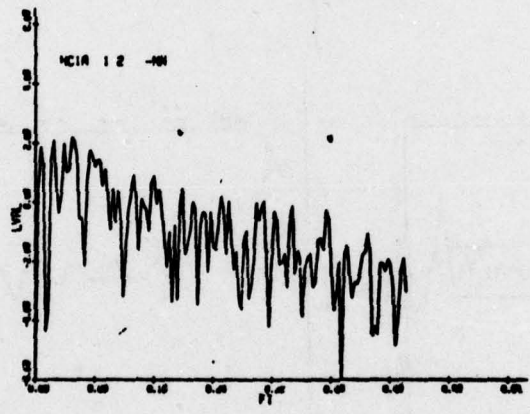
-18-



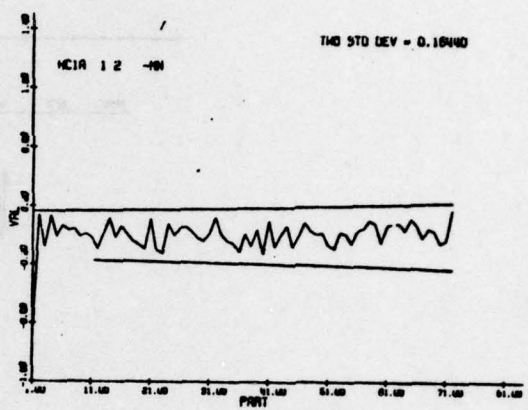
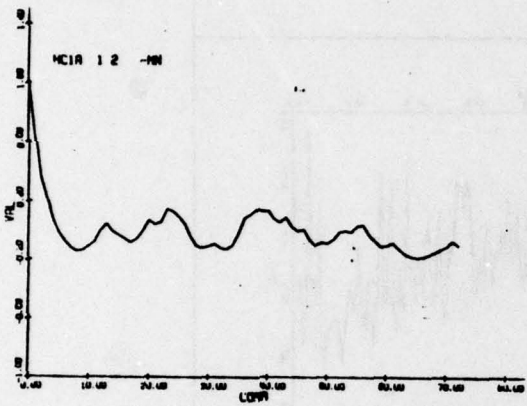
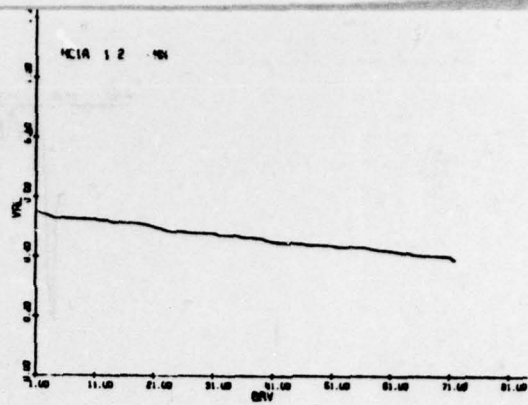
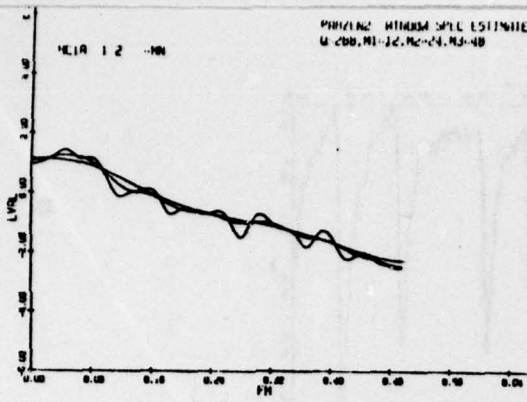
-19-



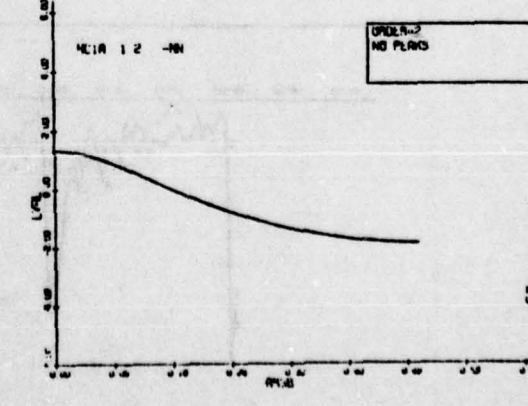
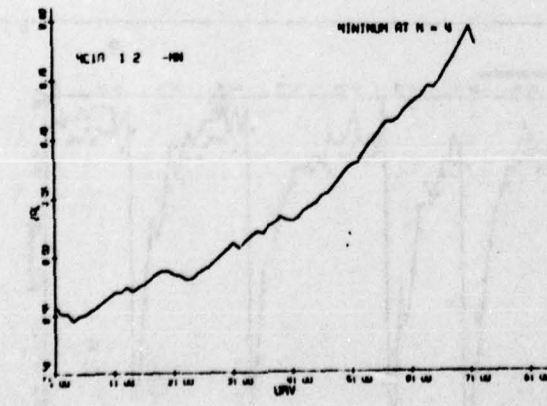
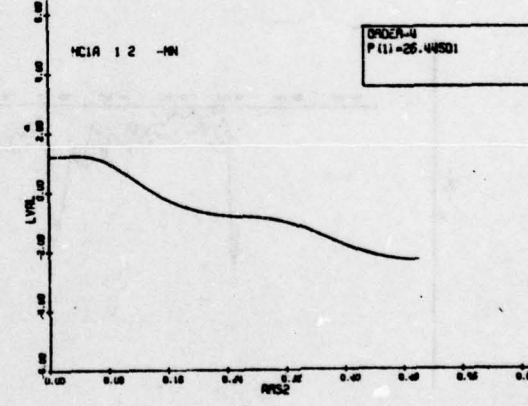
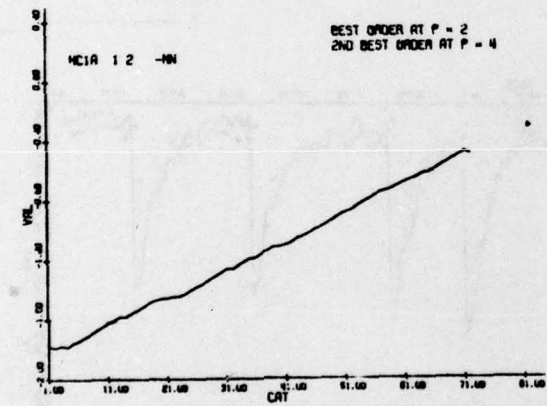
-20-



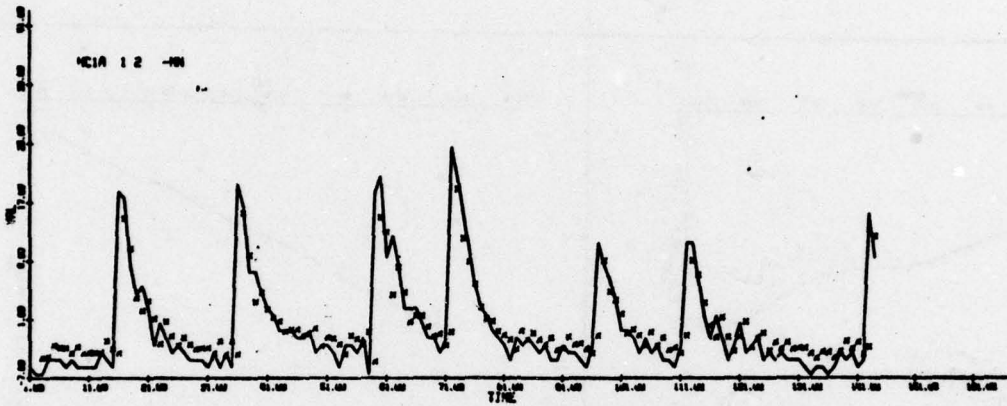
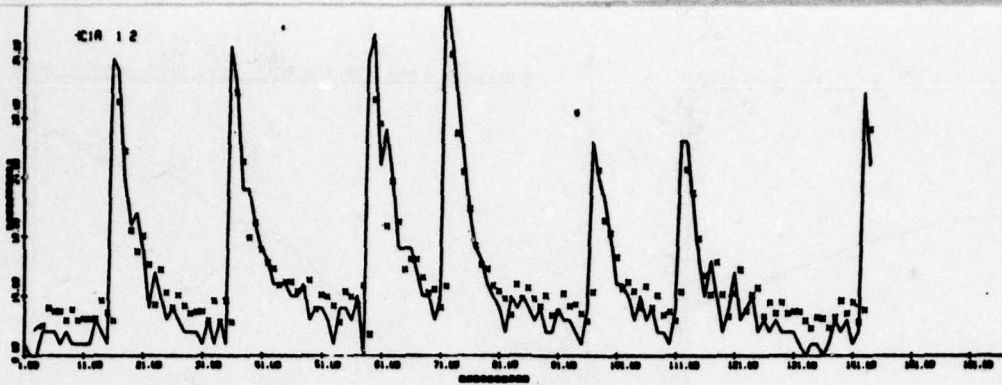
-21-



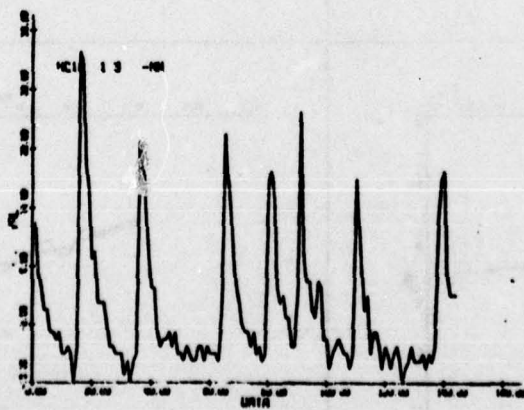
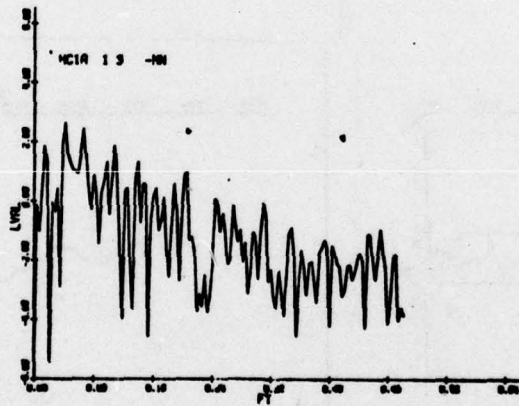
-22-



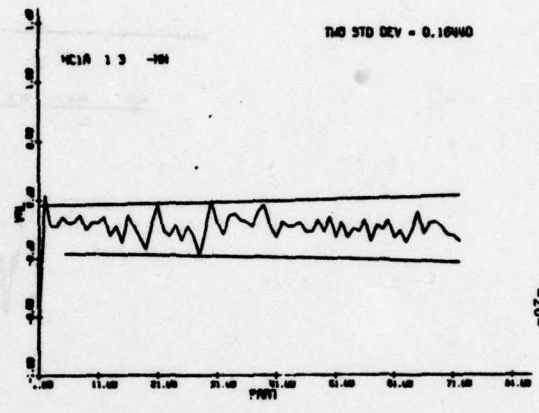
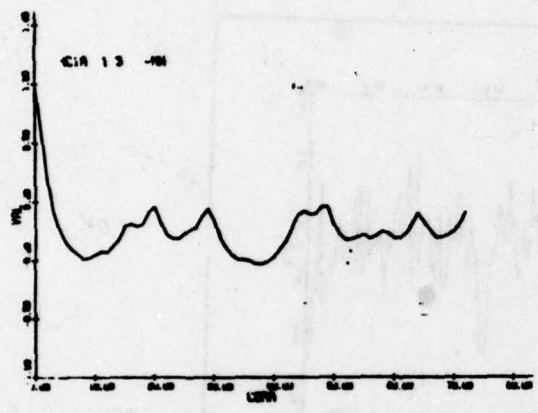
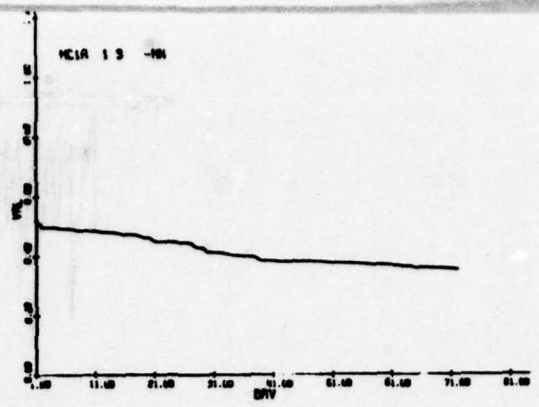
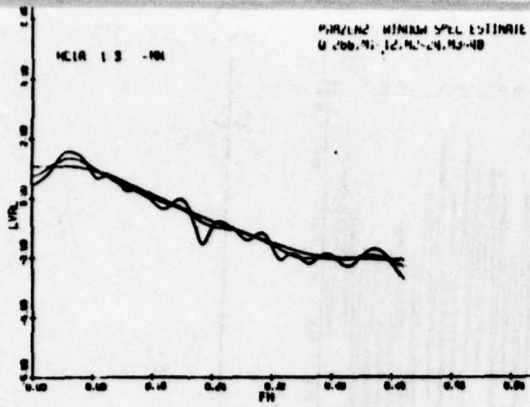
-23-



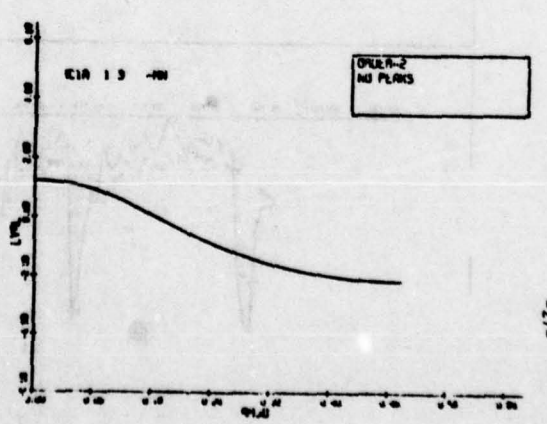
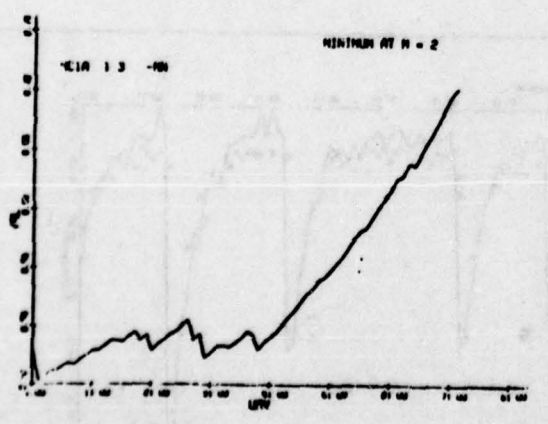
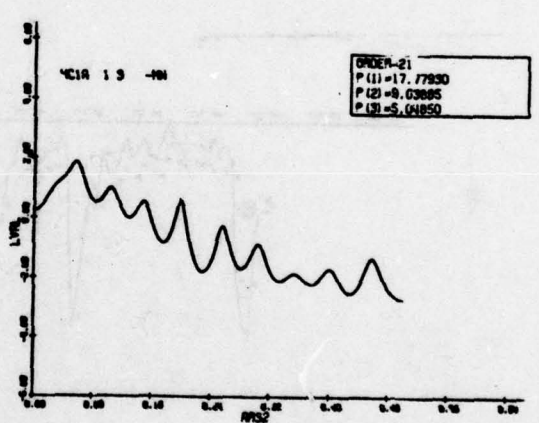
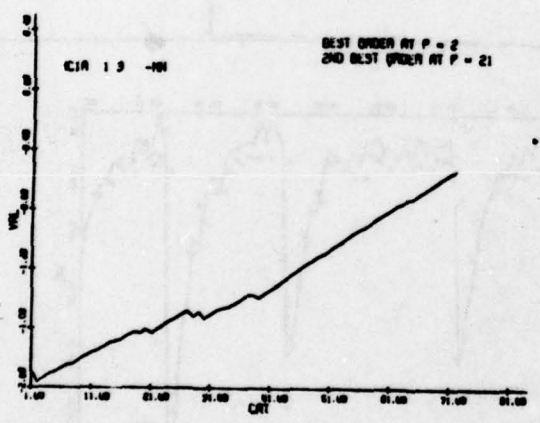
-24-



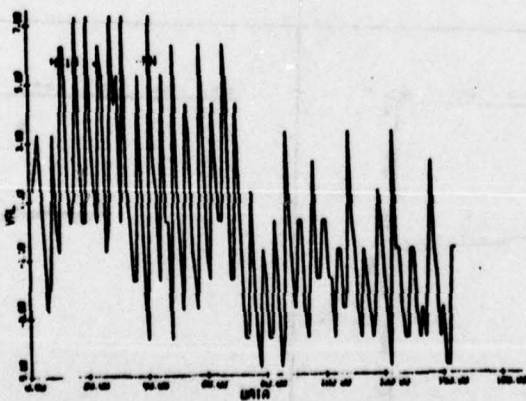
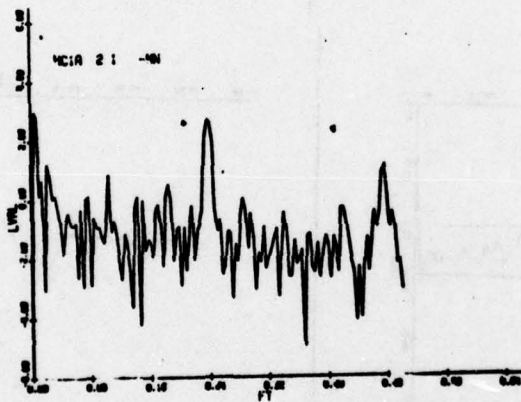
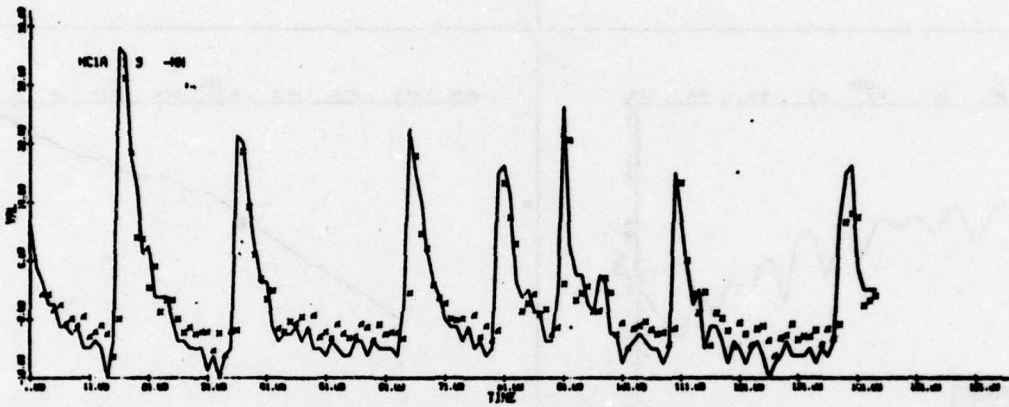
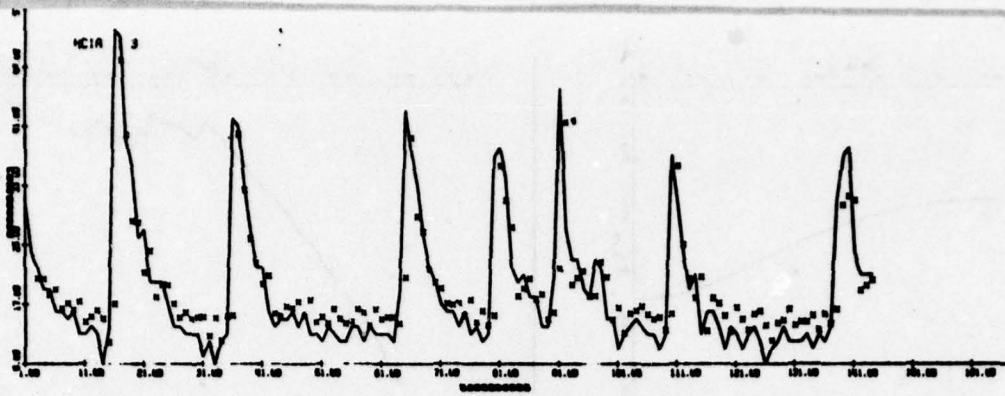
-25-

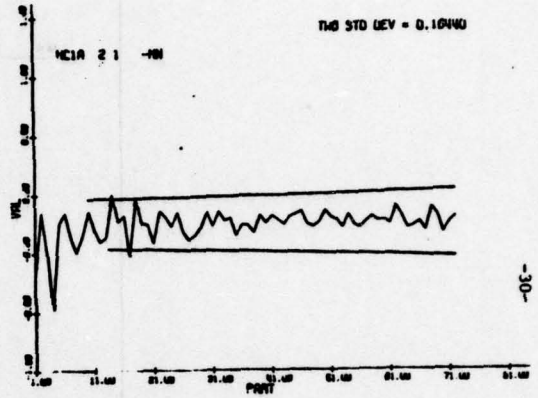
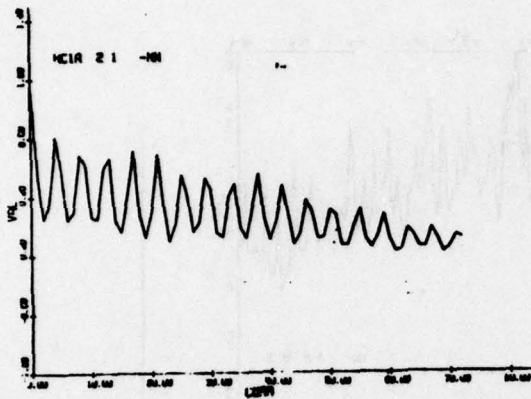
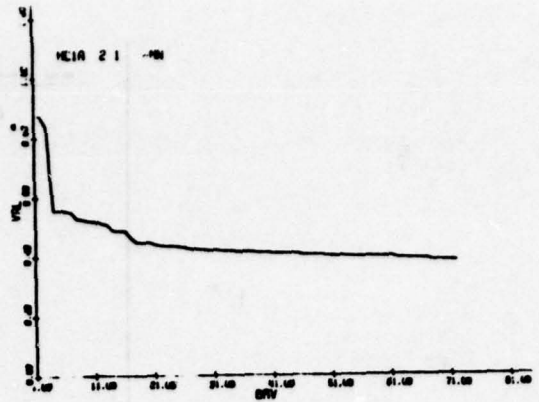
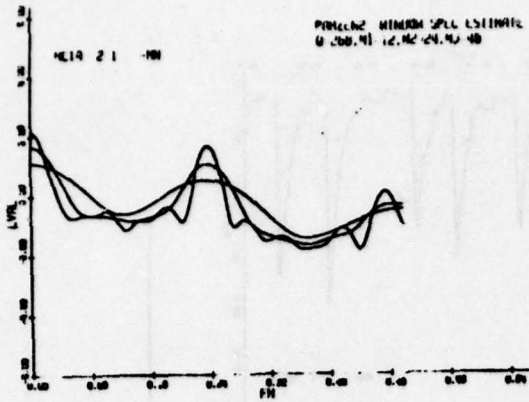


-26-

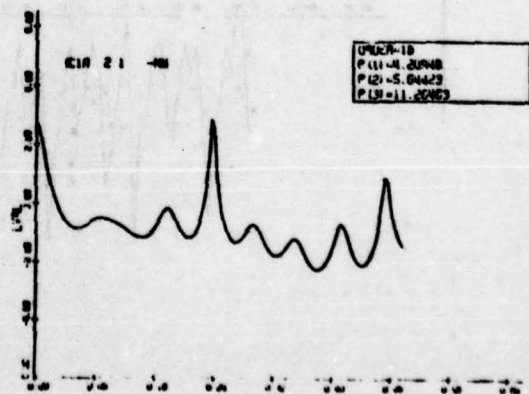
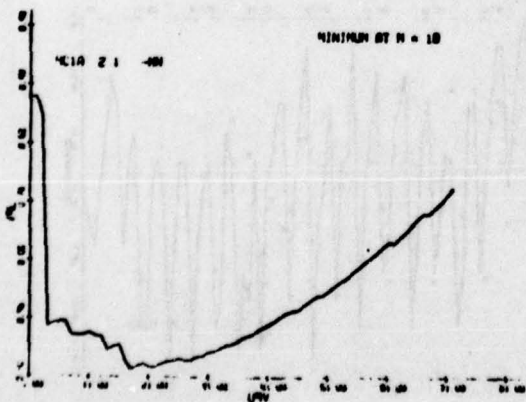
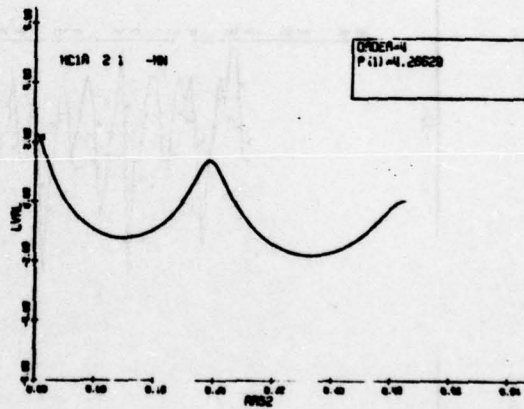
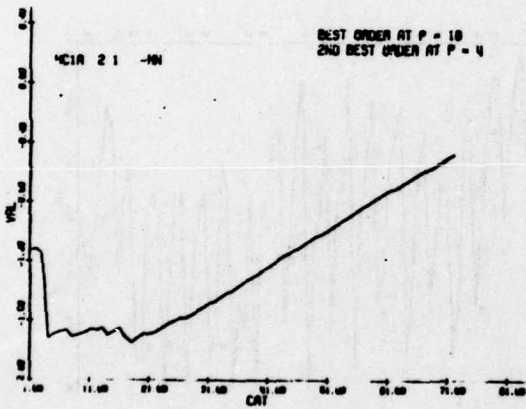


-27-

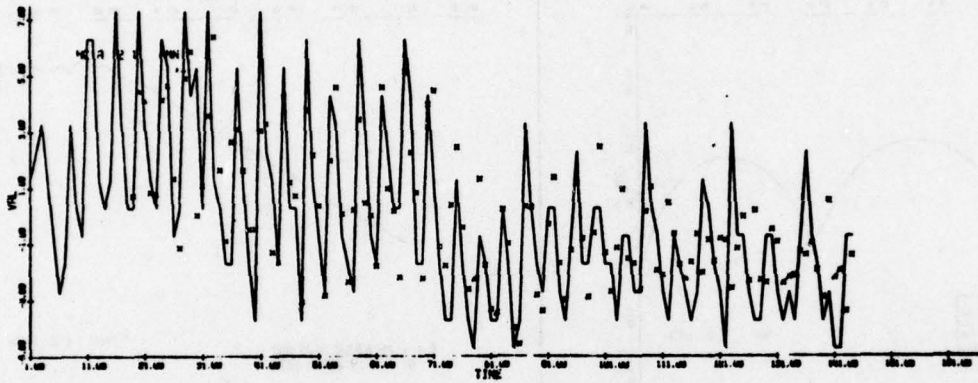
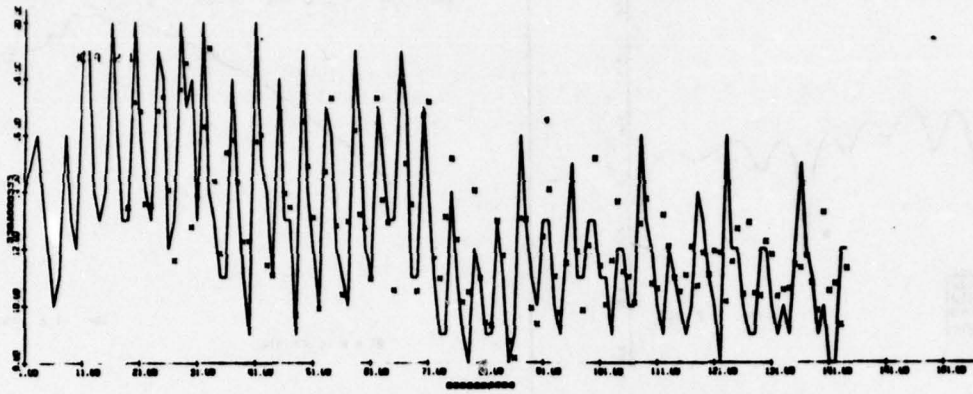




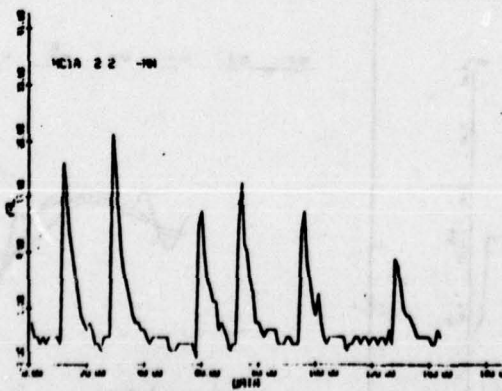
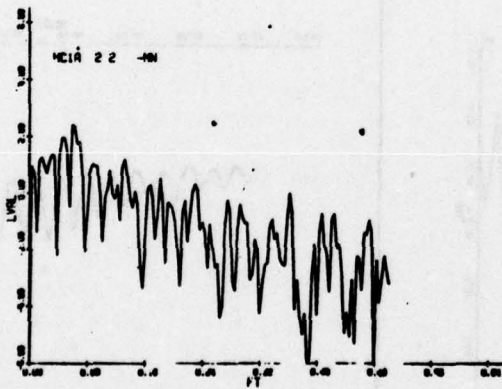
-30-



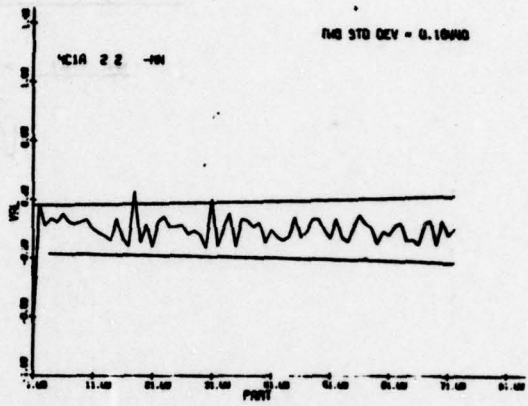
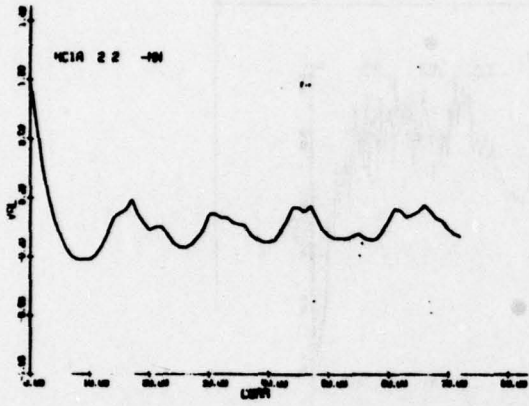
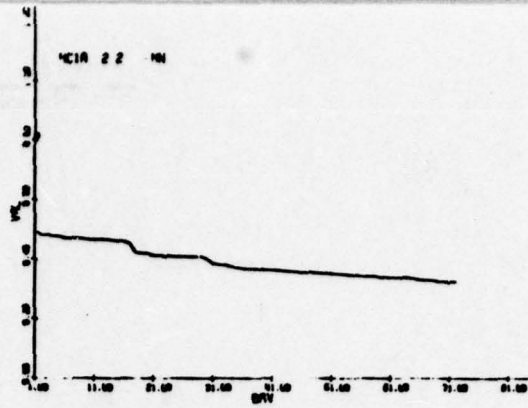
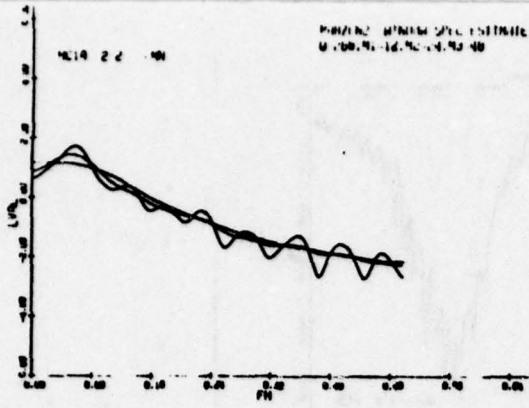
-31-



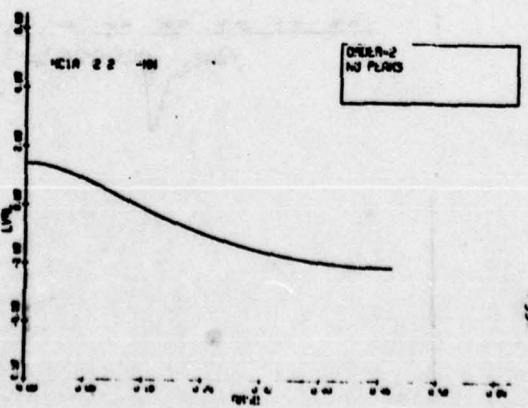
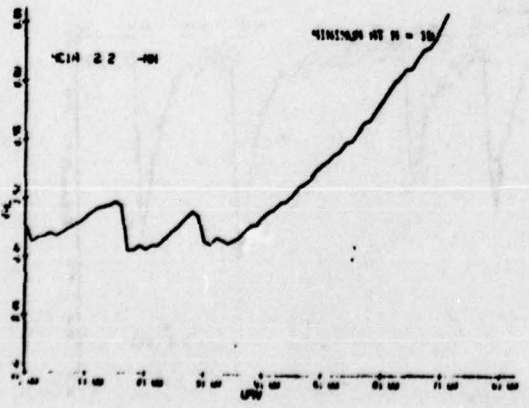
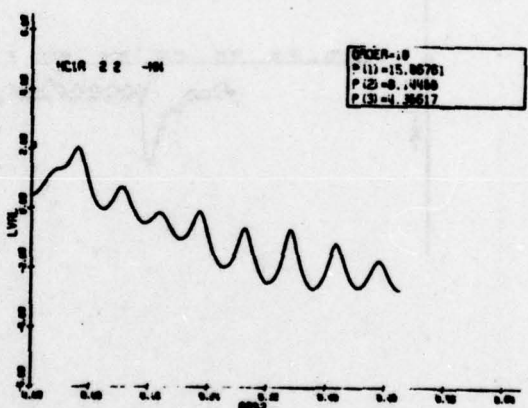
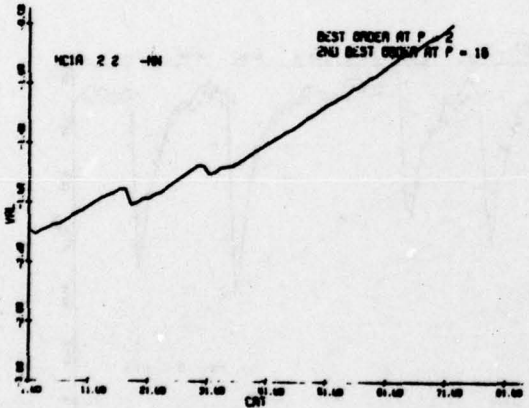
-32-



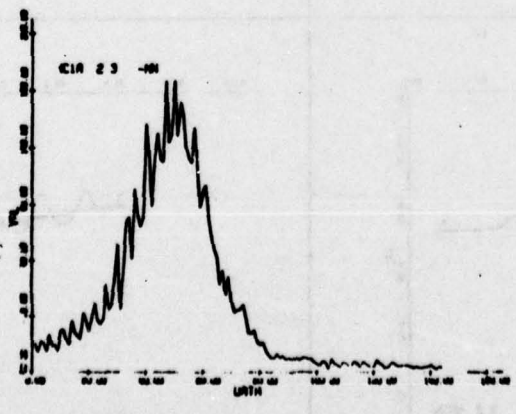
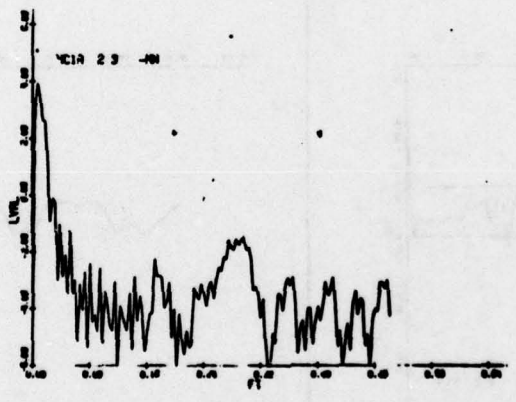
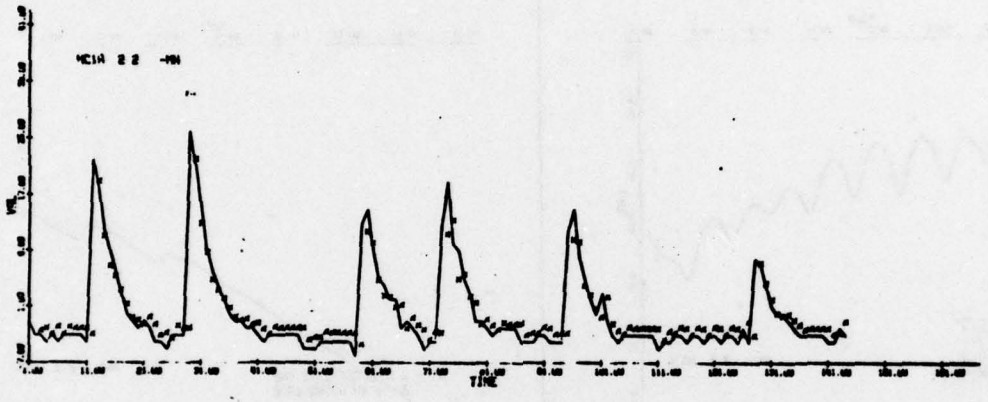
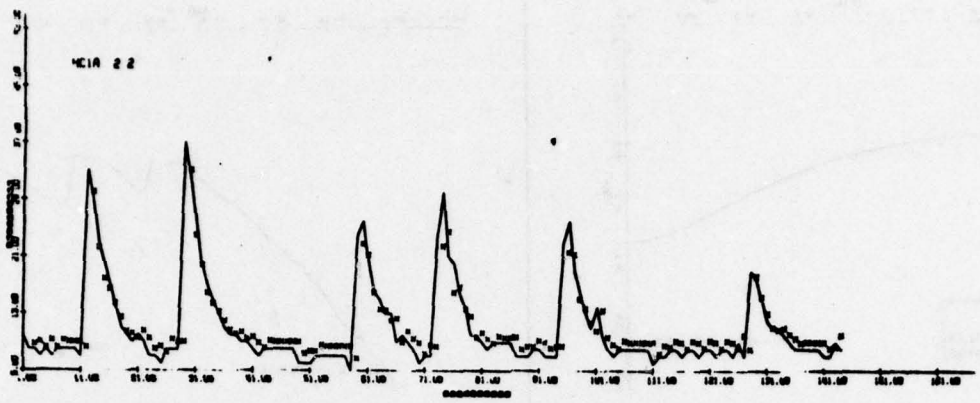
-33-

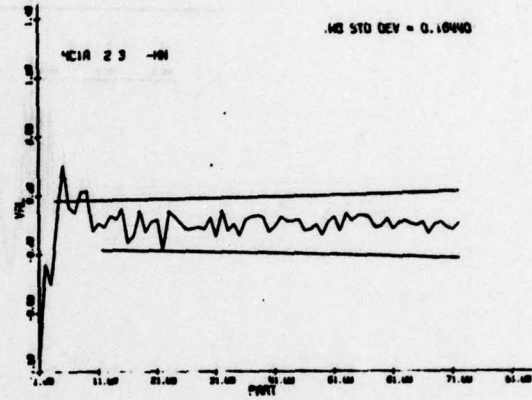
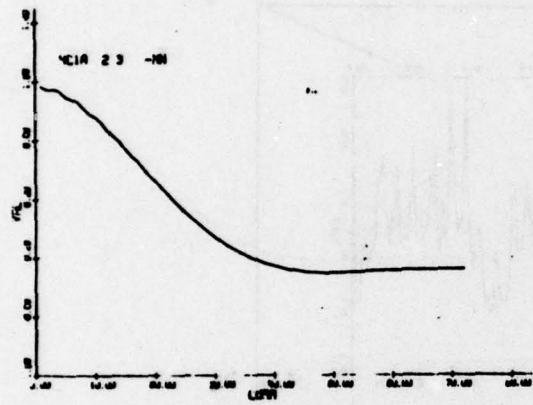
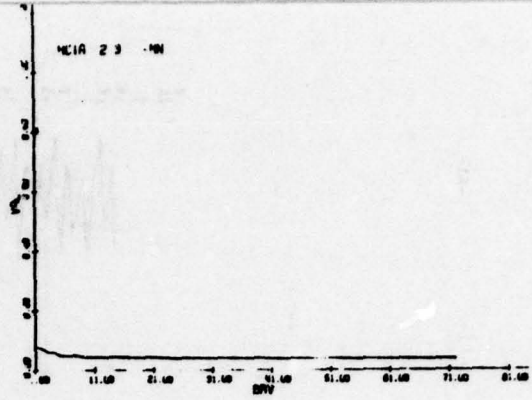
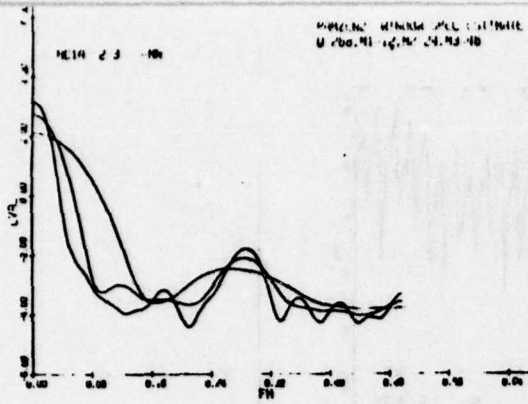


-34-

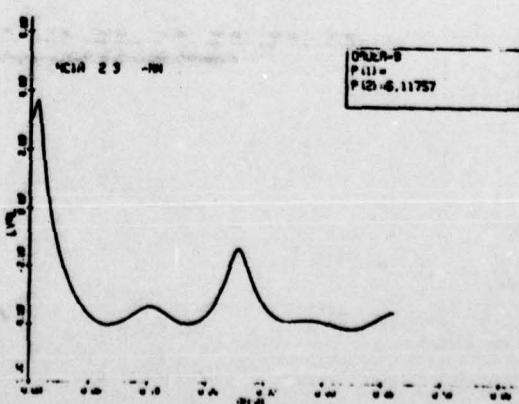
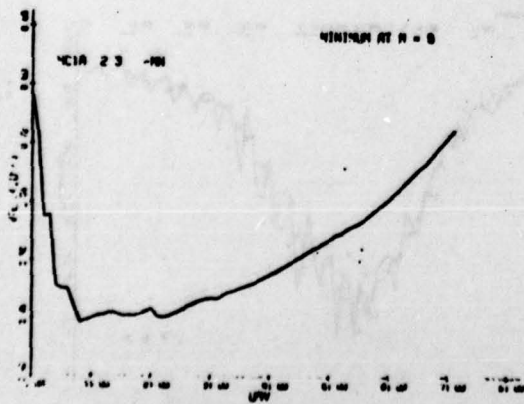
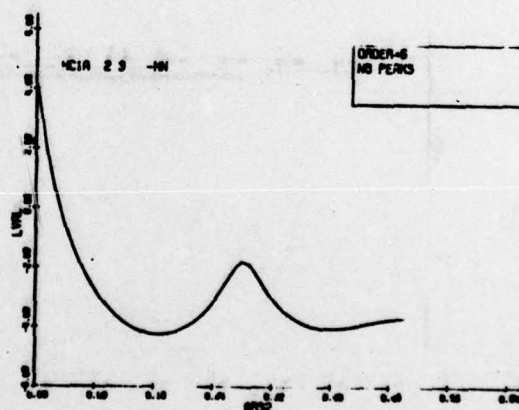
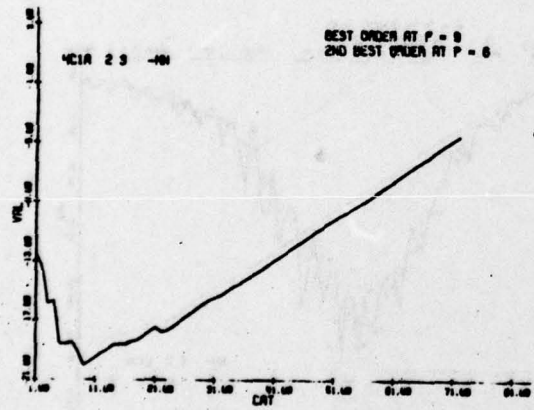


-35-

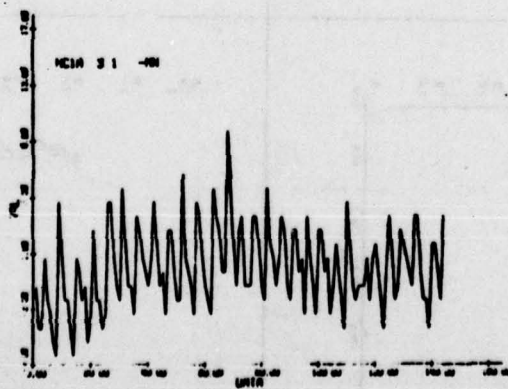
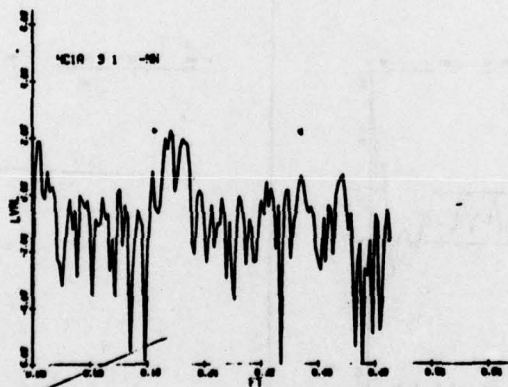
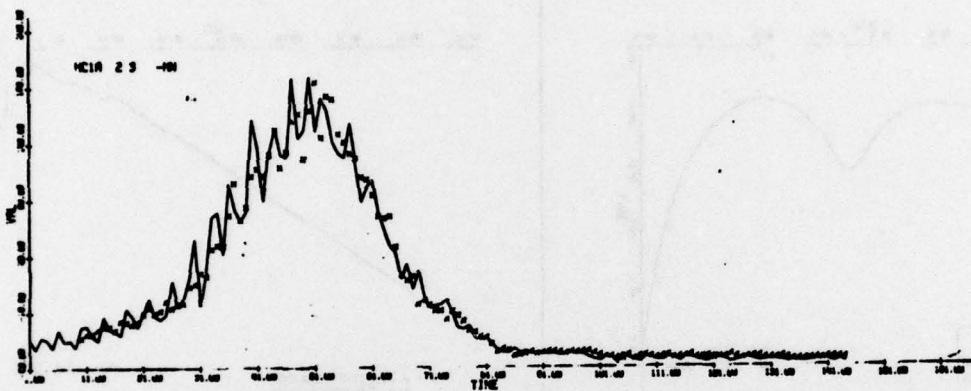
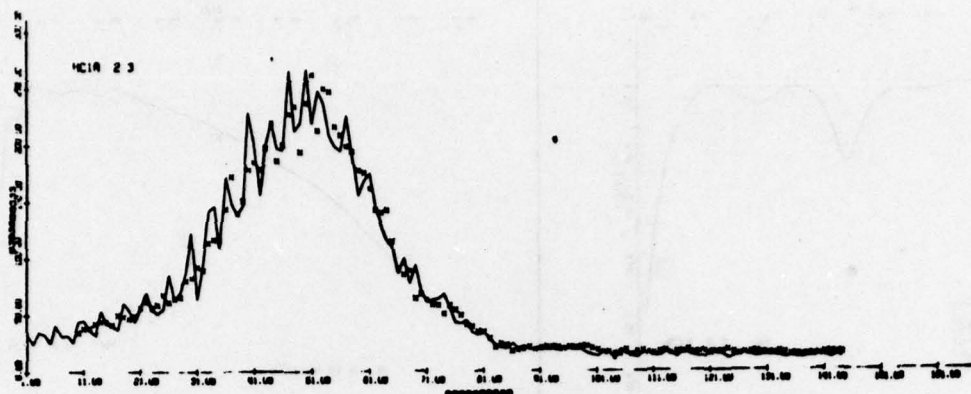


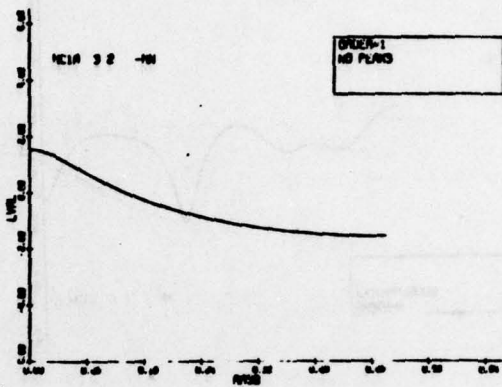
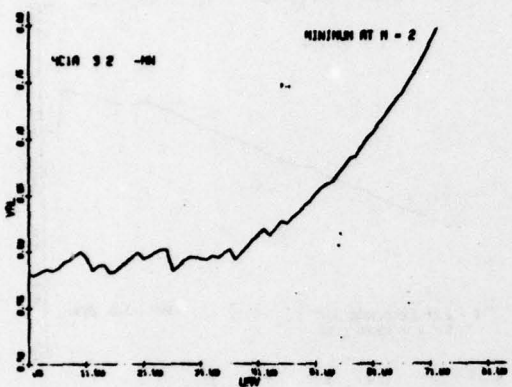
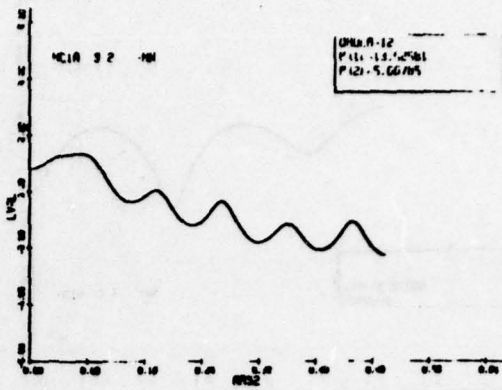
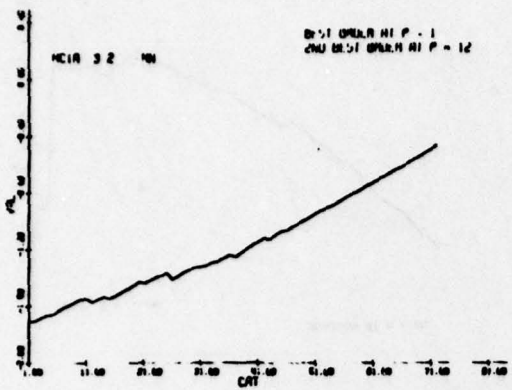


-38-

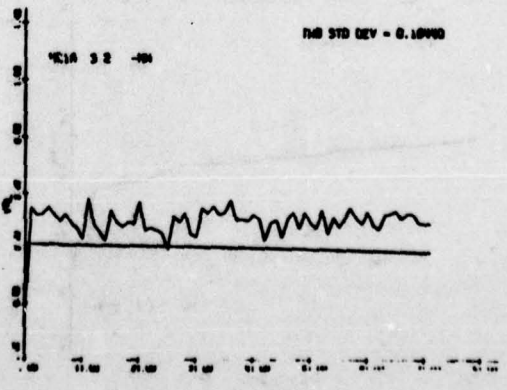
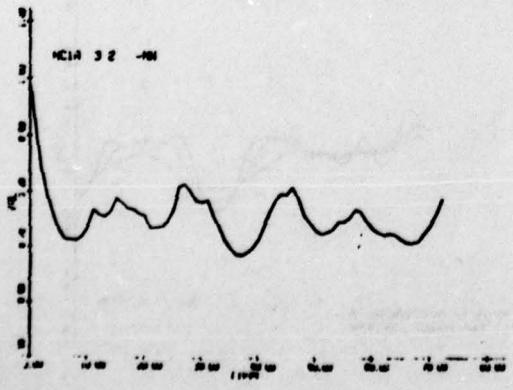
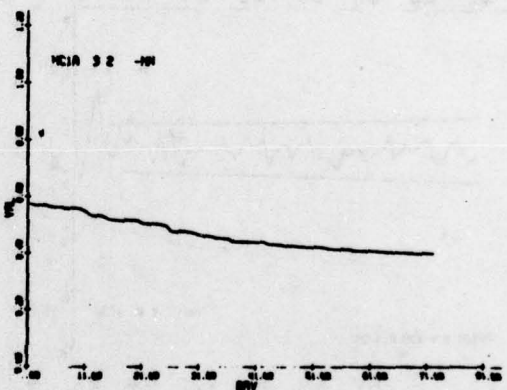
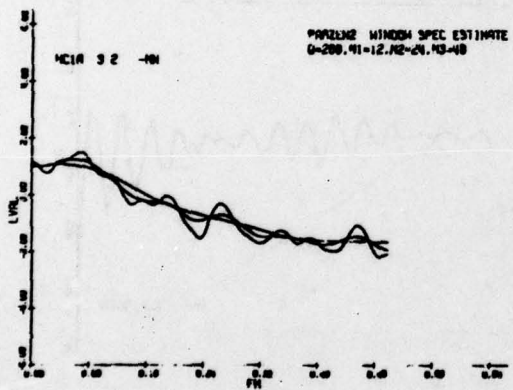


-39-

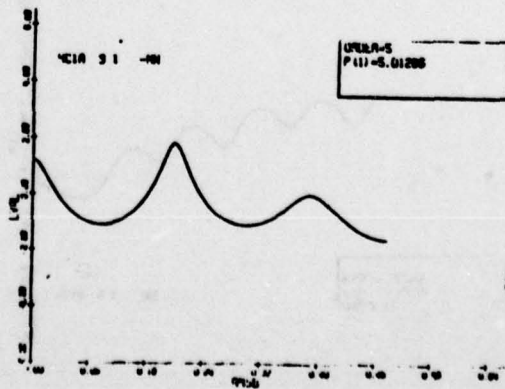
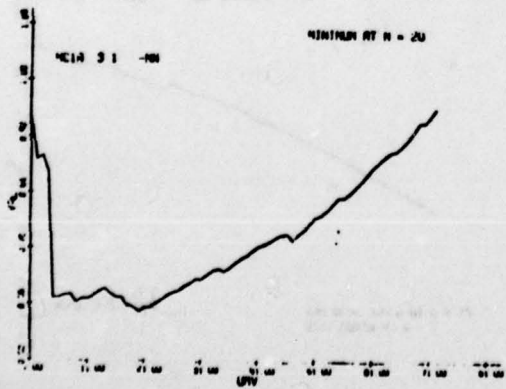
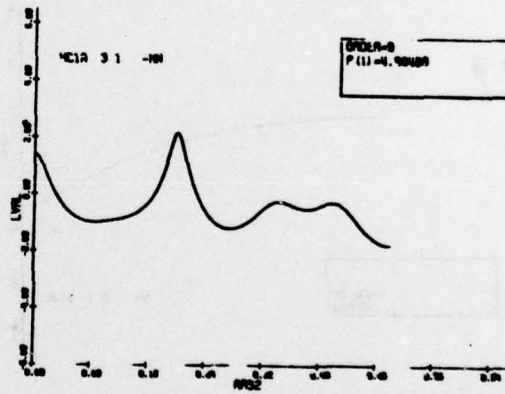
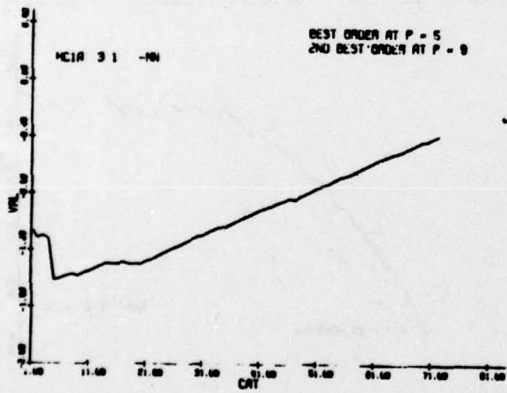
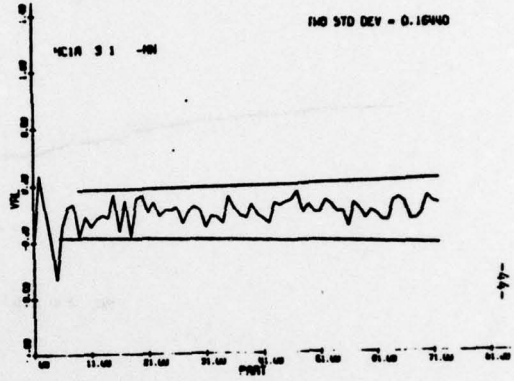
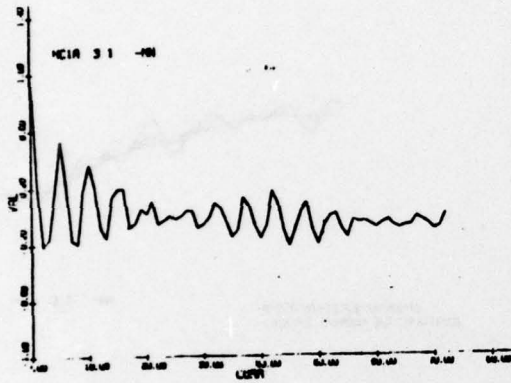
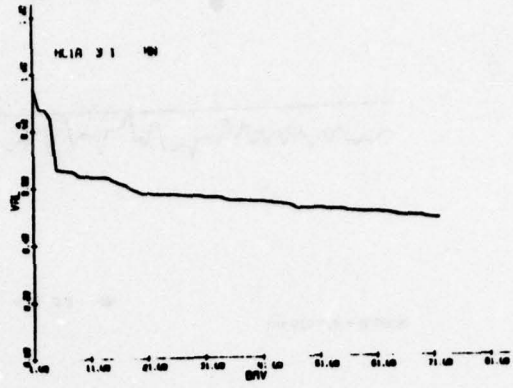
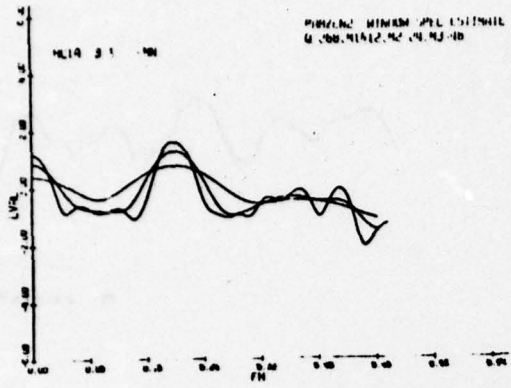


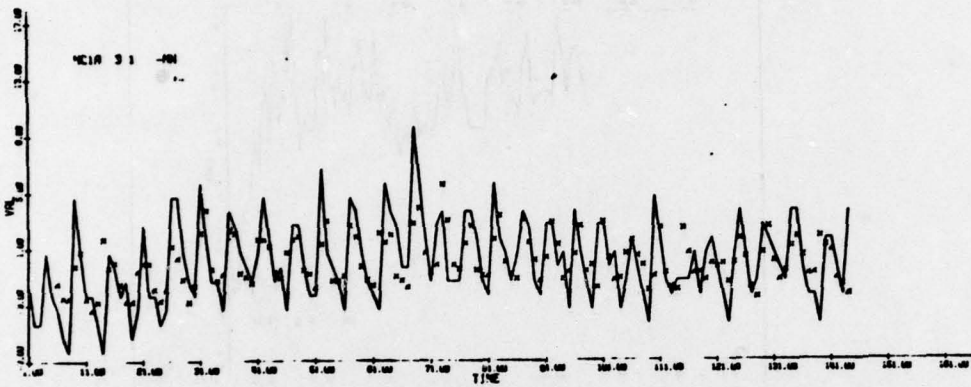
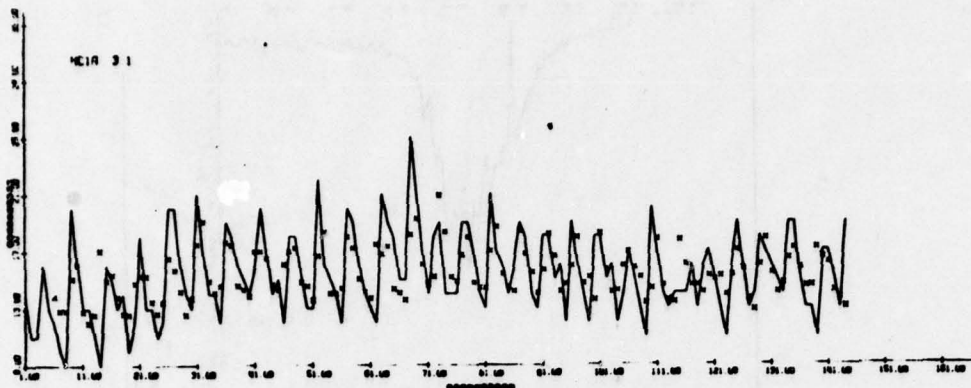


-2-

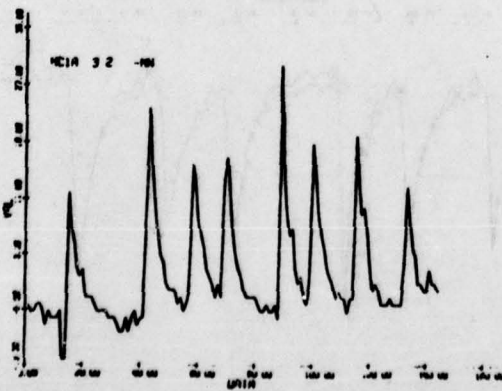
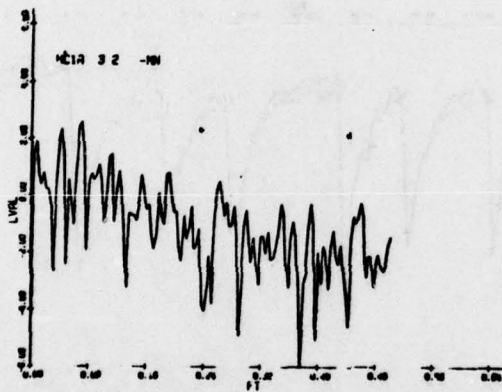


-3-

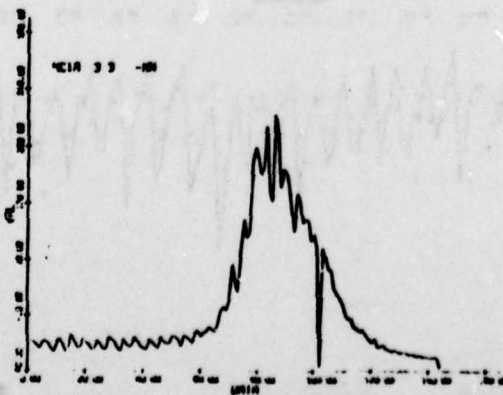
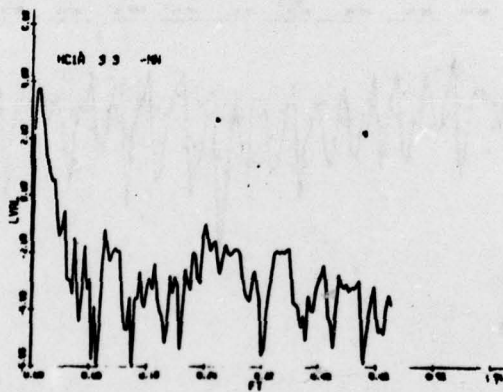
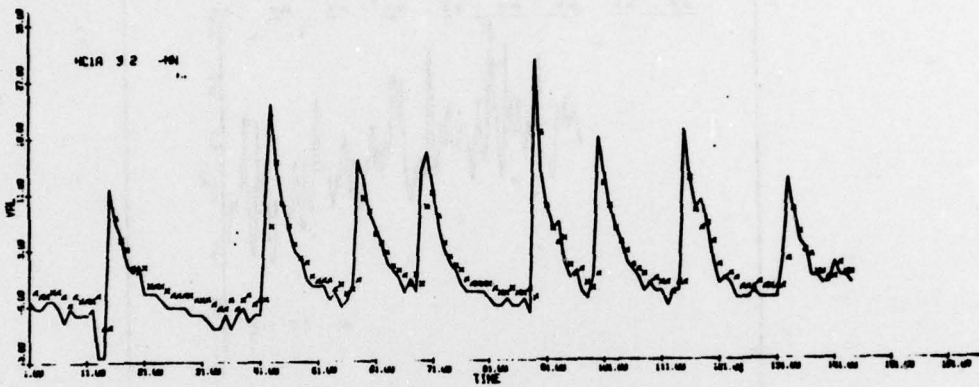
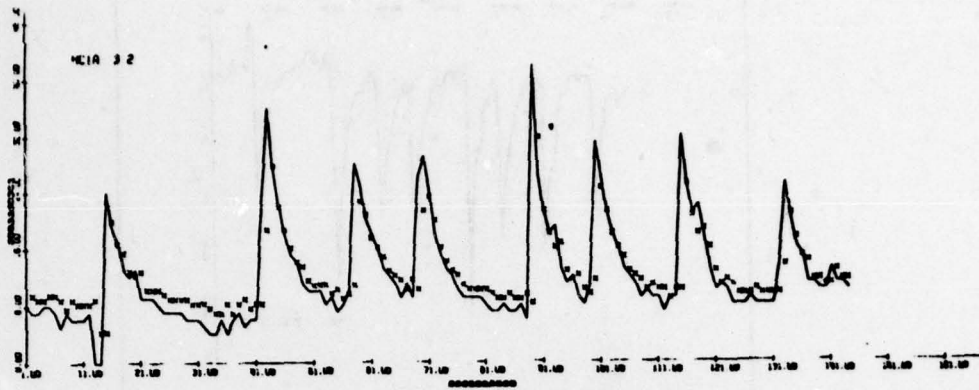


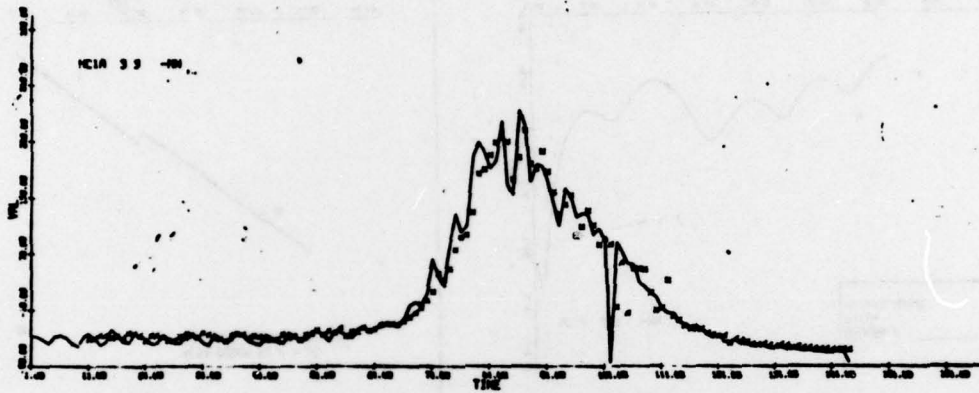
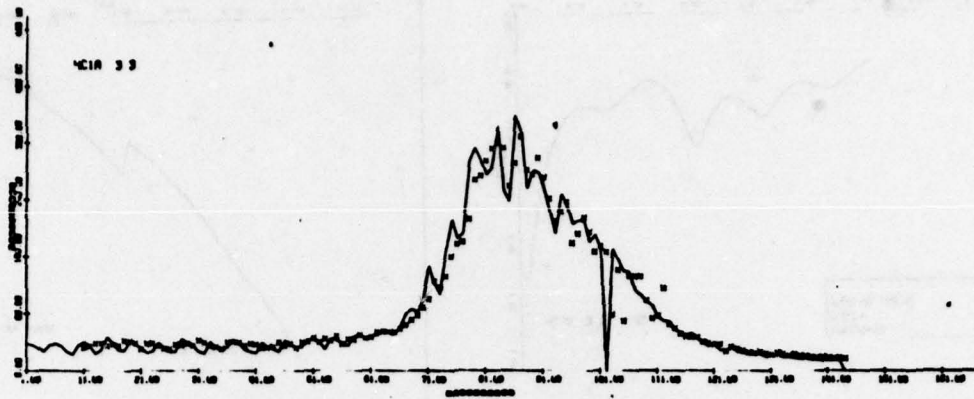


-46-

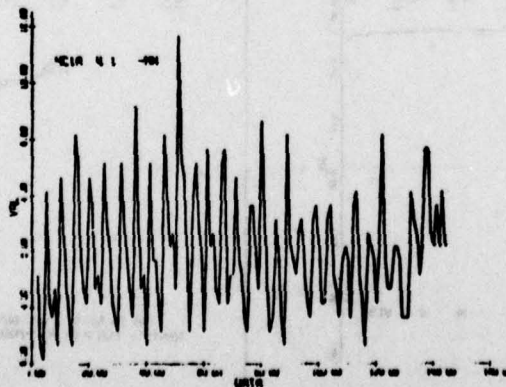
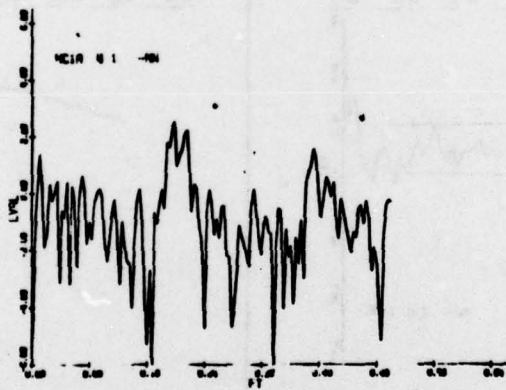


-47-

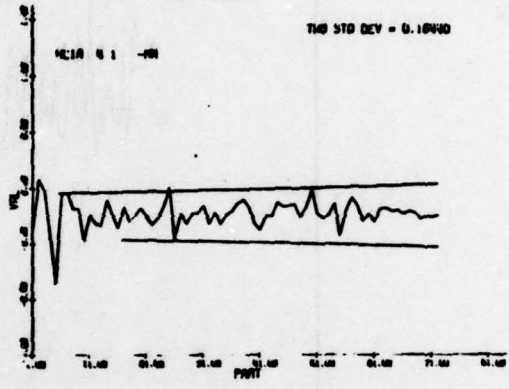
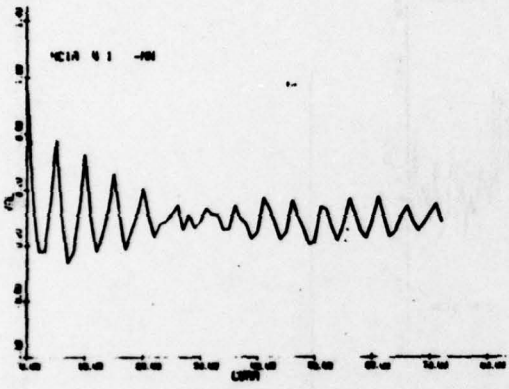
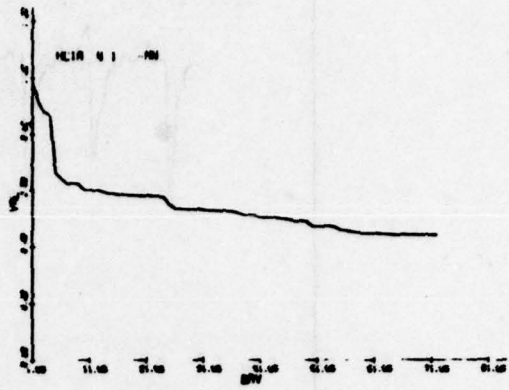
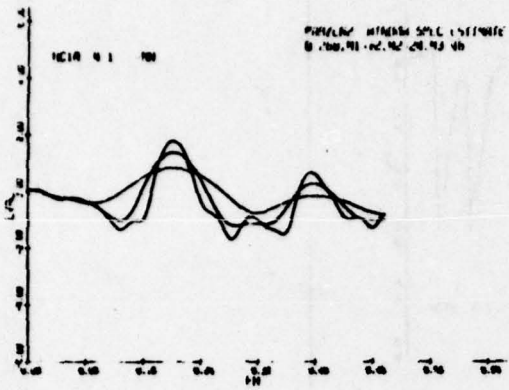




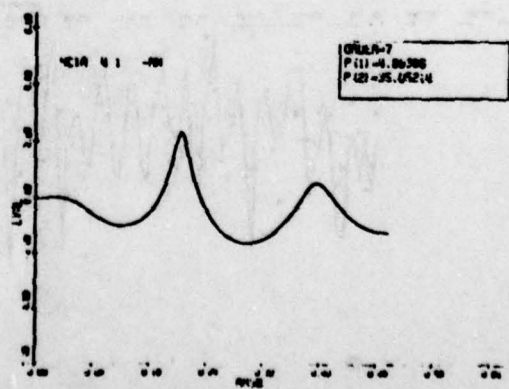
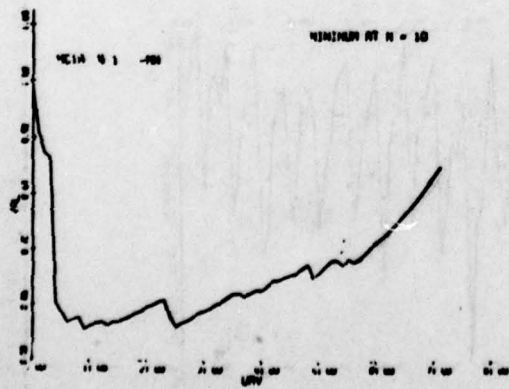
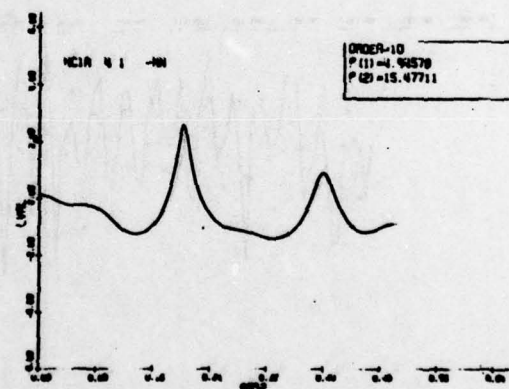
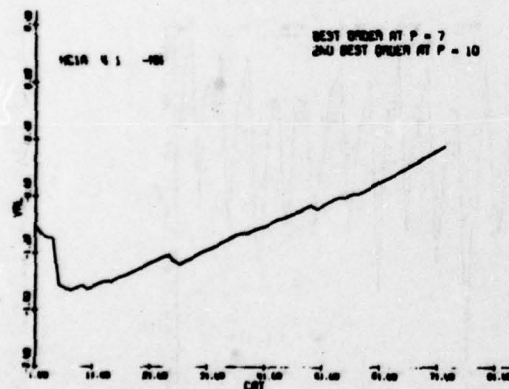
-52-



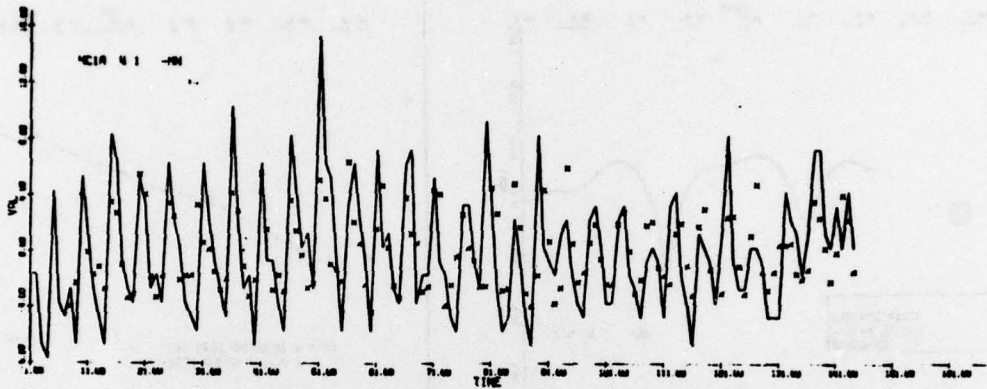
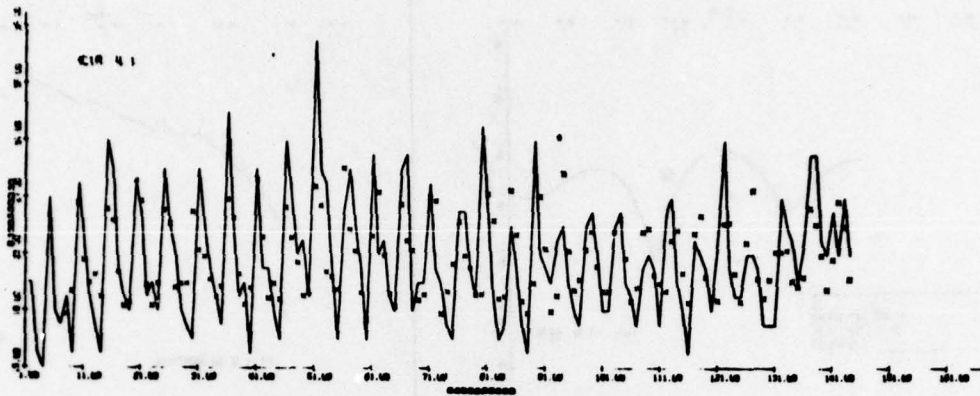
-53-



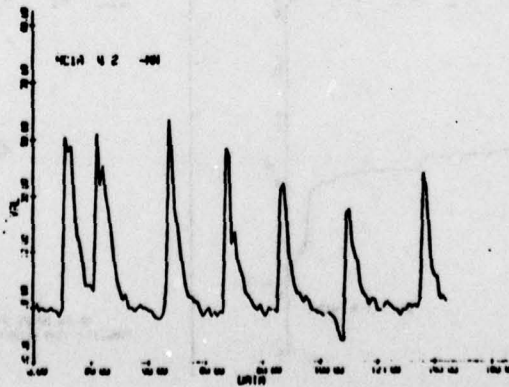
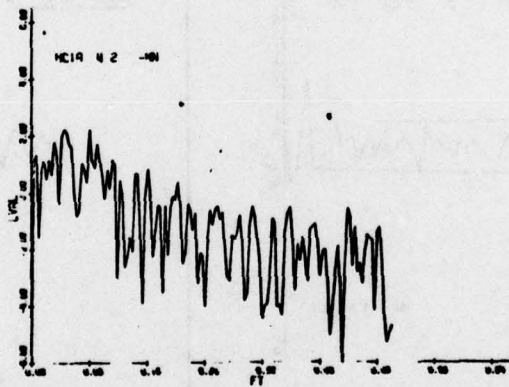
-54-



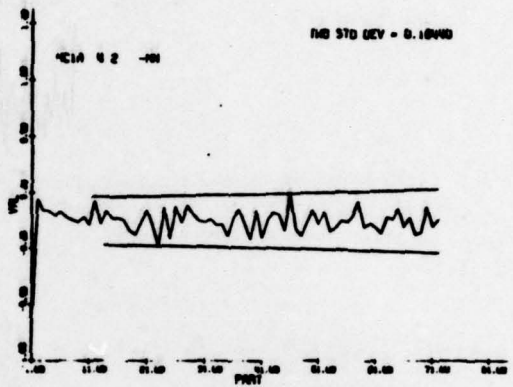
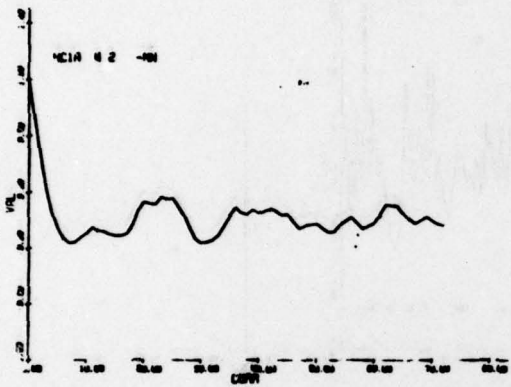
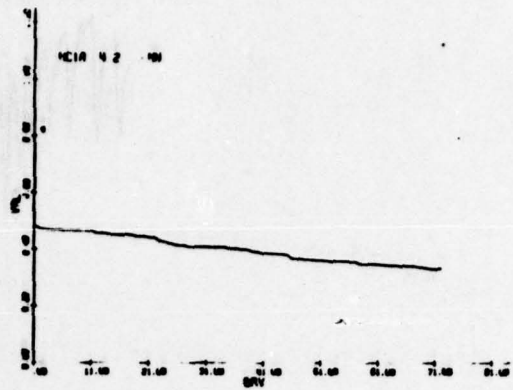
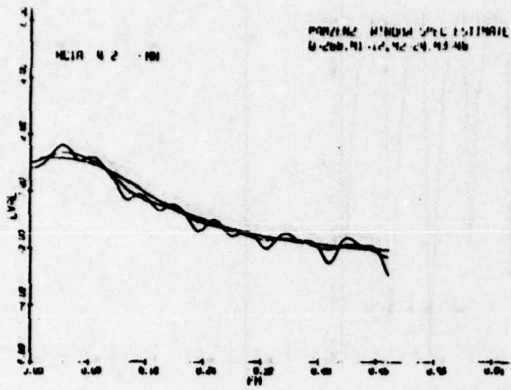
-55-



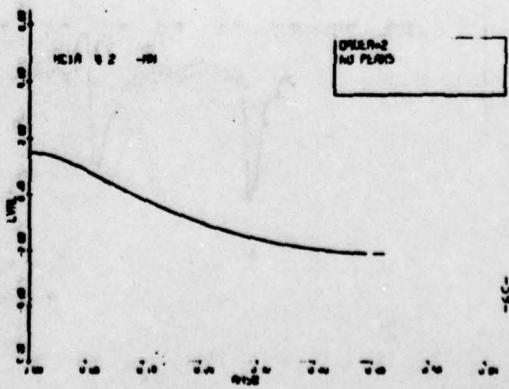
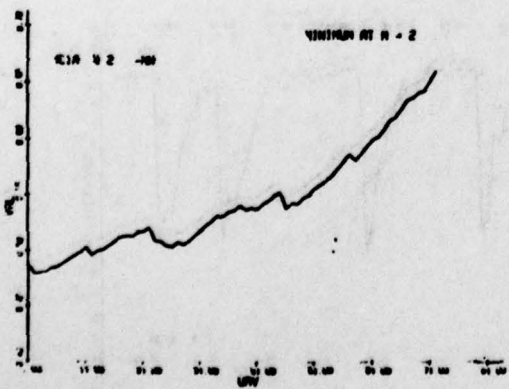
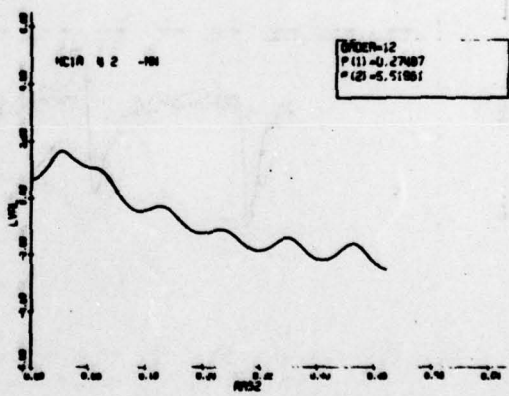
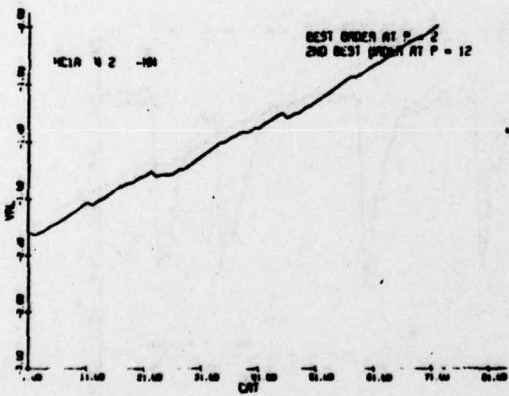
-56-



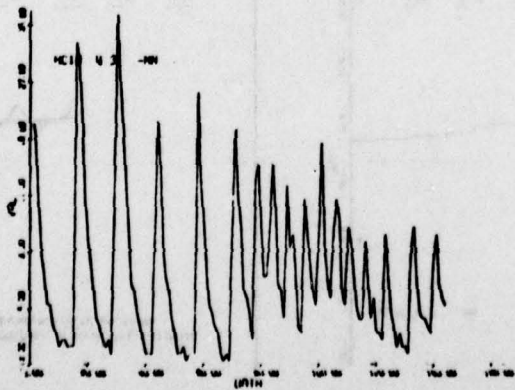
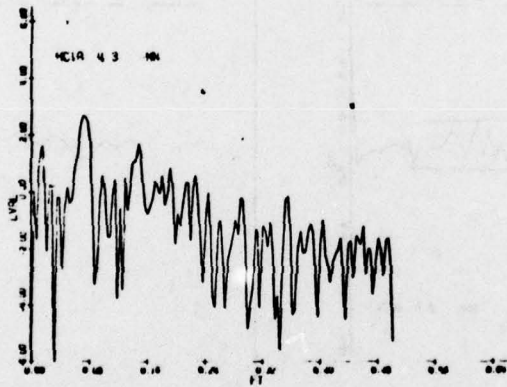
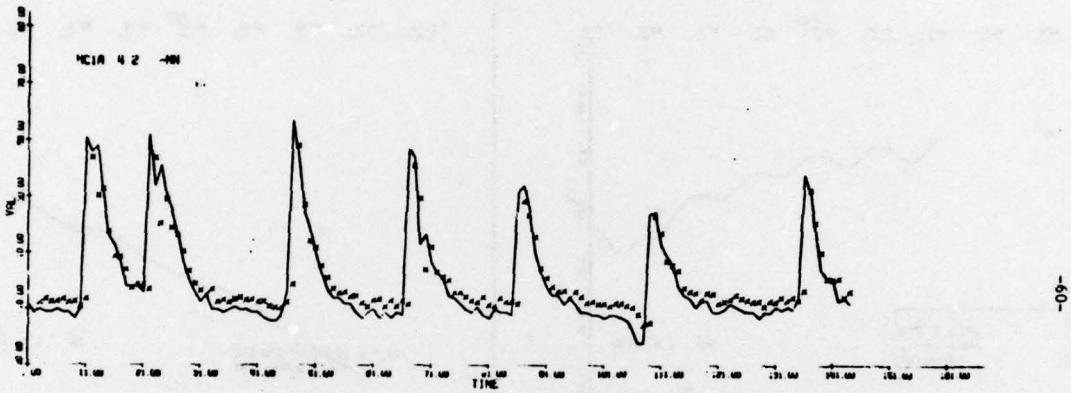
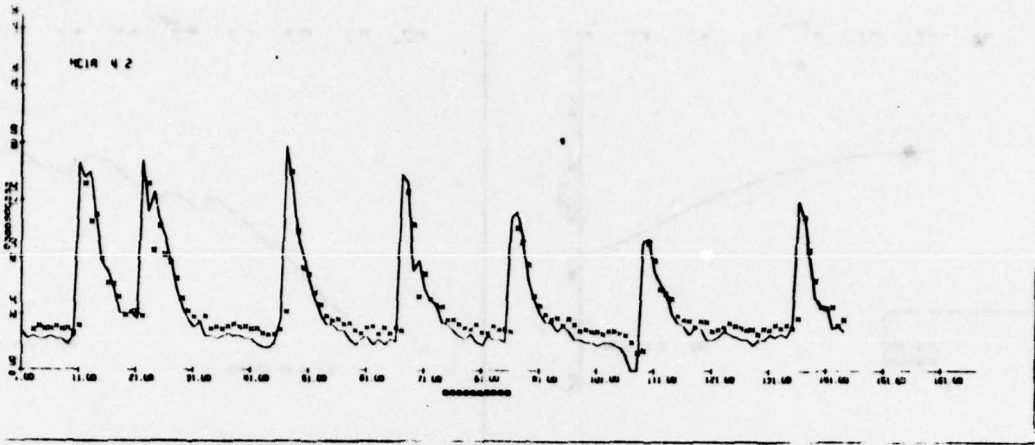
-57-

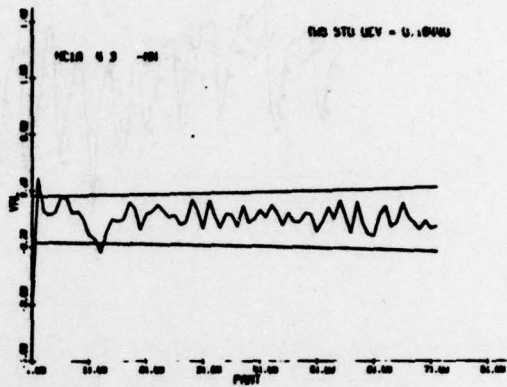
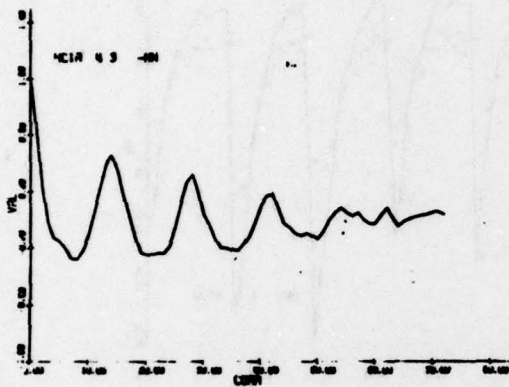
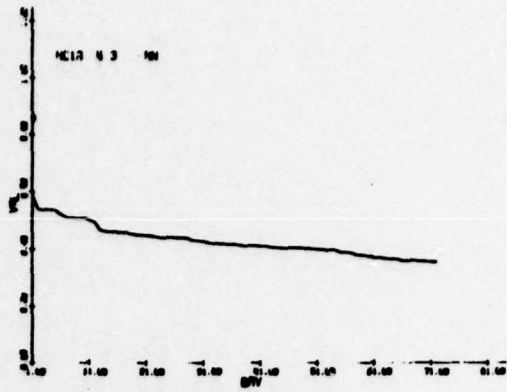
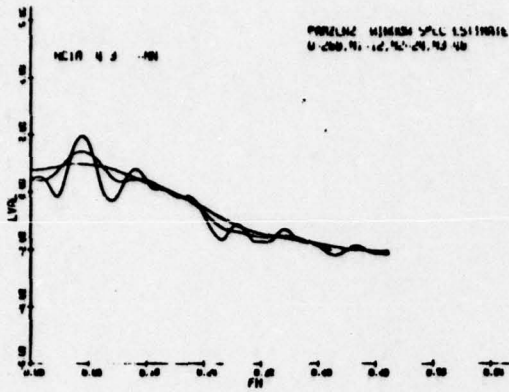


-58-

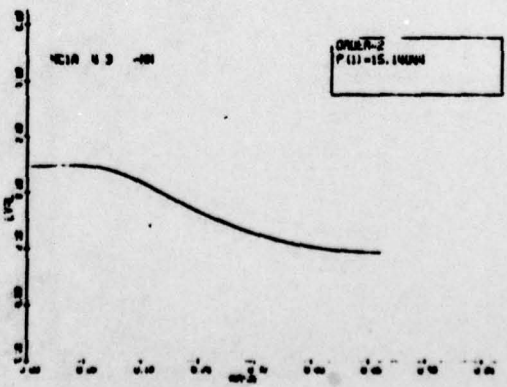
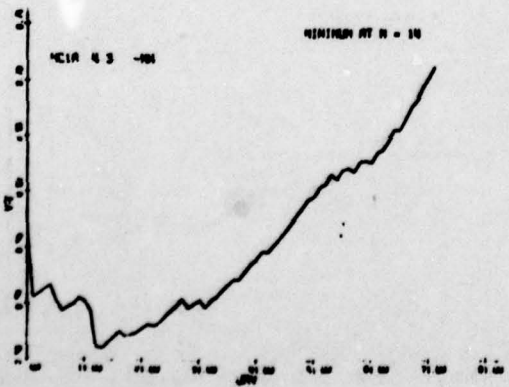
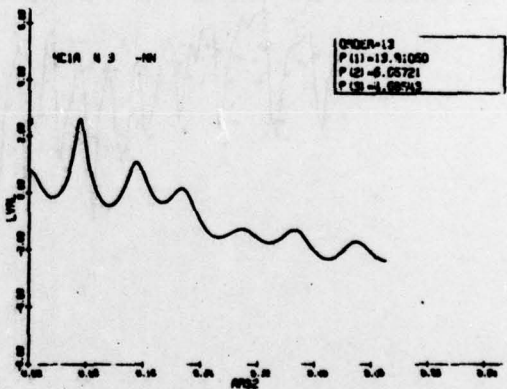
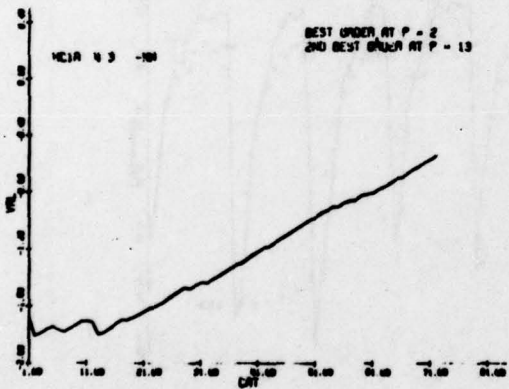


-59-





-02-



-03-

