AD-775 728

THE CHARACTERIZATION AND DISCRIMINATION OF EXOATMOSPHERIC OBJECTS THROUGH MULTIVARIATE EMPIRICAL ANALYSIS

D. C. Collins, et al

Massachusetts Institute of Technology
Lexington, Massachusetts

28 December 1973

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Technology Service Corporation under Purchase Order No. C-599 to M.I.T. Lincoln Laboratory	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

The Characterization and Discrimination of Exoatmospheric Objects Through Multivariate Empirical Analysis

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Subcontract Report (Final Report)

5. AUTHOR(S) (Last name, first name, initial)

Collins, D.C. and Meisel, W.S.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
28 December 1973	173	6

8a. CONTRACT OR GRANT NO. F19628-73-C-0002	9a. ORIGINATOR'S REPORT NUMBER(S) None
b. PROJECT NO. 7X263304D215 c. d.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) ESD-TR-74-42

10. AVAILABILITY LIMITATION NOTICES

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
None	Office of the Chief of Research and Development, Department of the Army

13. ABSTRACT

The objectives of the effort were to investigate implications of data on 1000 radar signatures as to (a) the level of discrimination of signatures of the target class (reentry vehicles) from signatures of the non-target class (other objects), and (b) to indicate the empirically derived decision rules and features of the signatures best for that purpose. The guiding philosophy was to develop an estimate of the best level of discrimination performance inherently possible, given the characteristics of the physical objects as evinced in the signatures, with minimal dependence on the physical model or on subjective inspection of the data. Methods used were the empirical analysis of multivariate data and, in particular, state-of-the-art pattern recognition techniques for feature creation, feature ranking, and discrimination algorithm development. The project demonstrated a systematic quantitative approach (1) to the rapid, cost-effective evaluation of the degree of difficulty of discriminating given non-target classes from given target classes, (2) to the determination of the characteristics of the classes which aid or inhibit discrimination, and (3) to the determination of the nature and degree of degradation of representative discrimination algorithms as the number of returns (pulses) was reduced from 200 down to 3 pulses.

14. KEY WORDS

radar signature studies
reentry vehicles
multivariate empirical analysis
discrimination techniques
exoatmospheric objects

1800

UNCLASSIFIED

Security Classification

ia

Technology Service Corporation
225 Santa Monica Boulevard
Santa Monica, California 90401
(213) 451-8778

THE CHARACTERIZATION AND DISCRIMINATION OF EXOATMOSPHERIC OBJECTS THROUGH MULTIVARIATE EMPIRICAL ANALYSIS

D. C. Collins
W. S. Meisel

28 December 1973

Final Report for

MIT Lincoln Laboratories

(Purchase Order No. C-599)

CONTENTS

CONTENTS (Cont.)

ACKNOWLEDGEMENT

The authors would like to express their appreciation to Harry Knobel and Mike Teener for efforts beyond the call of duty in obtaining some of the numerical results in this report. Conversations with Bob Stone also contributed to the quality of this work.

Dr. William Schoendorf of Lincoln Laboratories, the project monitor, participated in many useful discussions on project direction and provided valuable advice throughout the project.

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Cont.)

LIST OF FIGURES (cont.)

THE CHARACTERIZATION AND DISCRIMINATION OF EXOATMOSPHERIC OBJECTS THROUGH MULTIVARIATE EMPIRICAL ANALYSIS

1.0 INTRODUCTION

This is a Final Report on work under purchase order No. C-599 with MIT Lincoln Laboratories.

The objectives of the effort were to investigate the radar signature data provided by Lincoln Laboratores (A) to determine an estimate of the achievable level of discrimination of target class signatures from non-target class signatures, and (B) to indicate some general procedures for objectively (empirically) creating and/or selecting a set of features of the signature data which jointly contain the most discrimination information. The guiding philosophy was to develop an estimate of the best level of discrimination performance inherently possible, given the characteristics of the physical objects as evinced in the signatures, with minimal dependence on the physical model or on subjective inspection of the data. Methods to be used were the empirical analysis of multivariate data and, in particular, state-of-the-art computer-oriented pattern recognition/data analysis techniques. If successful, the project would demonstrate a systematic approach (1) to the rapid evaluation of the degree of difficulty of discriminating given non-target classes from given target classes, (2) to the determination of the characteristics of the signatures which allow discrimination, (3) to the determination of the characteristics of the signatures which make discrimination difficult, and (4) to the determination of the impact on discrimination performance of a reduction in the number of returns (pulses).

Section 2.0 will describe the problem context in more detail. Section 3.0 will outline a general technical approach, while section 4.0 will describe the specific techniques applied in this study. Section 5.0 will describe an approach to the selection of the "best" group of P features from a set of K features, where "best" is defined in terms of the performance of the resulting discrimination algorithm. Section 6.0 will indicate the degree to which the estimates of performance derived hold up under an independent test. Section 7.0 will discuss the impact of reducing the number of radar returns analyzed on the level of discrimination. Section 8.0 will indicate how one may analyze those samples which cause difficulty in discrimination to discover the physical characteristics and the non-target classes causing the discrimination difficulty; this is essentially the inverse of the problem of determining what aspects aid discrimination. Section 9.0 will discuss briefly a potpourri of topics of interest whose full exploration was beyond the scope of the present report. Finally, a conclusion will summarize the general results of the project.

2.0 THE PROBLEM CONTEXT

2.1 The Discrimination System Objectives and Constraints

The basic problem involves the discrimination of exoatmospheric objects into two classes: target and non-target, using information extracted from radar cross section measurements processed during a very short time period when an object is being tracked. System objectives include the following:

(1) extremely high probability of target detection;

(2) "manageable" probability of false alarm;

(3) extremely short time in which to make a decision; and

(4) high reliability that the decision algorithm developed will perform as expected.

System constraints include the following:

(1) a limitation on the number of measurements (pulses) the system can expend on a given object;

(2) a limitation on the amount of data processing resources and time that can be allocated to the decision function itself; and

(3) a limitation on the resources required to deal with objects deemed to be targets; this relates to the question of a "manageable" number of false alarms.

It is not the purpose of this study to design a final discrimination algorithm for use in an operating system, but the evaluation of potential discrimination performance must be consistent with these system objectives and constraints.

2.2 The Nature and Quantity of the Sensor Data

2.2.1 Hybrid Analysis. Actual radar data from exoatmospheric objects is limited. In situations involving the design of a discrimination system to be effective against a new threat (or, inversely, the design of a new threat), sensor data is often nonexistent. In this context the use of "hybrid analysis" techniques is called for [4].

Hybrid analysis is an approach to combining the advantages of statistical pattern recognition techniques and radar target modeling techniques. Quite simply, it consists of applying pattern recognition to artificial samples of radar returns generated by models of the objects to be recognized. Pattern recognition techniques nave achieved only moderate success in prior radar applications because of the paucity of samples available of the classes of objects to be identified; the number of samples required to take full advantage of statistical pattern recognition is often at least an order of magnitude larger than that available in the form of real data. However, an accurate target model can provide, at minimal expense, essentially unlimited numbers of samples of returns from a given target viewed from different aspects, with different levels of noise, and so on. On the other hand, a very detailed and competent radar model is not utilized fully if only a rough statistic derived from the output of that model is used; pattern recognition techniques tend to exploit the full complexity of the model.

2.2.2 The Nature of the Data. Hybrid analysis was employed to develop the sensor data for this project. The data was generated by a computer-based model which can produce simulated time histories of the radar cross section measurements from range measurements of exoatmospheric objects of particular design. Various assumptions concerning the nature of the objects, their motion through space and the impact of the radar system's relative location and particular design can be simulated (see Figure 2-1). The core data of the model are a series of static range measurements of radar cross section of particular objects. The model can be easily altered to model different threats, different geometries, and different radar systems.

The data employed by TSC in this study consisted of the principal polarization (PP) and the orthogonal polarization (OP) measurements of radar cross section in dbsm (see Figure 2-2). No phase data was employed. Each signature (time history of returns from a particular object) consisted of 200 RCS measurements (200 pulses) sampled equally spaced over a simulated time period of 10 seconds. Thus the "effective" pulse repetition frequency was 20 Hz. (The modeled radar system could have a higher PRF and simply be time sharing its resources as it tracks a number of objects.) In view of the system objectives, the impact on discrimination performance of other "effective" PRF's and other types of sampling is of some interest.

The "objects" modeled consisted of one object which can be described as the "target" and seven objects which can be described as "non-targets." Knowledge of the physical nature of these objects may be valuable both in guiding the analyst's intuition in formulating heuristic or subjective features and in interpreting the results of an inverse analysis to determine the physical characteristics which degrade performance of the decision algorithm for certain objects. However, the emphasis in this study was on objective feature creation and selection methods and thus an in-depth discussion of the physical "threat design" implications of the study is not included in this report.

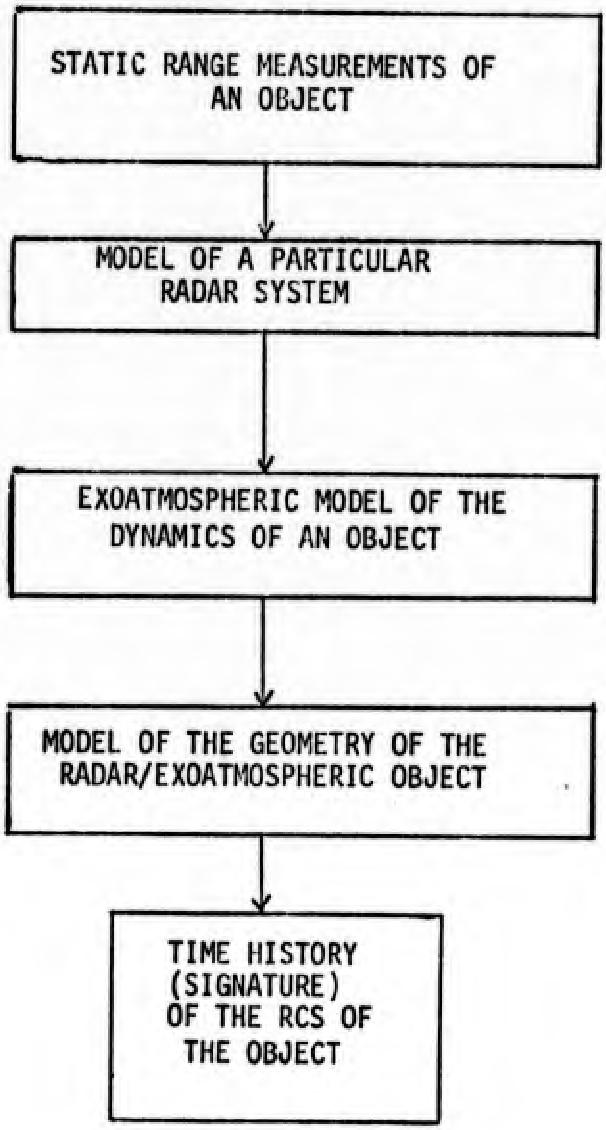


Figure 2-1: Structure of the model used to simulate the RCS signature (time history) of exoatmospheric objects.

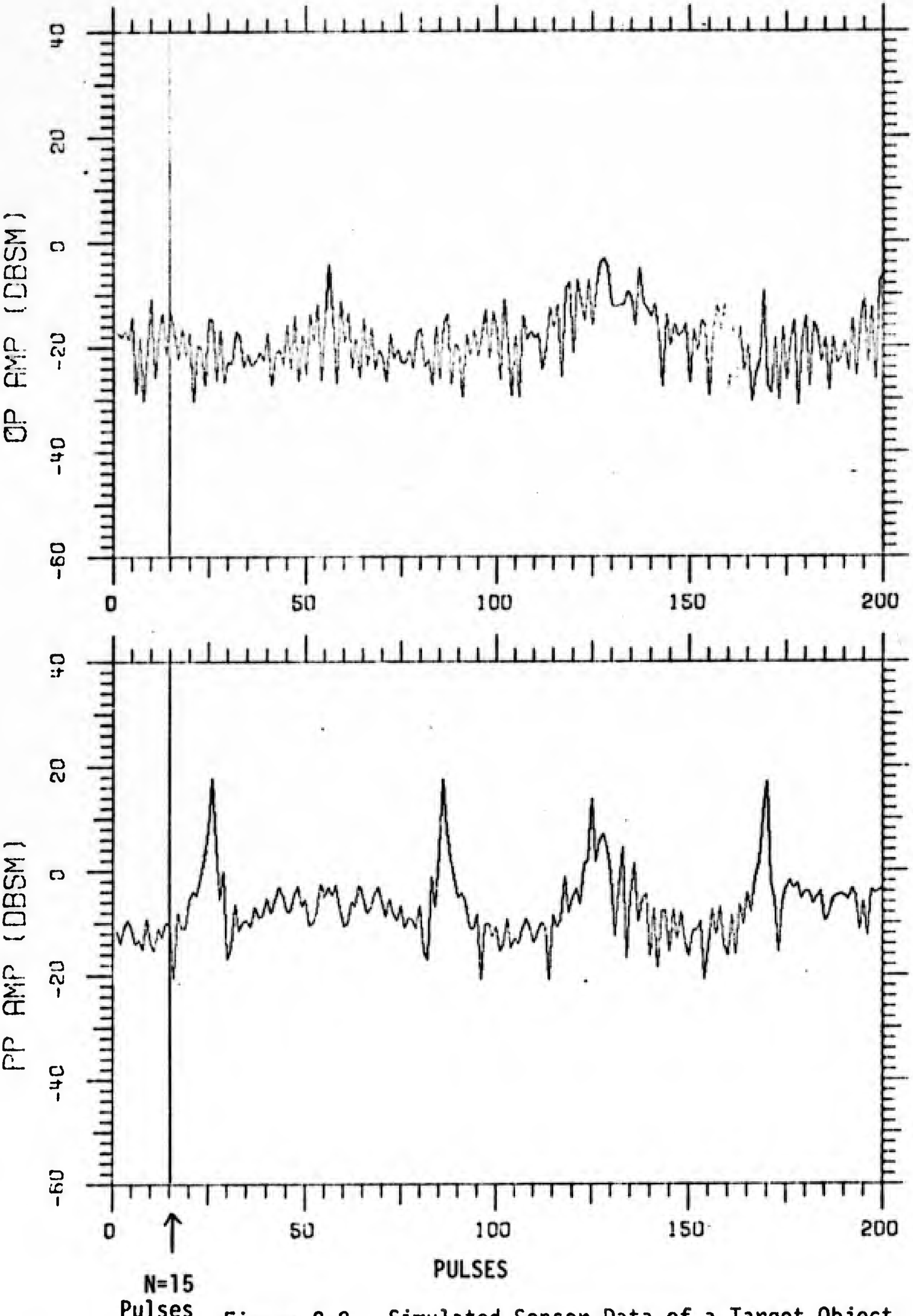


Figure 2-2: Simulated Sensor Data of a Target Object

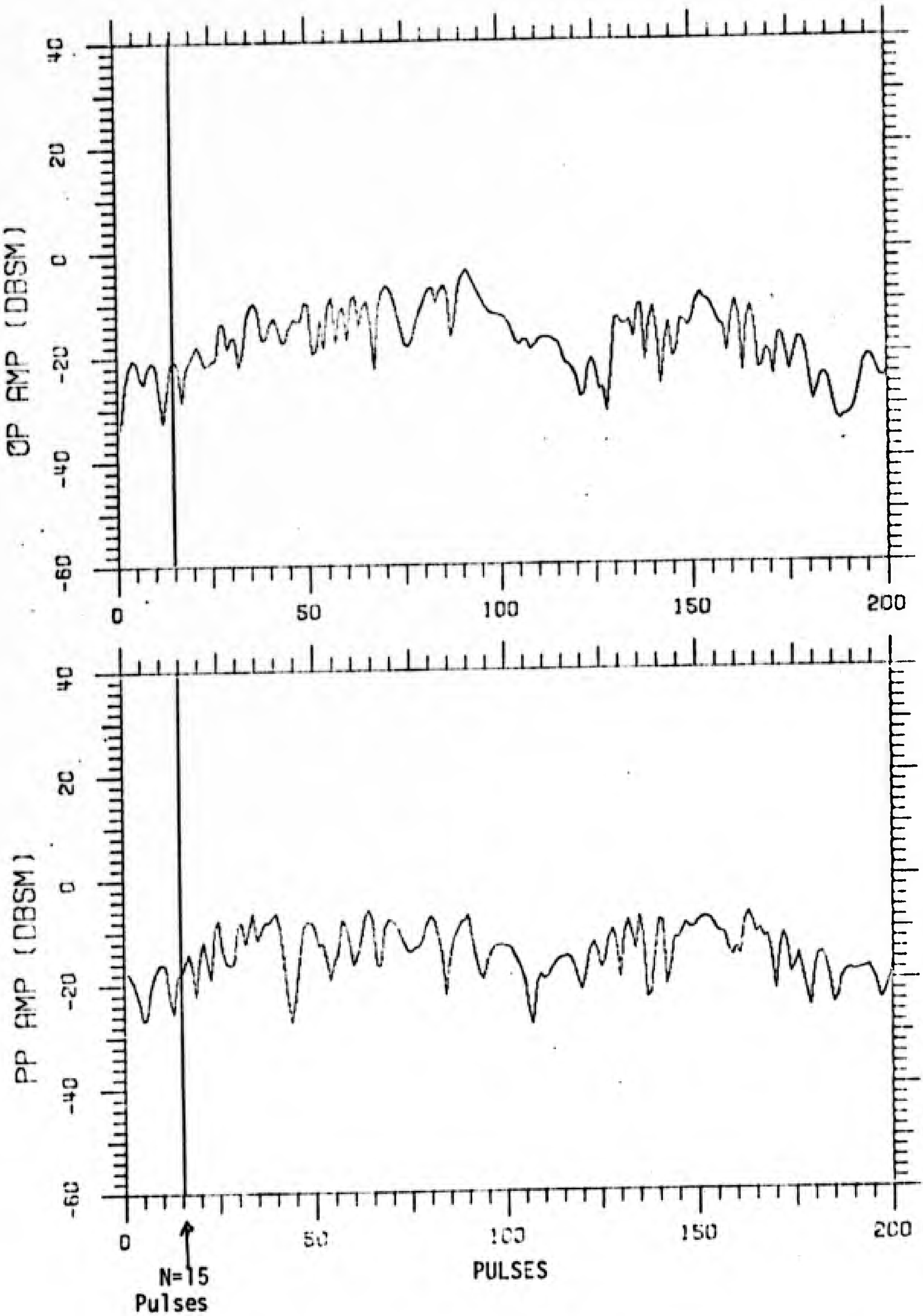


Figure 2-3: Simulated Sensor Data of a Non-Target Object

2.2.3 The Quantity of the Data Used. The data employed in the study consisted of a total of 500 signatures of the target object and 100 signatures each of the 7 non-target objects. Each signature consisted of 400 numbers: 200 radar cross section (RCS) amplitude measurements in dbsm for both principal (PP) and orthogonal (OP) polarizations. (A pair of numbers, OP and PP, for a particular point in time are referred to in the text as a "pulse", i.e., they represent the response to a single radar pulse.)

The data was divided into two sets: 80% of the signatures were placed in the "design set" (learning set, training set) and the remainder in the "test set." The design set then consisted of 400 target signatures and 560 non-target signatures. The test set was used to validate the decision algorithm developed on the design set. Figure 2-4 summarizes the above comments.

2.3 Objectives of the Data Analysis Study

2.3.1 General Objectives. Briefly, the objectives were to investigate the implications of the signature data in order to:

(1) Determine the extent to which a particular set of target objects can be distinguished from a particular set of non-target objects.

(2) Determine the definition of features with good discrimination capability. This involved:

(a) Definition of heuristic (or subjective) features

(b) Design of completely objective features

(c) Design of objective combinations of objective and subjective features.

(3) Determine the best groups of features of a given dimension (low dimensionality being related to efficiency and reliability of the final decision algorithm)

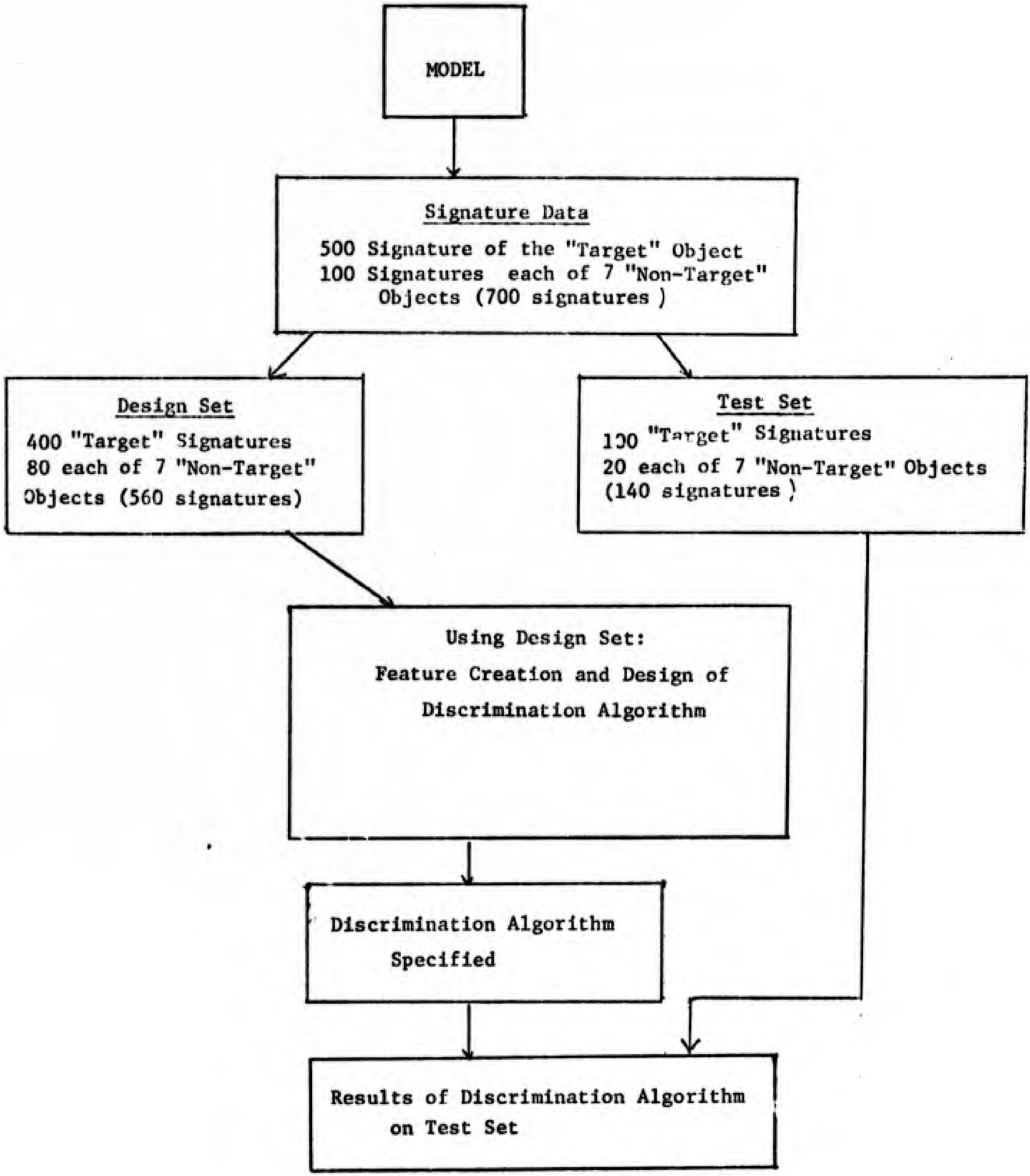


Figure 2-4: Design Data Set/Test Data Set Breakdown

(4) Determine the nature of the degradation of the discrimination algorithm's performance as the number of pulses decreases.

(5) Illustrate the manner in which data analysis techniques might be used to help indicate the design of non-target objects which are difficult to discriminate.

A topic which received special emphasis was the objective feature creation technique. The guiding philosophy of the project was to develop high-discrimination-information-content features with minimum dependence on the physics of the system and subjective inspection of the data.

2.3.2 Specific Objectives. Due to the usual constraint of a limited amount of resources which could be applied to the project, some narrowing of the possible generality of the above objectives was necessary. The following are the specializations of the objectives that were made:

a) Number of pulses (RCS measurements)

The basic data analysis tasks, including feature creation, feature ranking, decision algorithm development, and validation of results, were carried out using 15 pulses. That is, although the signature data provided contained OP and PP RCS measurements for 200 pulses, the study was focused on truncated signatures with 15 pulses (30 pieces of data, 15 OP and 15 PP). However, a great deal of additional analysis was performed on the cases where the number of pulses equaled 3, 6, and 200. (Quite good results were achieved for the 3 pulses case.) Analysis was performed to determine the degree to which the decision algorithm degraded as the number of pulses was reduced. This analysis is reported in Section 7.0.

b) Spacing of the returns (pulses)

Given that 200 time samples of RCS were available in each signature and that the study was to focus on 15, a question arose as to whether the 15 pulses selected should be maximally separated in time (and therefore maximally uncorrelated) or contiguous in time at the effective PRF of 20 Hz. It was decided to perform the analysis with contiguous pulses. (Preliminary analysis indicates improved performance can be achieved by a wider spacing of pulses.)

c) Combining of non-target subclasses

While some initial analysis was focused on treating each of the non-target objects separately (e.g., separate estimation of class-conditional probability densities, etc.), it was determined that for the purposes of the study, it was adequate to combine the non-target subclasses into one class. It was noted that, in general, the ensemble of non-target classes often had a simpler distribution in the feature space than did each subclass. The basic analysis is therefore performed in a two-class discrimination context. After the basic analysis was performed in the two-class context, however, a further analysis was performed to determine how (for the sets of features of interest) the class conditional densities of the individual non-target objects were distributed relative to the target class. These results are reported in Section 8.0 where the nature of the "inverse problem" of threat design is discussed.

3.0 AN OUTLINE OF AND MOTIVATION FOR A GENERAL TECHNICAL APPROACH

In the sections which follow we briefly describe the steps followed in the analysis of the data, the motivation for each step and some of the data analysis tools employed. Before proceeding, a few general aspects of the study are perhaps worth mentioning. On the one hand the existence of an already developed, implemented and tested set of analysis tools, TSC's Advanced Data Analysis Library, permitted the development and analysis of a large number of features (functional transformations of the signature data) in a rapid response/cost-effective fashion. On the other hand, however, the degree of refinement of the analysis procedures at each step had to be adjusted to conform to the scope and level of effort of the contract. The overall thrust of the study was to illustrate a complete data analysis procedure, from data acquistion to an analysis of the impact of the reduction of the number of pulses on the performance of a discrimination algorithm. It was intended that this study provide representative lower bounds on the level of discrimination achievable by a study of broader scope.

3.1 Reduction of Dimensionality, Feature Creation and Selection

The overall discrimination problem is illustrated in Figure 3.1. The objective is to design an algorithm (black box) which takes as inputs M raw sensor signature data measurements and outputs a correct classification: "target" or "non-target." In the study the signature data contained 200 returns and the original M was 400 (200 PP, 200 OP) data measurements. This data was reduced to M=30 (15 PP, 15 OP) data measurements by focusing the initial portion of the study on the first 15 returns or pulses.

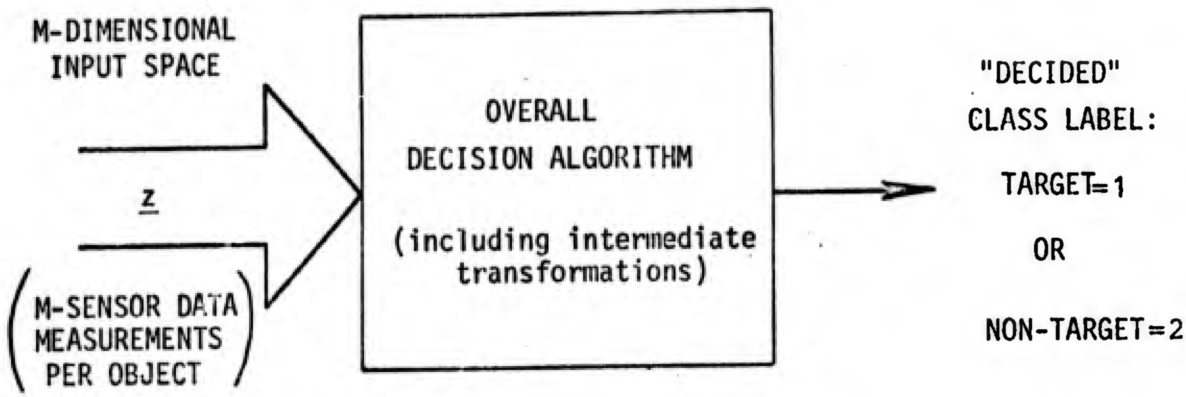


FIGURE 3.1 THE OVERALL DISCRIMINATION PROBLEM

There exist a large number of pattern recognition techniques (described in [1]), both direct and indirect, that can be used to develop a discrimination rule which will yield a class label when presented with a set of M data measurements as inputs. Unfortunately these techniques all have the same characteristic: they are vulnerable to the "curse of dimensionality." (See [1, Chapter 1] for a discussion of this problem.)

The curse of dimensionality relates to the fact that the development of a decision algorithm which performs a mapping directly from the M-dimensional measurement space to a 1-dimensional, class label space, often can require an "infeasibly large" number of samples or signatures of each class of objects if M is "sufficiently large." Of course the number of samples required for a given M depends on a variety of considerations such as the nature of the decision algorithm being developed, the nature of the distribution of the data itself, and its intrinsic dimensionality (see [1, Chapter 1]). The question of what constitutes an "infeasibly large" number of signatures for a given M can only be answered in terms of the cost of acquiring the samples and the impact of the sample size on the computational cost involved in developing or designing the decision rule. In those cases where M is sufficiently large to evoke the curse of dimensionality, the usual remedy employed is to effectively reduce the dimensionality of the problem by transforming (preprocessing) the M-dimensional measurement space into a P-dimensional feature or pattern space where P is much smaller than M. The decision rule is then designed using the set of sample signatures (represented in lower P-dimensional feature space) as input (see figure 3.2). The topics of "feature extraction" and "feature selection" can be subsumed in the general problem of the design of a dimensionality reducing transformation.

3.2 Ingredients of a Feature Creation/Selection Transformation Design Procedure

There are an infinity of possible transformations from an M-dimensional space to a P-dimensional space; a natural question arises as to how one goes about designing the "best" such transformation. The criteria that might be involved in defining what is meant by "best" might include:

1) Performance: Measures of the degree to which the transformation retains the discrimination information present in the measurement data.
2) Validity: Measures of the ability of the decision rule, whose design is based on the representation of the design set samples in the P dimensional feature space, to perform on samples in the test set.
3) Cost: Cost in time and other system resources to perform the transformation.

An additional criteria might involve the impact of the transformation on the complexity of the decision rule and thus on the amount of system resources required for its implementation.

Given a set of R sample signatures (in our case $R \cong 1000$) divided between the two classes (target and non-target) each signature being represented by a point in a measurement space of dimension M (in our case most often M= 30), the design of a particular dimensionality reducing transformation involves a procedure consisting of the following steps (see figure 3.3):

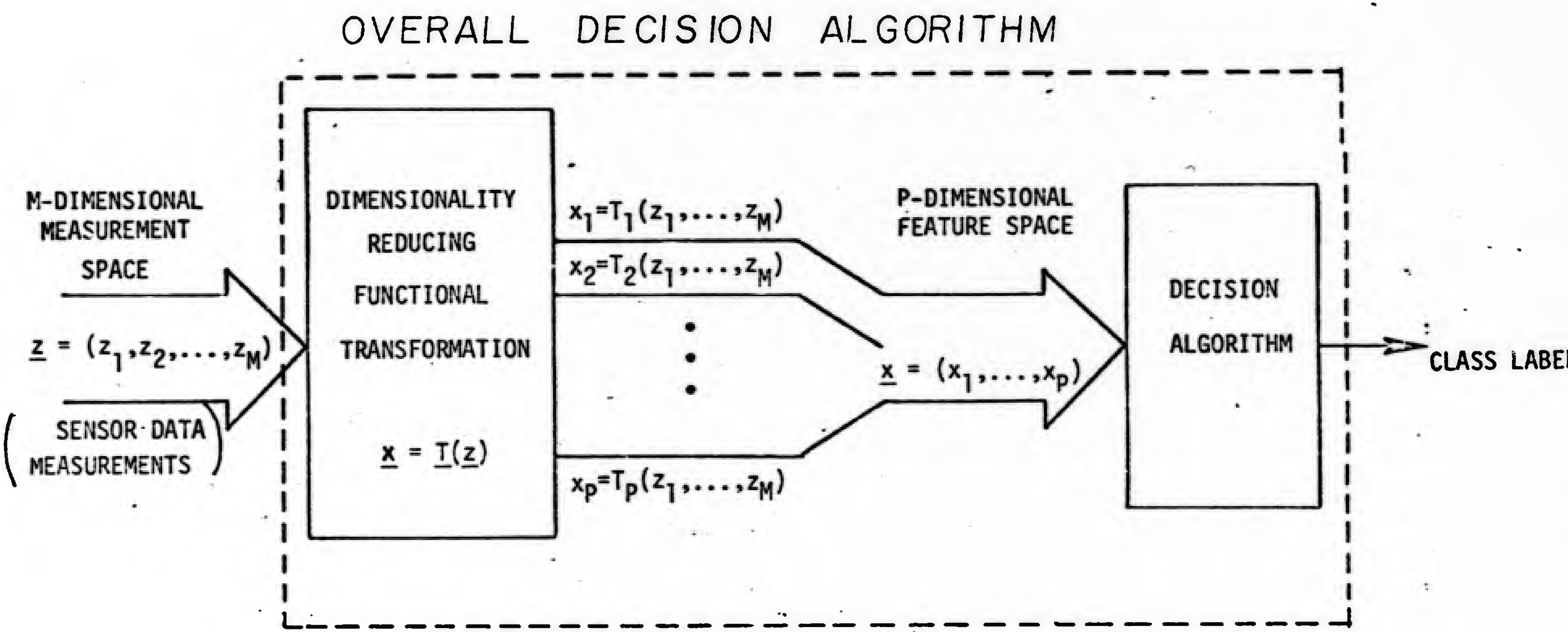


FIGURE 3.2 A DIMENSIONALITY REDUCING TRANSFORMATION

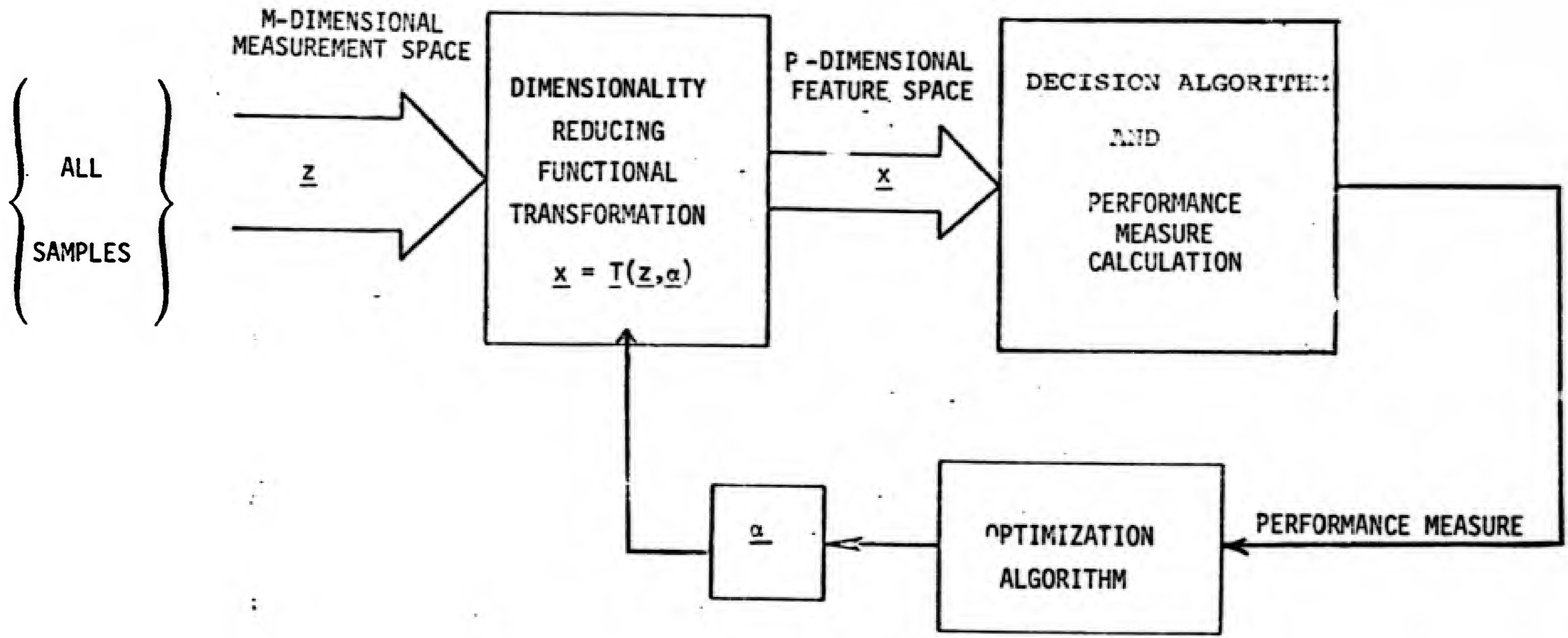


Figure 3.3 The $\underline{\alpha}$ parameters determine the design of dimensionality reducing transformation.

(1) Specification of the nature of the functional form of the dimensionality reducing transformation which will map the M-dimensional measurement space into a P-dimensional feature space. This includes specification of those aspects or parameters of the functional form which are design variables to be determined by an optimization procedure.

(2) Specification of a single criterion which measures the relative "goodness" of the dimensionality reducing transformations generated by variations in the designable aspects or parameters of the transformation. The criterion can involve measures of performance of a decision algorithm on the design set of samples.

(3) Specifications of an optimization technique or approach to select the best transformation relative to the criteria specified in (2).

We note for the sake of clarity that heuristic feature definition does fit this formulation as discussed in the following section; however, in this case, the "parameters" in step(1) are somewhat artificial.

4.0 THE SPECIFIC TECHNICAL APPROACH OF THIS STUDY

Specifications of a particular set of chioces of each of the previously discussed three ingredients (the form of the transformation, performance measure, and optimization procedure) result in a particular data analytic approach to the basic discrimination problem. Particular individual choices are predicated on the nature of the data and the discrimination problem, the nature of other ingredients in the procedure, the level of sophistication necessary, and the amount of resources that can be applied to the problem. In this illustrative study, several simplifications and specializations of the most powerful and general approaches were at times necessary to remain within the level of effort of the study. In what follows we describe the selections and specializations made in this study.

4.1 Specification of the Nature of the Data-to-Feature Transformation

In this study we primarily limited ourselves to combinations of four types of transformations:

1) A set of heuristically defined transformations of the measurement data into single features.

2) A parameterized, continuous, piecewise linear transformation of the measurement data into a single feature.

3) A parameterized linear transformation of heuristic features into a single feature.

4) A fourth transformation which amounts to a selection of P features from a set of K features (heuristically or objectively defined). This transformation can be represented as a special matrix product of the K dimensional feature vector (see figure 4.1). That is,

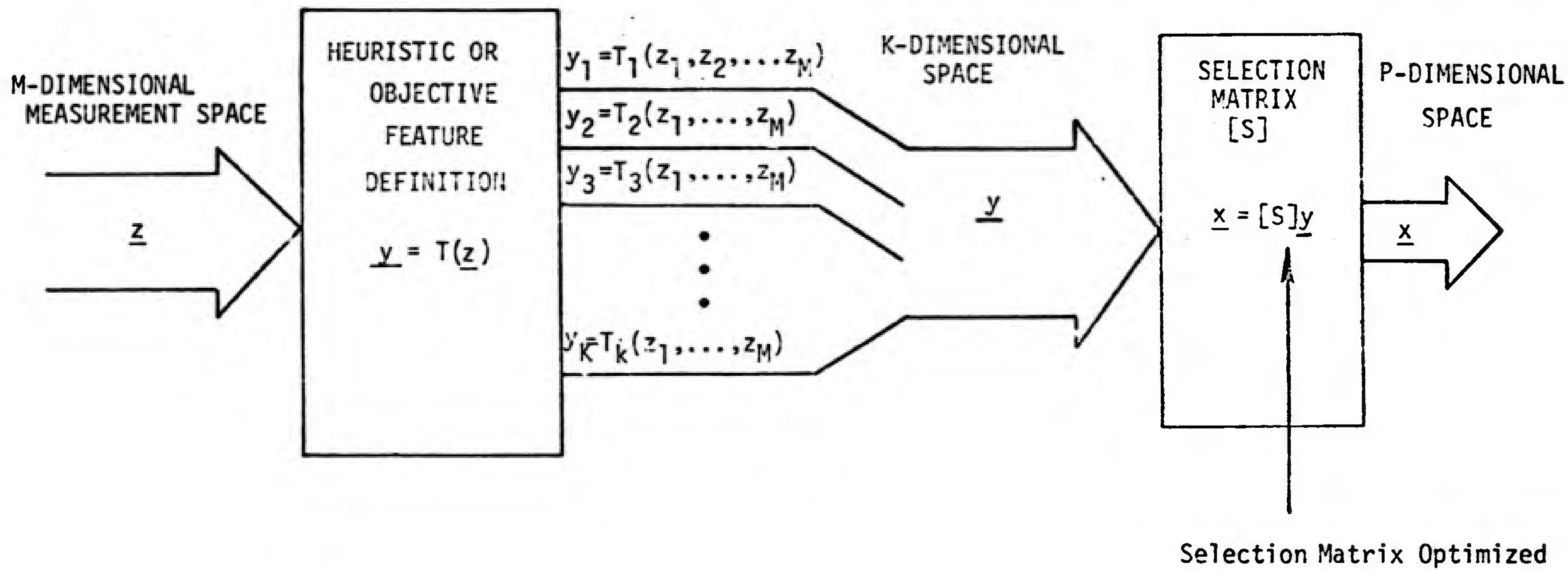
$$\underline{x} = [S]\underline{y} \qquad \text{where}$$

$$\underline{x} \text{ is P dimensional, } \underline{y} \text{ is K dimensional}$$

and [S] is a special selection matrix of dimension PxK made up of "1"'s and "0"'s; a single "1" per row and no more than a single "1" per column.

The manner in which a design of this selection matrix was determined is described in Section 5.0.

Each of the first three types of transformation is described below with the particular functional forms specified.



The PXK dimensional selection matrix maps the K-dimensional feature space into a P-dimensional space made up of a selection of P of the K features.

e.g., $P=3$, $K=8$

$$[S] = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\underline{x} = [S]\underline{y} = \begin{pmatrix} y_1 \\ y_3 \\ y_6 \end{pmatrix}$$

Figure 4.1 - Feature selection optimization

4.1.1 Heuristically Defined Features

The heuristic features we chose for evaluation are listed in Table 4-1. They are, for the most part, straightforward functions of the signature. They include:

(a) means of various types,

(b) measures of variance, third, and fourth moments,

(c) variation of the signatures,

(d) kurtosis and skewness, and

(e) relative (ratio) measures involving means, variances, etc., for the principal polarization versus the orthogonal polarization.

Frequency domain features were considered initially but discarded after some testing due to the fact that the study focused on numbers of pulses in the three to fifteen range. Additional heuristic features were suggested by the analysis of the data but were not implemented and tested due to limitations on time and scope of the project. (Note that the formulas include the number of pulses (N) as a parameter and thus are defined for N=3, 6, 15, 200, etc).

Quite a few of the heuristic features that were analyzed proved to be extremely valuable in terms of their information content. Many have not, to the authors' knowledge, been studied previously.

Table 4-1

Definition of Heuristic Features

N = Number of pulses (returns) used in the feature.

OP = RCS of orthogonal polarization in dbsm

PP = RCS of principal polarization in dbsm

XSQ = 10**(OP/10) = RCS of orthogonal polarization in M^2

YSQ = 10**(PP/10) = RCS of principal polarization in M^2

ASQ = XSQ + YSQ

X = 10**(OP/20)

Y = 10**(PP/20)

ANGLE = atan $\left[\frac{Y}{X}\right]$

Means - Arithmetic and Geometric

PPOW = mean PP(RCS) in M^2 $\quad \frac{1}{N}\sum_{i=1}^{N} \text{YSQ}$

OPOW = mean OP(RCS) in M^2 $\quad \frac{1}{N}\sum_{i=1}^{N} \text{XSQ}$

TPOW = PPOW + OPOW $\quad \frac{1}{N}\sum_{i=1}^{N} \text{ASQ}$

PGM = mean PP(RCS) in dbsm $\quad \frac{1}{N}\sum_{i=1}^{N} \text{PP}$

OGM = mean OP(RCS) in dbsm $\quad \frac{1}{N}\sum_{i=1}^{N} \text{OP}$

TGM = mean of "total" OP + PP(RCS) in dbsm $\quad \frac{10}{N}\sum_{i=1}^{N} \log_{10}[\text{ASQ}]$

Table 4-1 (Continued)

Measures of Dispersion

VPPOW	-	variance of PP in M^2	$\left[\frac{1}{N}\sum_{i=1}^{N} YSQ^2\right] - PPOW^2$
VOPOW	-	variance of OP in M^2	$\left[\frac{1}{N}\sum_{i=1}^{N} XSQ^2\right] - OPOW^2$
VTPOW	-	variance of "total RCS" (PP + OP) in M^2	$\left[\frac{1}{N}\sum_{i=1}^{N} ASQ^2\right] - TPOW^2$
VPGM	-	variance of PP in dbsm	$\left[\frac{1}{N}\sum_{i=1}^{N} PP^2\right] - PGM^2$
VØGM	-	variance of OP in dbsm	$\left[\frac{1}{N}\sum_{i=1}^{N} OP^2\right] - OGM^2$
VTGM	-	variance of "total RCS" (OP + PP) in dbsm	$\left[\frac{100}{N}\sum_{i=1}^{N} (\log_{10} ASQ)^2\right] - TGM^2$
MDPP	-	mean deviation of PP	$\frac{1}{N}\sum_{i=1}^{N} \lvert PP - PGM\rvert$
MDØP	-	mean deviation of OP	$\frac{1}{N}\sum_{i=1}^{N} \lvert OP - OGM\rvert$
MVARPP	-	mean variation of PP	$\frac{1}{N}\sum_{i=2}^{N} \lvert PP(i) - PP(i-1)\rvert$
MVAROP	-	mean variation of OP	$\frac{1}{N}\sum_{i=2}^{N} \lvert OP(i) - OP(i-1)\rvert$
PP3M	-	3rd moment of PP	$\frac{1}{N}\sum PP^3$
OP3M	-	3rd moment of OP	$\frac{1}{N}\sum OP^3$

Table 4-1 (Continued)

PP4M	- 4th moment of PP	$\frac{1}{N}\sum PP^4$
OP4M	- 4th moment of OP	$\frac{1}{N}\sum OP^4$
PPSKEW	- skewness of PP	$\frac{PP3M}{VPGM^{3/2}}$
PPKUR	- kurtosis of PP	$\frac{PP4M}{VPGM^2} - 3$
OPSKEW	- skewness of OP	$\frac{OP3M}{VOGM^{3/2}}$
OPKUR	- kurtosis of OP	$\frac{OP4M}{VOGM^2-3} - 3$
FLPKR	- ln'|PPKUR| where	$\ln'\lvert a\rvert = \left\{\begin{array}{ll} (\ln a) + 1, & \text{if } a > 1 \\ a & \text{, if } \lvert a\rvert < 1 \\ -(\ln(-a)) - 1, & \text{if } a < -1 \end{array}\right\}$
FLOPSK	- ln'|OPSKEW|	

Table 4-1 (Continued)

Various Ratios

RATVDP	-	ratio of variation of OP to PP	$\frac{MVAROP}{MVARPP}$
RATVPVP	-	ratio of variation of PP to variance of PP	$\frac{MVARPP}{VPGM}$
RATVOVO	-	ratio of variation of OP to variance of OP	$\frac{MVAROP}{VOGM}$
RVPPOW	-	relative dispersion of PP	$\frac{VPPOW^{1/2}}{PPOW}$
RVOPOW	-	relative dispersion of OP	$\frac{VOPOW^{1/2}}{OPOW}$
RVPGM	-	relative dispersion of PP	$\frac{VPGM^{1/2}}{PGM}$
RVOGM	-	relative dispersion of OP	$\frac{VOGM^{1/2}}{OGM}$
XANGL	-	average of angle	$\frac{1}{N}\sum_{i=1}^{N} ANGLE$
VANGL	-	variance of ANGLE	$\left[\frac{1}{N}\sum_{i=1}^{N} ANGLE^2\right] - XANGL^2$
XMPA	-	angle associated with the ratio of √(PPOW/OPOW)	$\tan^{-1}\left[\left(\frac{PPOW}{OPOW}\right)^{1/2}\right]$
OPRAT2	-	average of ratio of OP (M^2) to PP (M^2)	$\frac{1}{N}\sum_{i=1}^{N} \frac{XSQ^2}{YSQ}$
OPRAT	-	ratio of mean OP to mean PP	$\frac{OPOW}{PPOW}$
OPSQR2	-	average of ratio of squared OP to PP	$\frac{1}{N}\sum_{i=1}^{N} \frac{XSQ^2}{YSQ}$
OPSQR	-	ratio of average OP squared to average of PP	$\frac{OPOW^2}{PPOW}$
OPRAV	-	ratio of OP to mean PP	$\frac{OGM}{PGM}$
VARRAT	-	ratio of variance of OP to variance of PP	$\frac{VOPOW}{VPPOW}$

Table 4-1 (Continued)

OPSAV	-	ratio of average of squared OP (M^2) to mean PP (M^2)	$\left(\frac{1}{N}\sum_{i=1}^{N} XSQ^2\right) \Big/ PPOW$
RATSKEW	-	ratio of OP skewness to PP skewness	$\frac{OPSKEW}{PPSKEW}$
RATKUR	-	ratio of OP kurtosis to PP kurtosis	$\frac{OPKUR}{PPKUR}$

4.1.2 A Continuous Piecewise Linear Transformation of the Sensor Data

For this transformation the input data was limited to the information in the first three pulses (i.e., PP(1), OP(1), PP(2), OP(2), PP(3), OP(3)),and thus M=6. The transformation itself was a very general functional non-linear form which mapped the M=6 dimensional space to a single dimension or feature(see figure 4.2). A description of this functional form follows:

Continuous Piecewise Linear Functions

A piecewise linear function is a function for which one can find a partition of the space of independent variables such that the function is linear on every subregion. If the function is continuous piecewise linear, there are no discontinuities in the function at the boundaries between subregions. A continuous piecewise linear function of one variable js shown in Figure 4.3. Figure 4.4 (a), (b), and (c) illustrate continuous piecewise linear functions of two variables. In both cases the continuity constraint requires that the hyperplanes defining the function in any subregion meet at the boundaries of the subregions. Thus, in Figure 4.3 the values of the linear functions on the first and second subregions must be the same at the boundaries between those subregions, i.e., the point a, and the values of the linear functions on the second and third subregions must be the same at the boundary between those subregions, i.e., the point b. In higher dimensions the subregions can become

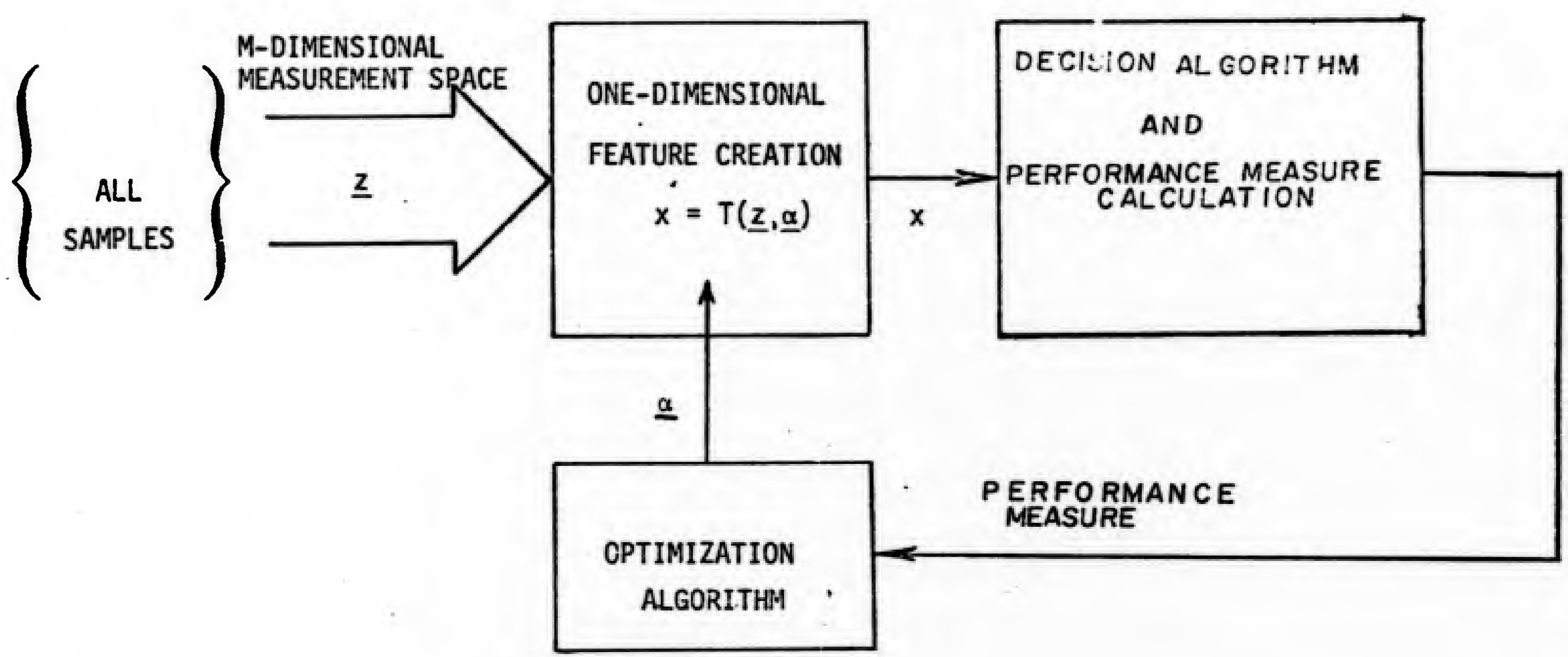


Figure 4.2 - Single feature creation: an objective combination of the sensor data.

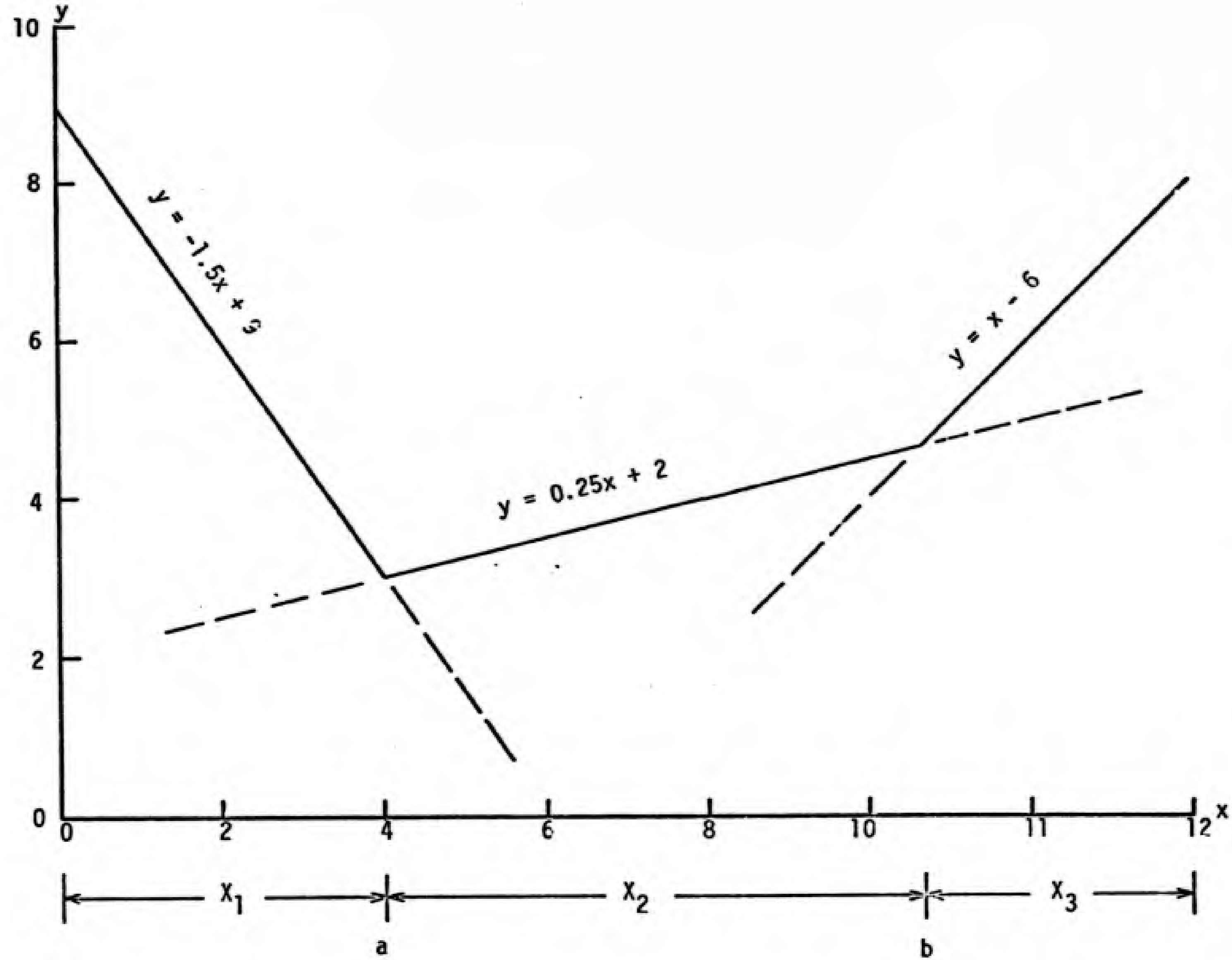


Figure 4.3

A CONTINUOUS PIECEWISE LINEAR FUNCTION OF ONE VARIABLE

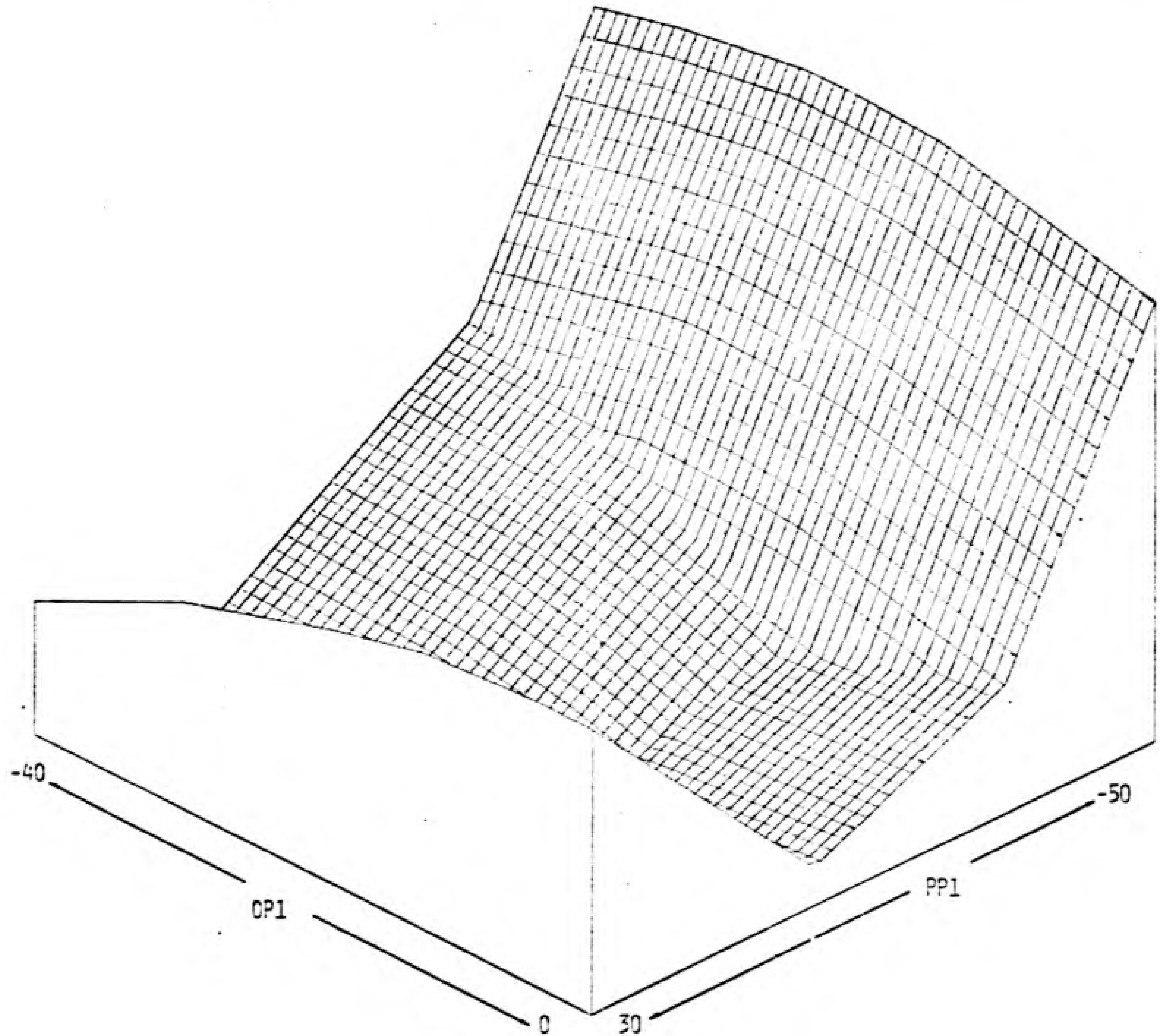


Figure 4.4a Two-Dimensional Piecewise Linear Form

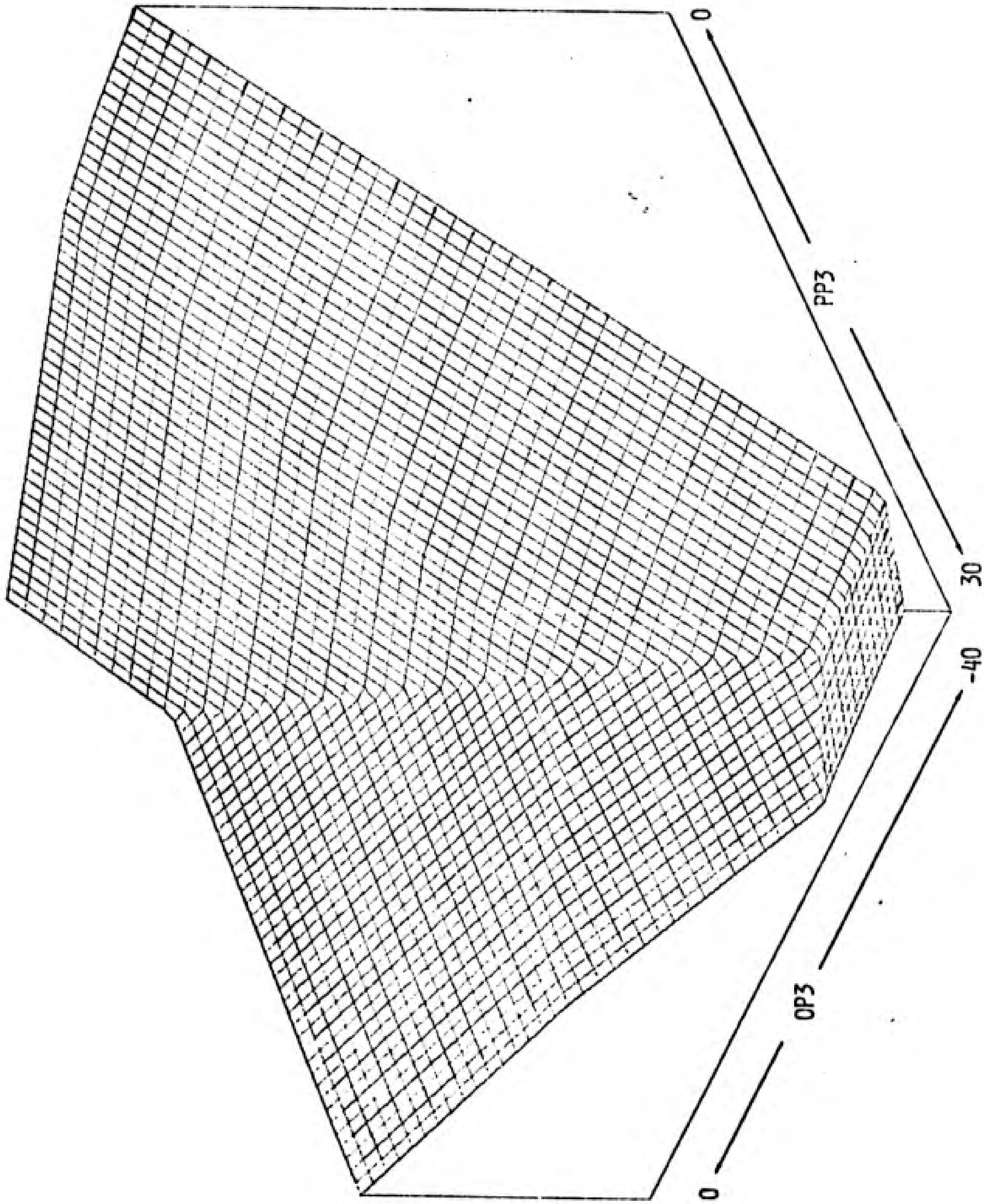


Figure 4.4b

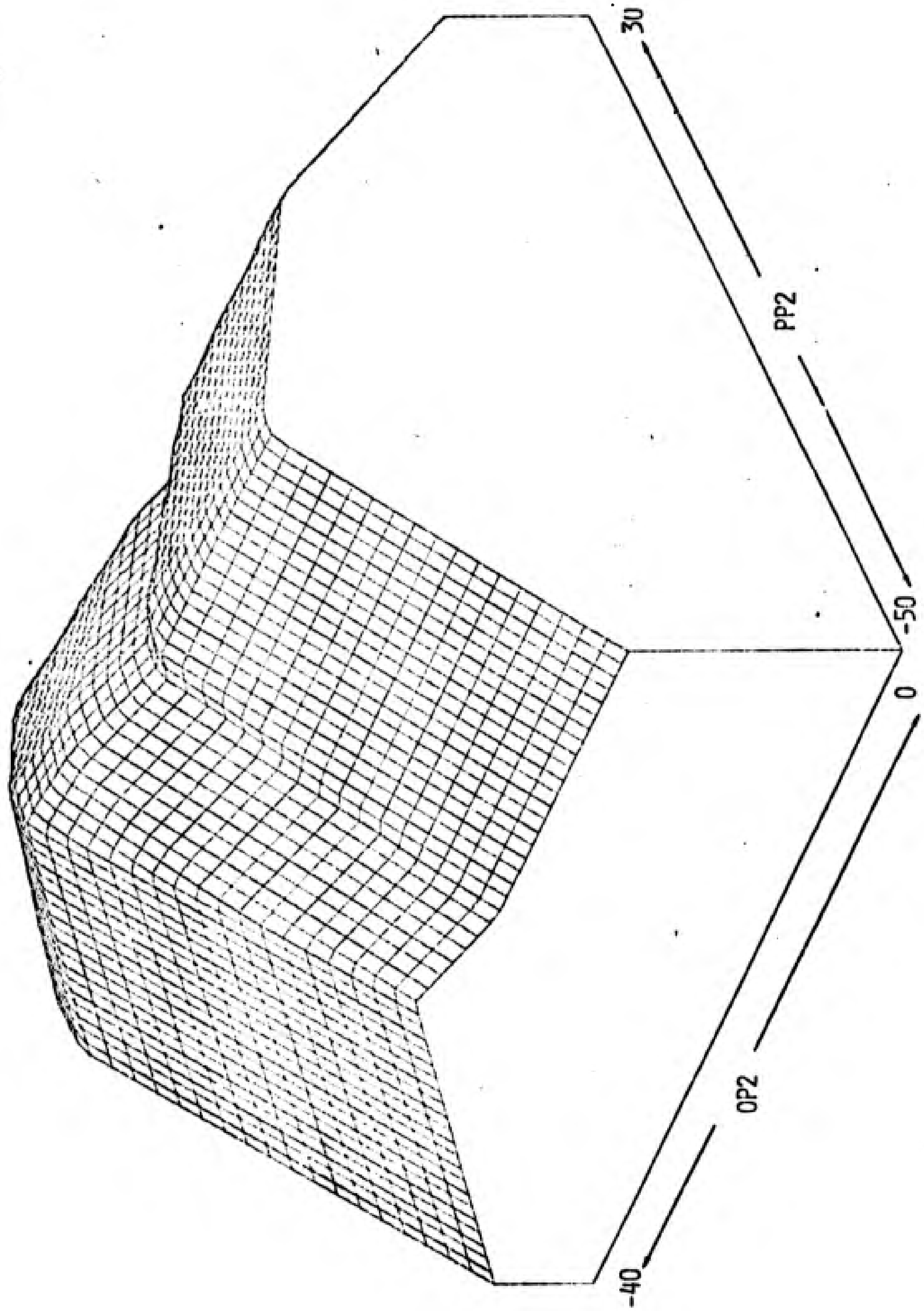


Figure 4.4c

considerably more complex, as indicated in Figure 4.5, and the problem of ensuring continuity is a more difficult technical problem. The general formula for a piecewise linear function is given by

$$F(x_1,\dots,x_n) = \begin{cases} \sum_{j=1}^{n} b_{1j}x_j + b_{1,n+1} & \text{for } \underline{x} \text{ in } X_1 \\ \vdots & \\ \sum_{j=1}^{n} b_{Rj}x_j + b_{R,n+1} & \text{for } \underline{x} \text{ in } X_R \end{cases} \tag{4.1}$$

where $\underline{x} = (x_1,x_2,\dots,x_n)$ and $X_1,X_2,\dots,X_R$ are subregions partitioning the space.

For any given set of subregions $X_1,X_2,\dots,X_R$, one could (with difficulty) find the optimal coefficients b_{ij} with a constraint of continuity at the boundaries. Since the choice of subregions is not obvious, the problem of simultaneously finding the optimal subregions makes the procedure quite difficult. The approach employed in the present work is the specification of the piecewise linear function in an alternate form which insures continuity as the parameters are varied and which defines the boundaries of the subregions implicitly as a function of the parameters defining the linear function on each subregion.

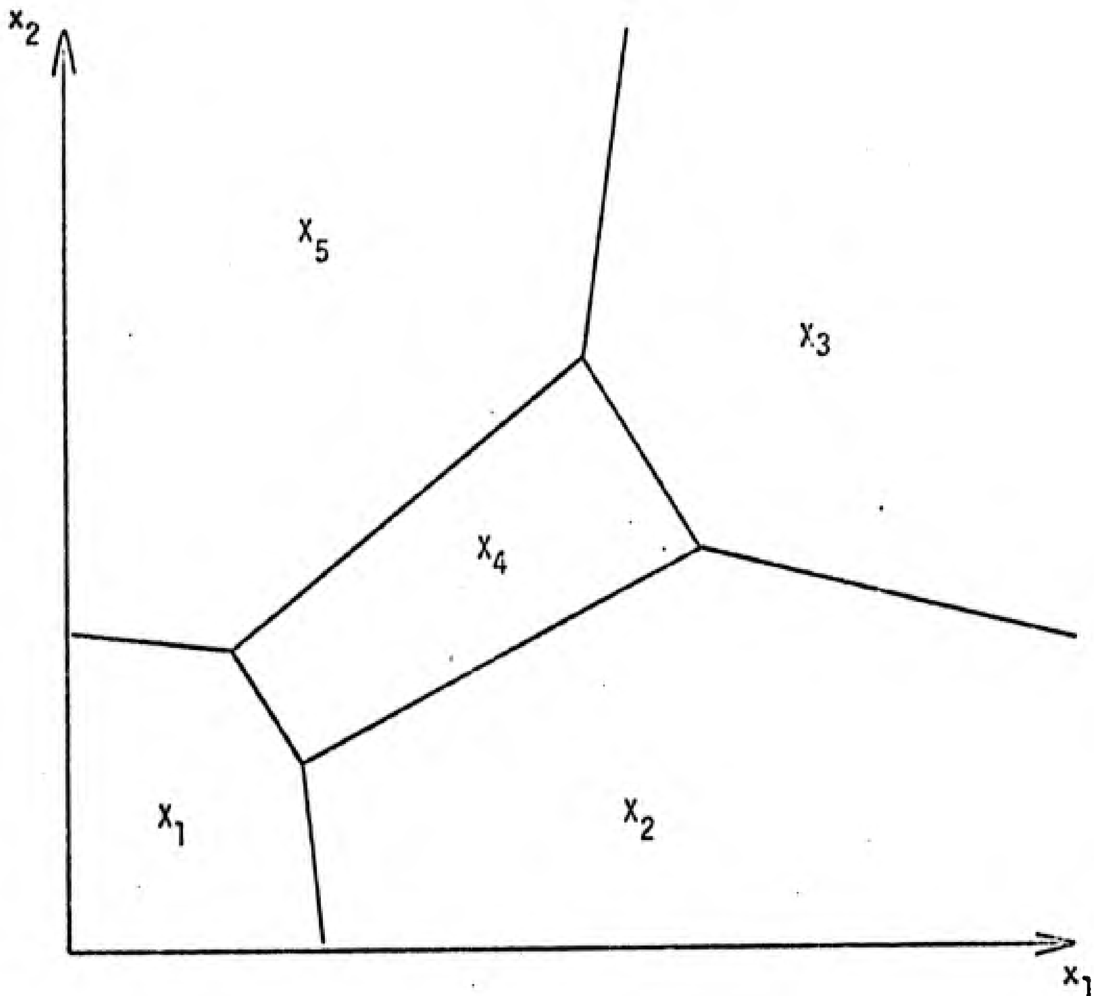


FIGURE 4.5

AN EXAMPLE OF POSSIBLE SUBREGIONS FOR A TWO-VARIABLE CONTINUOUS PIECEWISE LINEAR FUNCTION

Specifically, equation (4.2) defines a continuous convex piecewise linear function:*

$$P(x_1,x_2,\ldots,x_n) = \underset{i=1,2,\ldots,K}{\mathrm{Max}} \left\{ \sum_{j=1}^{n} a_{ij}x_j + a_{i,n+1} \right\} \tag{4.2}$$

Referring to equation (4.1), note that $F(x_1,\ldots,x_n) = P(x_1,\ldots,x_n)$ if $b_{ij} = a_{ij}$ and X_i is the region where the i^{th} hyperplane is largest, i.e.,

$$X_i = \left\{ \underline{x} = (x_1,x_2,\ldots,x_n) \middle| \sum_{j=1}^{n} a_{ij}x_j + a_{i,n+1} \geq \sum_{j=1}^{n} a_{kj}x_j + a_{k,n+1} \text{ for all } k \right\}.$$

Figures 4.3 and 4.4 (c) illustrate convex and non-convex piecewise linear functions respectively. Figure 4.3 illustrates this definition graphically. Note that the value of the function P(x) is obtained quite simply by calculating the values of the three linear functions

$$g_1(x) = -1.5x + 9$$

$$g_2(x) = 0.25x + 2$$

$$g_3(x) = x - 6$$

*A convex function is roughly, one which has the property that all the points on a line connecting two points on the surface of the function take values greater than or equal to the function.

and taking the largest value which results. The subregions are defined implicitly; for example, in Figure 4.3, X_2 is the region where

$$0.25x + 2 \geq x - 6$$

and

$$0.25x + 2 \geq -1.5x + 9$$

A simple extension of the approach will yield non-convex functions:

$$F(x_1,\ldots,x_n) = \sum_{k=1}^{N} w_k P_k(x_1,\ldots,x_n) \quad , \tag{4.3}$$

where

$$P_k(x_1,\ldots,x_n) = \max_{i=1,2,\ldots,K_k} \left\{ \sum_{j=1}^{n} a_{ij}^{(k)} \cdot x_j + a_{i,n+1}^{(k)} \right\} ,$$

i.e., F is a sum of functions of the form (4.2). The function F may be non-convex if the weights w_k differ in sign. Note that F is a "parameterized" function: to fully specify F, we must choose the values $w_1,\ldots,w_N$ and $a_{ij}^{(k)}$ for $k=1,2,\ldots,N$; $i=1,2,\ldots,K_k$; and $j=1,2,\ldots,n$. Some of these parameters are redundant; the total number of free parameters is

$$(n+1)\left(\sum_{k=1}^{N} K_k\right) . \tag{4.4}$$

The parameters b_{ij} in equation (4.1) are related to the parameters w_k and a_{ij} by a linear equation on each subregion.

The procedure used was to test whether a convex function of the form in (4.2) was sufficient to represent the input/output relationship; this would be the case only if the relationship itself were convex or nearly so. If a convex function was insufficient, then a functional approximation of the form (4.3) was employed. This procedure yields the fringe benefit of detecting whether the model input/output relationship is itself essentially convex.

Development of EFFAP3

The particular transformation which was determined by optimizing the parameters of the piecewise linear form was called EFFAP3. Its parameters are listed in figure 4.6. The performance criteria used (described later) was the percent of correctly classified samples.

Figures 4.4abc show the nature of the final EFFAP3 surface, obtained by fixing four of the measurement variables and displaying the variation in the feature value as a function of the two remaining variables. Note that the surface is relatively smooth and, with some computational effort, could be represented by a continuous piecewise linear function with considerable fewer parameters.

INPUT VARIABLES

x_1 = PP(1)	x_2 = OP(1)	x_3 = PP(2)	x_4 = OP(2)	x_5 = PP(3)	x_6 = OP(3)	CONSTANT	
NO. 1 WITH WEIGHT = 4,798E-01 = w_1							
-2,015E-01	4,101E-03	-2,766E-03	4,857E-03	-1,089E-03	8,759E-03	-3,732E+00	$[a_{ij}^{(1)}]$ $i=1 \downarrow i=4$, $j=1 \rightarrow j=7$
-5,526E-02	-4,878E-02	-1,464E-02	-3,270E-02	-5,158E-02	-2,177E-02	-3,257E+00	
-4,053E-02	7,600E-02	-2,209E-02	1,064E-01	8,457E-03	7,835E-02	4,183E+00	
1,160E-01	5,506E-02	4,887E-03	-4,875E-03	2,535E-02	-2,083E-02	8,370E-01	
NO. 2 WITH WEIGHT = -7,826E-02 = w_2							
-5,958E-03	-2,552E-01	-4,671E-03	4,306E-03	-1,722E-03	3,611E-03	-6,612E+00	$[a_{ij}^{(2)}]$
6,403E-03	-4,161E-02	3,769E-04	5,311E-03	2,143E-04	6,694E-03	-5,220E-01	
4,352E-03	4,119E-02	2,549E-03	-1,258E-03	5,952E-03	-3,176E-03	8,689E-01	
-4,954E-03	2,572E-01	-4,584E-03	-8,143E-03	-7,159E-03	-2,894E-03	3,337E+00	
NO. 3 WITH WEIGHT = -1,914E-01 = w_3							
-1,274E-02	-7,493E-03	-1,889E-01	-4,880E-03	-9,845E-03	-1,032E-02	-4,689E+00	$[a_{ij}^{(3)}]$
8,492E-03	7,102E-03	-2,262E-02	6,037E-03	1,086E-02	7,585E-04	1,442E-01	
1,499E-02	1,254E-02	4,357E-02	1,643E-02	9,796E-03	1,949E-02	1,991E+00	
-1,141E-02	-1,297E-02	1,570E-01	-2,074E-02	-1,532E-02	-7,640E-03	-6,222E-01	
NO. 4 WITH WEIGHT = -2,165E-01 = w_4							
-1,821E-02	8,627E-03	-1,569E-02	-2,516E-01	-1,509E-02	5,065E-03	-6,784E+00	$[a_{ij}^{(4)}]$
2,993E-03	2,402E-02	1,433E-02	-3,746E-02	1,337E-04	1,906E-02	2,914E-01	
1,646E-02	-1,193E-02	3,059E-03	4,552E-02	7,487E-03	-2,425E-02	4,954E-01	
-2,979E-03	-2,127E-02	-1,898E-02	2,393E-01	-6,349E-04	5,335E-03	2,716E+00	

CONSTANT TERM = 1,280E+00 = w_0

$$\mathrm{EFFAP3}(\overline{x}) = w_0 + \sum_{k=1}^{4} w_k P_k(\overline{x})$$

$$P_k(x) = \underset{i=1,4}{\mathrm{Max}} \left\{ \sum_{j=1}^{6} a_{ij}^{(k)} x_j + a_{i7} \right\}$$

FIGURE 4.6 Parameters of the continuous piecewise linear transformation: EFFAP 3

4.1.3 A Parameterized Linear Combination of Heuristically Defined Features

In this case a quite simple functional form was selected for illustration: a parameterized linear form which maps five heuristically defined features into a single feature via the transformation (see figure 4.7)

$$x = T(\underline{y},\underline{\alpha}) = \sum_{i=1}^{5} \alpha_i y_i + \alpha_6 \quad . \tag{4.5}$$

In particular the form used was

$$\text{REGR} = \alpha_1 \cdot \text{MVAROP} + \alpha_2 \cdot \text{VANGL} + \alpha_3 \cdot \text{PGM} + \alpha_4 \cdot \text{TPOW} + \alpha_5 \cdot \text{PP3M} + \alpha_6 \tag{4.6}$$

The five heuristic features have been defined previously in Table 4.1. The method of selecting these particular five variables from all those listed in the table involved the use of a stepwise regression procedure which is briefly described below.

Stepwise regression is a powerful variation of multiple regression which provides a means of selecting features (independent variables) which will provide close to the best prediction possible (given the assumed functional form of the relationship, in this case linear) of a dependent variable. In this case, the dependent variable employed was the numerical value assigned to the known class label (target object = 1, non-target object = 2) associated with each sample signature in the

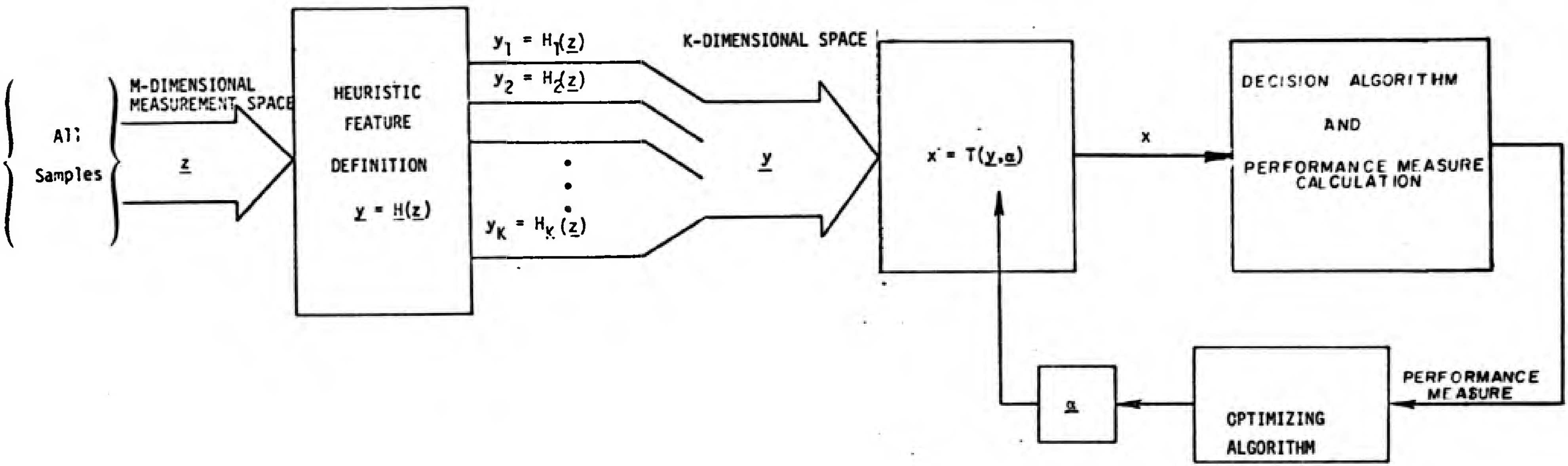


Figure 4.7 - Single feature creation: an objective combination of heuristic features

design set. The stepwise regression algorithm proceeds by first selecting the single feature (independent variable) which best explains the dependent variable. Next additional features are selected sequentially based on two criteria:

1) The value of the normalized regression coefficient (a_i) that the prospective feature would have if brought into the functional form as an additional term.

2) A measure, tolerance (T), of how nearly the new feature is expressable as a linear combination of the previously selected features. (A large value of T implies that a truly independent dimension is being added).

The prospective feature with the largest value of $a_i^2 T$ is selected as the next feature to be added to the set. Execution of the stepwise regression technique resulted in the selection of features listed in equation 4.6; they were selected in the order listed.

The end result of a subsequent optimization to determine the best parameters of the functional form was:

$$\mathrm{REGR} = -.12 \times \mathrm{MVAROP} + 2.26 \times \mathrm{VANGL} - 0.4 \times \mathrm{PGM} + $$

$$+ \; .029 \times \mathrm{TPOW} + .00002 \times \mathrm{PP3M} + 1.466$$

4.2 Specification of a Decision Algorithm and Performance Measure

In order to evaluate the "goodness" of a particular dimensionality reducing transformation which generates a set of P features, it is necessary to define a scalar valued performance measure or criterion. The performance measure used should reflect the overall objectives and constraints of the discrimination system (see section 2). For this study the criterion was a measure of how a Bayes classifier would perform given the transformed samples as inputs. The use of the Bayes decision algorithm involved estimating the class conditional probability densities, $p(\underline{x}|i)$, from the samples of the target and non-target classes in the design set. Given these estimates of the densities the decision rule was

$$DC(\underline{x}) = \begin{cases} 1 \text{ (target class) if } & \dfrac{p(\underline{x}|1)}{p(\underline{x}|2)} \geq T \\ 2 \text{ (non-target class)} & \text{otherwise} \end{cases} \tag{4.7}$$

where the threshold value, T, depends on the a priori probabilities of classes 1 and 2 $(p(1),p(2))$ and the cost (weights) associated with making an incorrect classification (h_{FA} = cost of a false alarm, h_{MD} = cost of a missed detection.) The threshold can be expressed as:

$$T = \frac{p(2)h_{FA}}{p(1)h_{MD}} \tag{4.8}$$

If these costs and a priori probabilities can be specified beforehand, the performance can be determined for the particular resulting T. We will discuss below methods for evaluating performance when these terms are unspecified.

Two methods of estimating the densities were used. Generally the density estimates were derived by assuming the samples to be normally distributed and using maximum likelihood estimates of the mean vector and covariance matrix. Where there was strong evidence that the normality assumption might lead to large errors, a non parametric (Parzen) estimate of the densities was used (see [1, chapter VI]). This technique of density estimation is highly computational and was only used sparingly in the feature ranking procedure described in the next section. Recourse to the Parzen estimator technique was made when there was a significant difference between the actual error rate (experienced when using the decision rule generated by the normal assumption on the samples of design set) and the estimated error rate (based on an exact method [6] of determining the overlap of two multivariate normal densities.) Figures 4.8 and 4.9 indicate for two features how the estimates obtained using the normal assumption compare with estimates using Parzen estimators. Note that differences in the density estimates themselves do not always imply significant differences in the decision boundaries. Whatever method was used in estimating the densities, however, the densities were then used to create decision rules which were applied to the samples to obtain performance measures. In particular, error rates might improve in going from the normal assumption to Parzen estimators; they will not get worse.

Given estimates of the class conditional probability densities, equation 4.7 can be used to determine a "decided" class label -- decided by the algorithm. Of course, each sample of the design set has a known

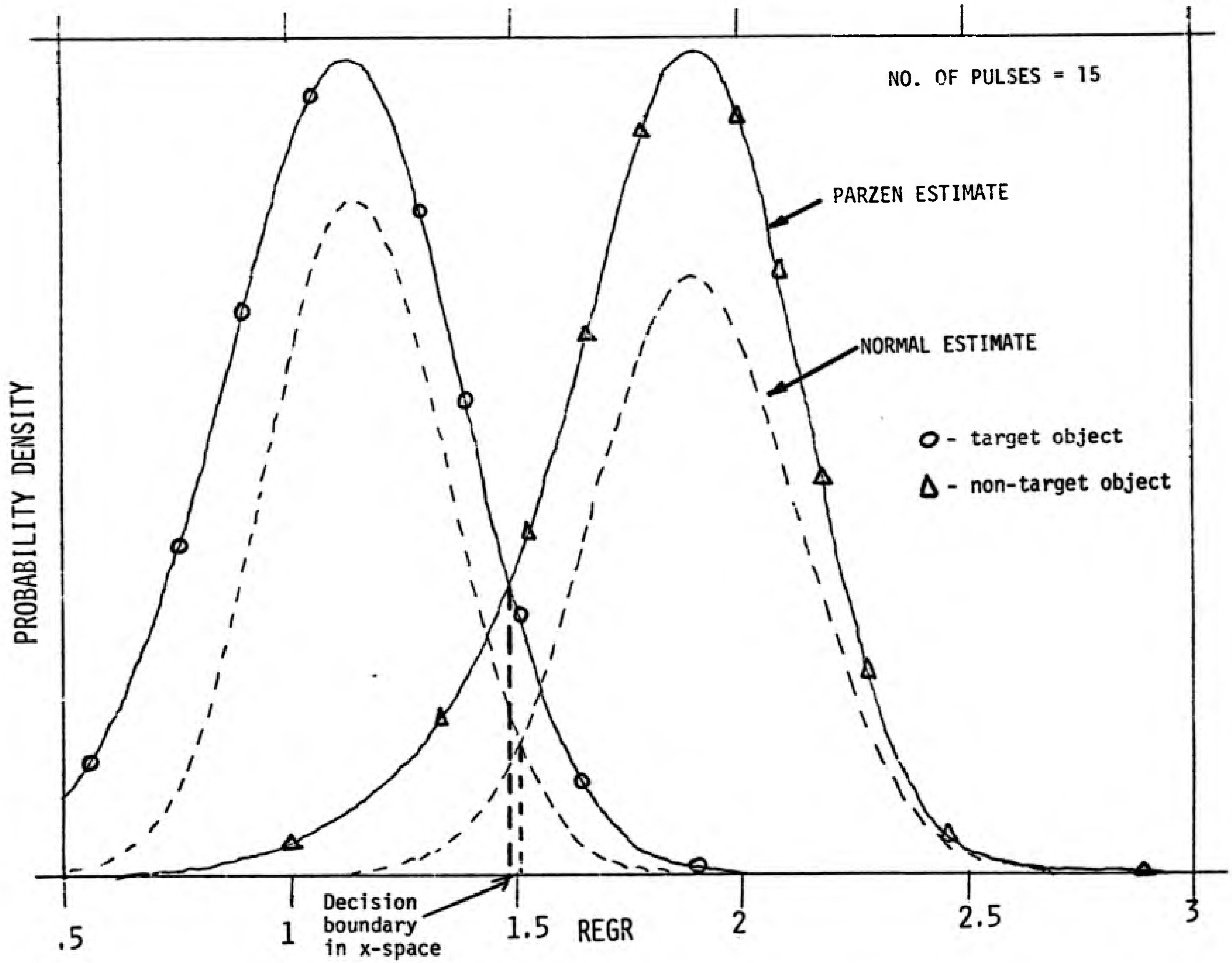


Figure 4.8: Two Density Estimates (Estimates are scaled differently)

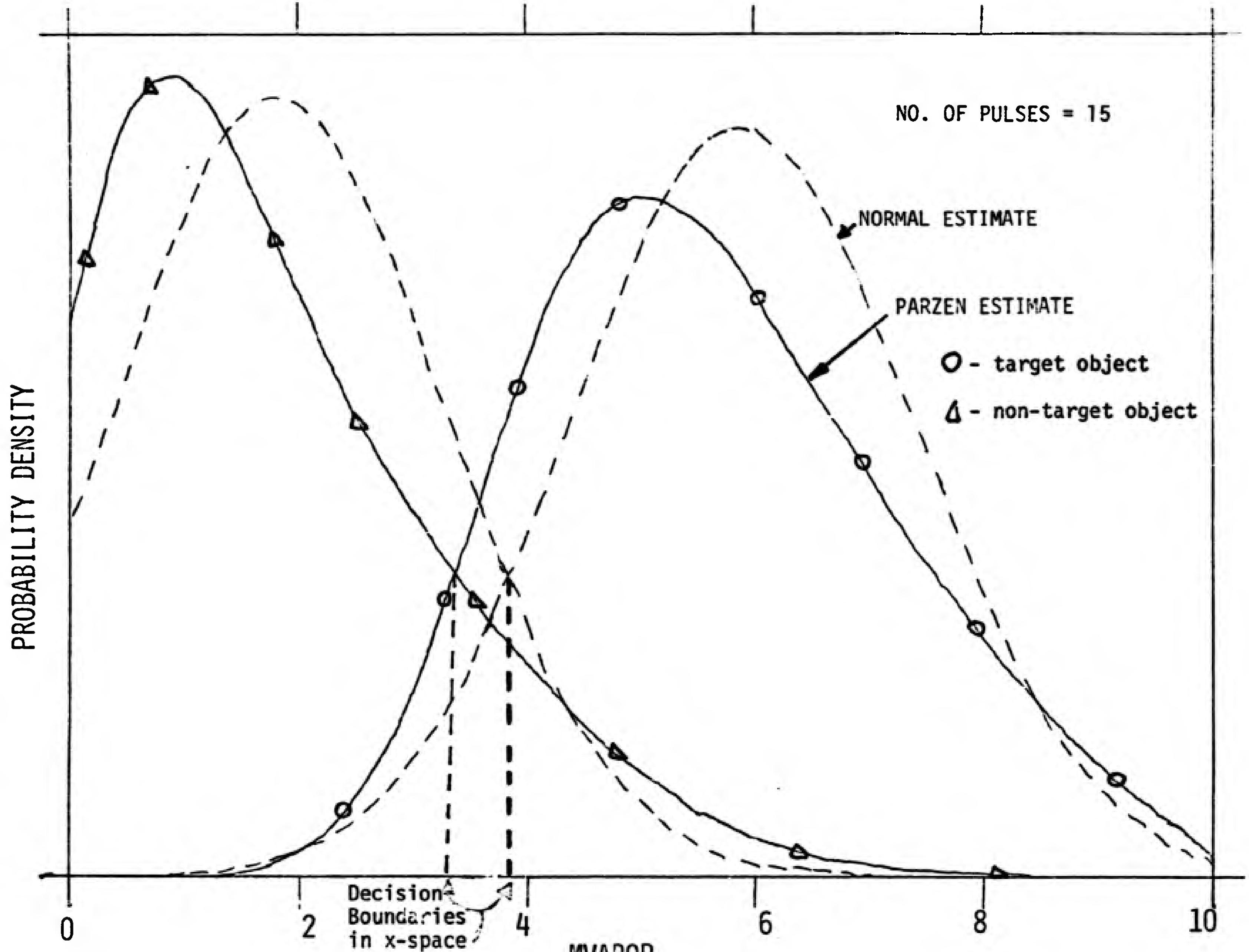


Figure 4.9: Two Further Density Estimates

CLASS INDICATED BY DECISION RULE:

DECIDED CLASS

RUE
LASS

		(Decided Target) 1	(Decided Non-target) 2
(True Target)	1	A_{11}	A_{12} (missed detection)
(True Non-Target)	2	A_{21} (false alarm)	A_{22}

ENTRIES ARE NUMBER OF SAMPLES IN GIVEN CELL.

FIGURE 4.10: A CONFUSION MATRIX

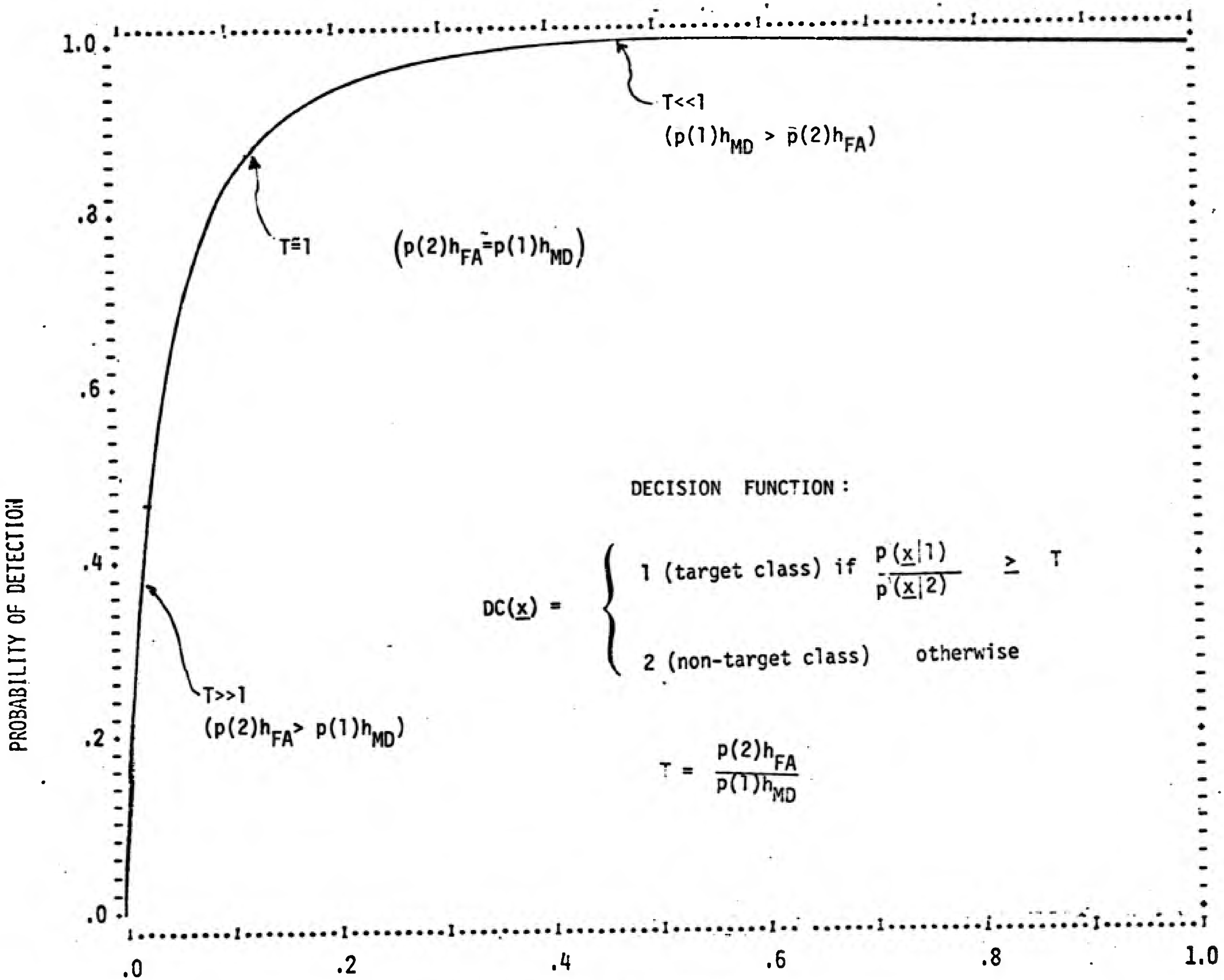


FIGURE 4.11 A PD/PFA CURVE

true class label associated with it. A confusion matrix (see figure 4.10) summarizes the distribution of the pairs of labels (true class, decided class) over the entire design (or test) set of samples.

Below we define several measures of performance from the data contained in the confusion matrix:

$$\text{Error rate (\%)} = 50 \left[\frac{a_{12}}{a_{11} + a_{12}} + \frac{a_{21}}{a_{21} + a_{22}} \right]$$

$$\text{Percent Correctly Classified} = 100 - \text{Error Rate}$$

$$\text{Probability of False Alarm (PFA)} = \frac{a_{21}}{a_{21} + a_{22}}$$

$$\text{Probability of Missed Detection (PMD)} = \frac{a_{12}}{a_{11} + a_{12}}$$

Probability of Detection (PD) = 1 - Probability of Missed Detection

(Notice that all of these measures of performance are defined in terms of the action of the derived decision rule on a set of labeled samples.)

Changing the threshold, T, in equation 4.7 will change the confusion matrix and hence the above measures; the performance of the decision rule is thus to some extent a function of T. As T is changed from large values ($T \gg 1$) to small values ($T \ll 1$) the resulting values of probability of detection (hereafter, PD) and probability of false alarm (hereafter, PFA) change in a characteristic fashion (see figure 4.11) and trace out a "PD/PFA" curve.

A detailed examination of the quality of a set of features might best be carried out in terms of the entire PD/PFA curve; for this reason PD/PFA curves of several of the more interesting sets of features were developed and can be found in the later sections. However, a simple scalar valued performance measure is a necessity for the efficient design optimization of the parameterized transformations mentioned in section 4.1 as well as for the feature ranking procedure involved in the determination of the best selection of P features from a set of K features (see figure 4.1 and section 5). Two such simple measures are:

1) Percent Correctly Classified for T=1

2) Probability of Detection for a Fixed Probability of False Alarm

Figure 4.12 illustrates [1] how the PD/PFA curves would change as an optimization of transformation proceeds. Note that for the purposes of the optimization, employing the probability of detection for a fixed probability of false alarm (a Neyman-Pearson type formulation) as the performance measure would yield similar results to those achieved by employing the percent correctly

---

[1] Note, this figure is the result of an actual optimization of a transformation of a portion of the design set. This was part of a preliminary analysis and does not correspond to any feature sets reported herein.

classified for a fixed T as the performance measure. Figure 4.13[1] illustrates another way to visualize what is occurring as a parameterized transformation is optimized with respect to a performance measure. In this figure we indicate that as the optimization proceeds the two class conditional densities change shape so as to minmize their overlap (maximize their separation).

For the purposes of this study the performance measure employed in all optimization procedures was the percent correctly classified for T=1. Specifically, this performance measure was used in the derivation of REGR and EFFAP 3. Figures 4.14 and 4.15 show the PD/PFA curves for the final design of the two transformations EFFAP 3 and REGR. Also included in the figures are the PD/PFA curves for the samples represented in the higher dimensional space which served as input to the designed transformation: in the case of EFFAP 3 the input space was 6 dimensional consisting of the OP and PP measurements of the first three pulses; while in the case of REGR, the input space was a 5 dimensional vector of heuristic features, themselves transformation of a 30 dimensional input space consisting of the OP and PP measurements from 15 pulses.

[1]This figure is again for illustrative purposes only.

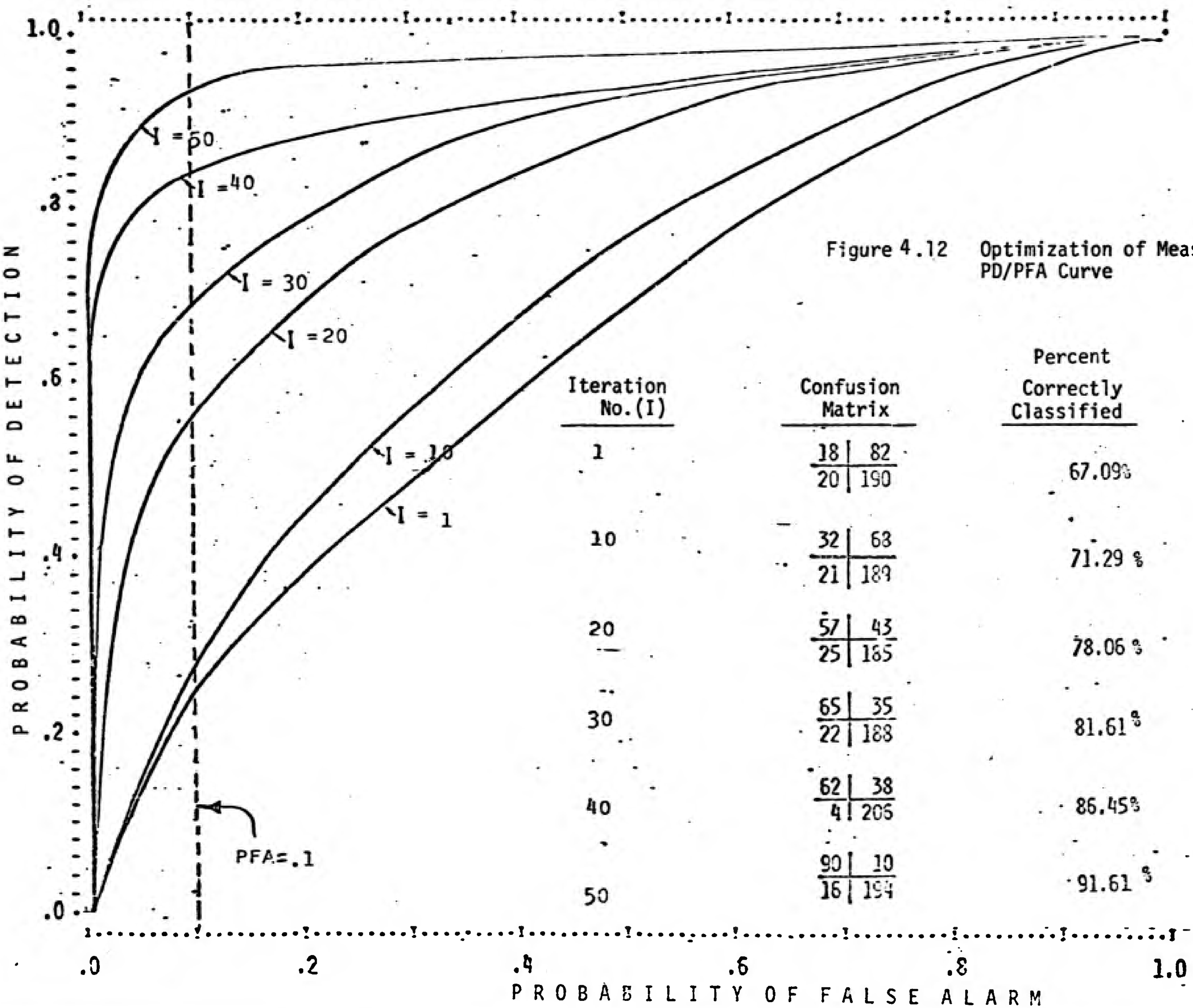


Figure 4.12 Optimization of Measures of a PD/PFA Curve

Iteration No.(I)	Confusion Matrix	Percent Correctly Classified
1	18 | 82 / 20 | 190	67.09%
10	32 | 68 / 21 | 189	71.29 %
20	57 | 43 / 25 | 165	78.06 %
30	65 | 35 / 22 | 188	81.61 %
40	62 | 38 / 4 | 206	86.45%
50	90 | 10 / 16 | 194	91.61 %

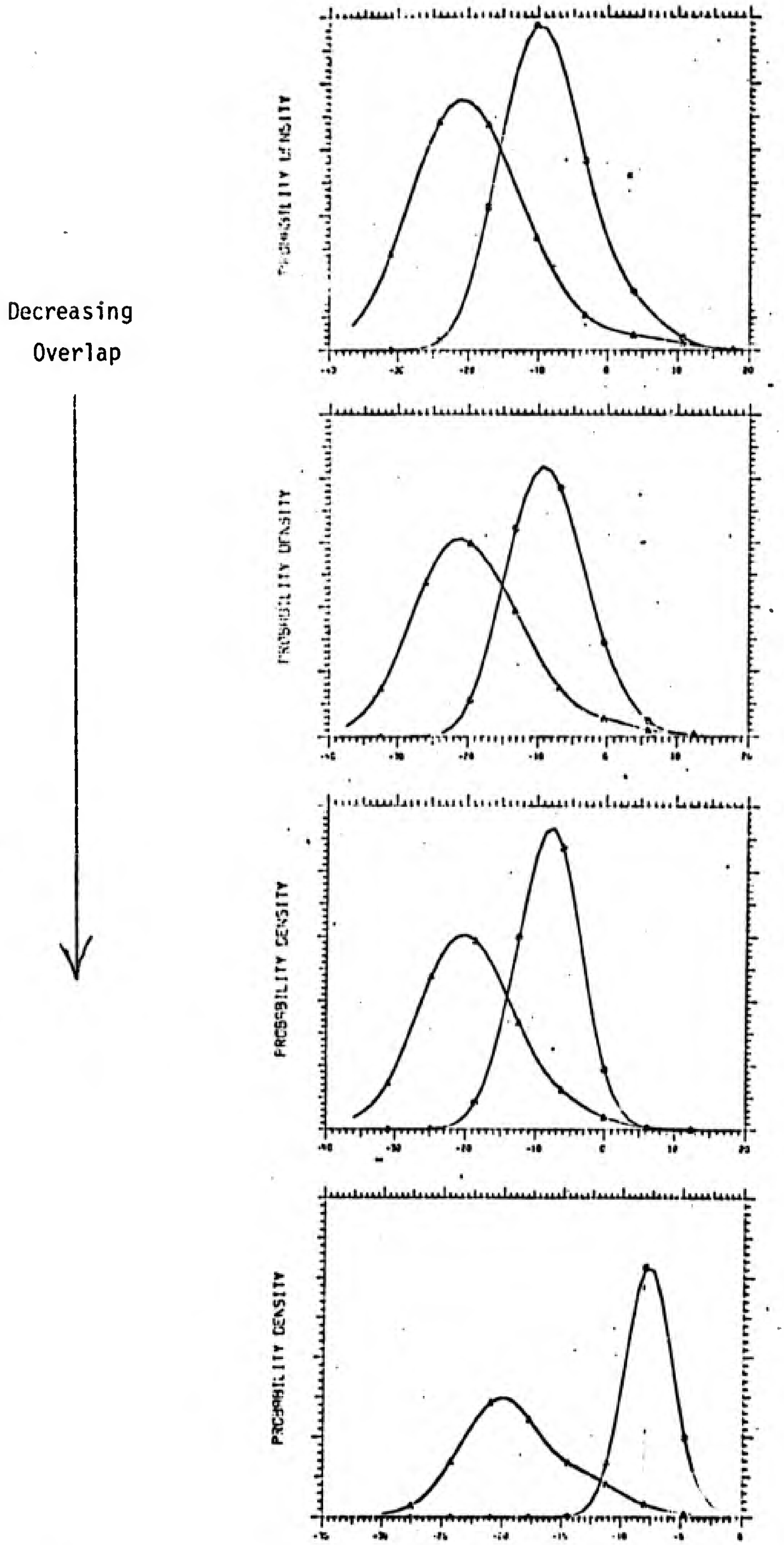
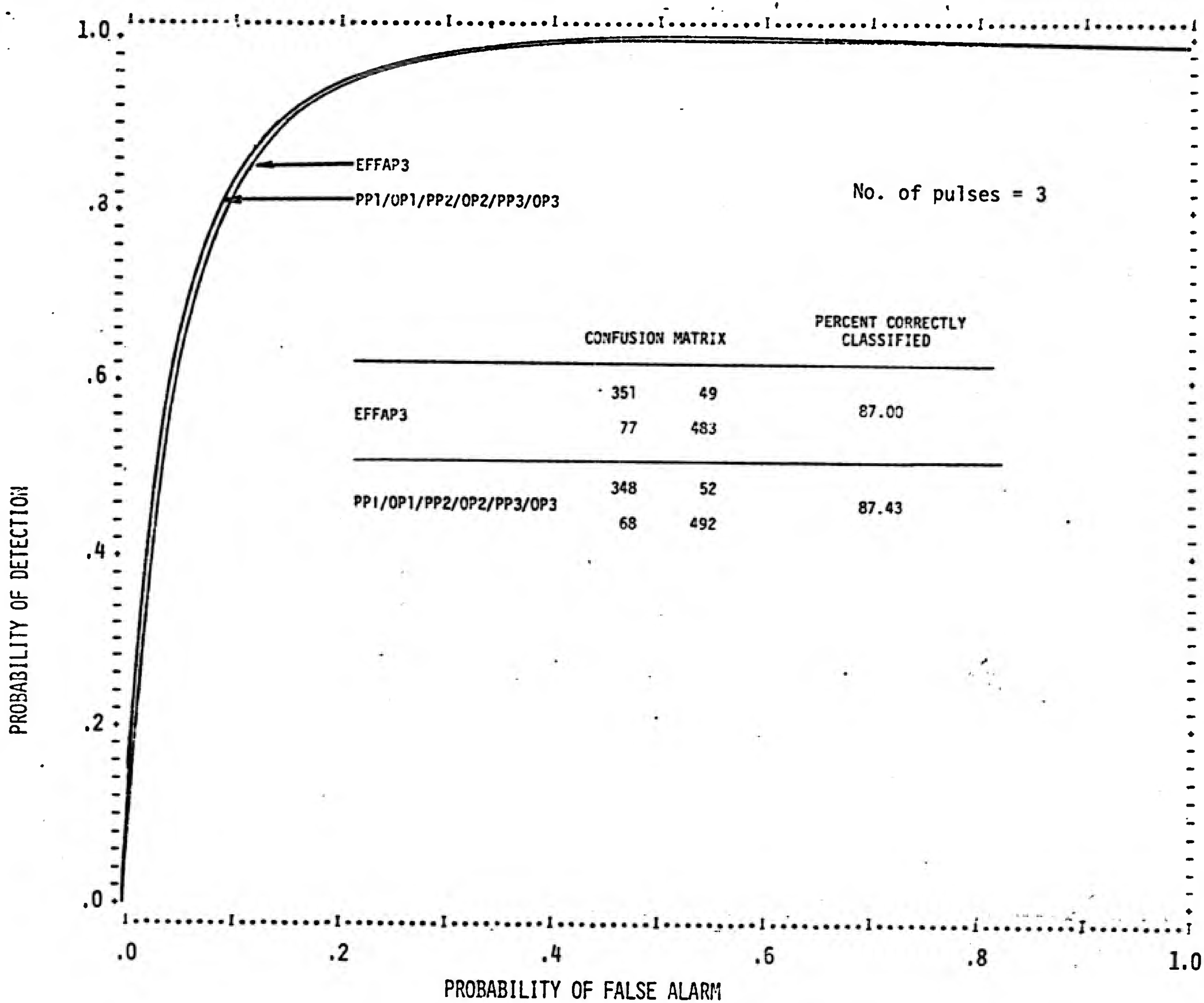


Fig. 4.13 Illustration of the effect of minimizing a measure of the overlap of two-class conditional probability densities.



	CONFUSION MATRIX		PERCENT CORRECTLY CLASSIFIED
EFFAP3	351	49	87.00
	77	483	
PP1/OP1/PP2/OP2/PP3/OP3	348	52	87.43
	68	492	

Figure 4-14. PD/PFA curve for EFFAP3

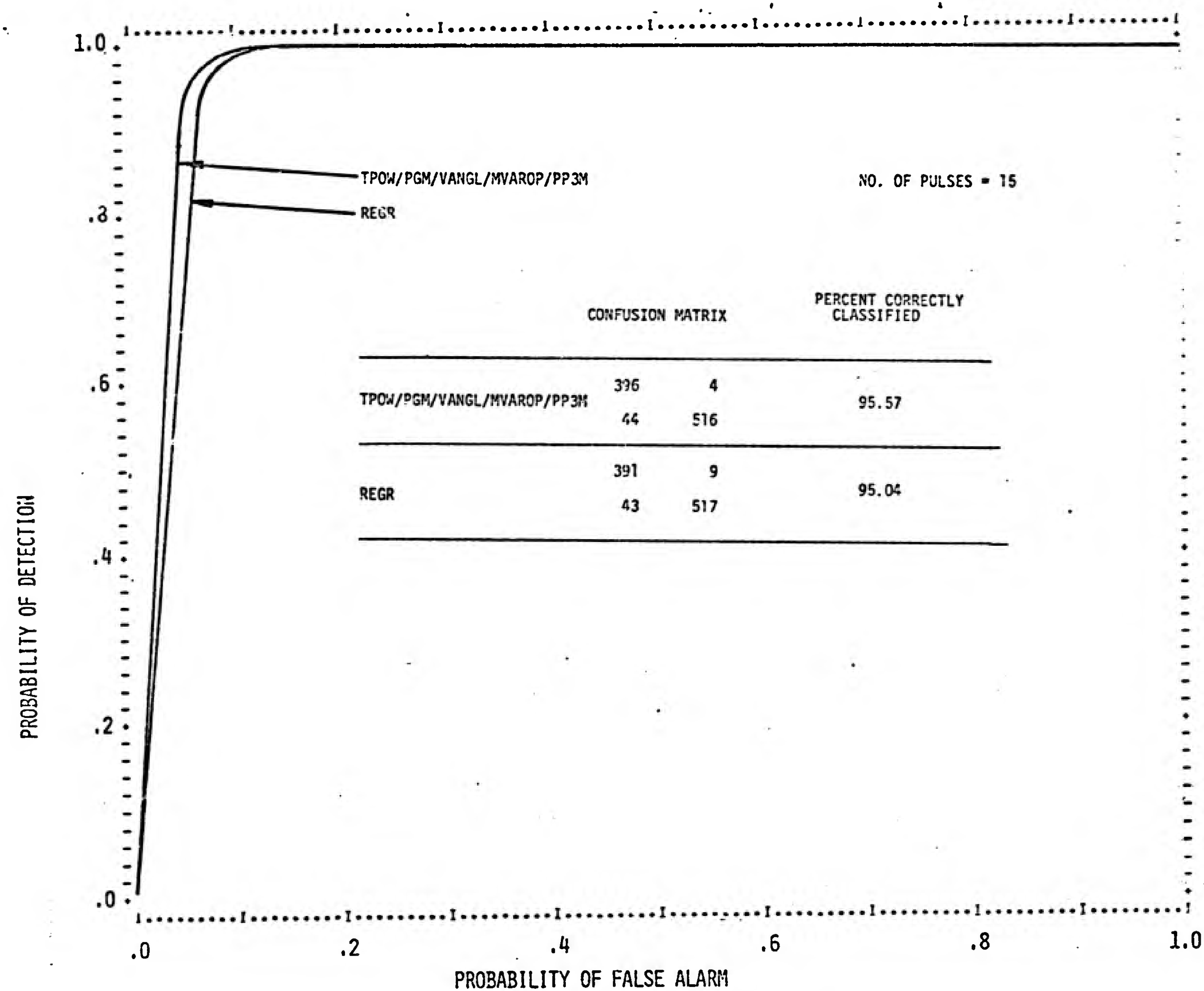


Figure 4.15 - PD/PFA curve for REGR

4.3 Specification of an Optimization Algorithm

The relationship between the parameters of the functional forms used in the design of the parameterized transformations which generated the EFFAP3 and REGR features and the performance measure selected, percent of samples correctly classified, was quite complex. For this reason an efficient random search optimization algorithm, structured random search, (see [1, chapter 3]) was selected to determine the optimal design of the parameters of the transformations.

4.4 Relationship to a More Extensive Approach

As described in earlier sections several specializations of the general approach (described in figure 3.2) to the optional design of a dimensionality reducing transformation were employed in this study. In particular the REGR and EFFAP3 parameterized transformations could be viewed as transformations from an M-dimensional measurement space to a one-dimensional feature space (see figures 4.2 and 4.7). Subsequently these single features were considered as candidates, along with a set of heuristically defined features, for selections via an optimization of the selection matrix (see figure 4.1) resulting in a P-dimensional feature space. A more general approach (and one that could yield improved performance measures) would involve optimization of a general nonlinear transformation <u>in the context of</u> the other features to be included in the final P-dimensional feature space (see figure 4.16).

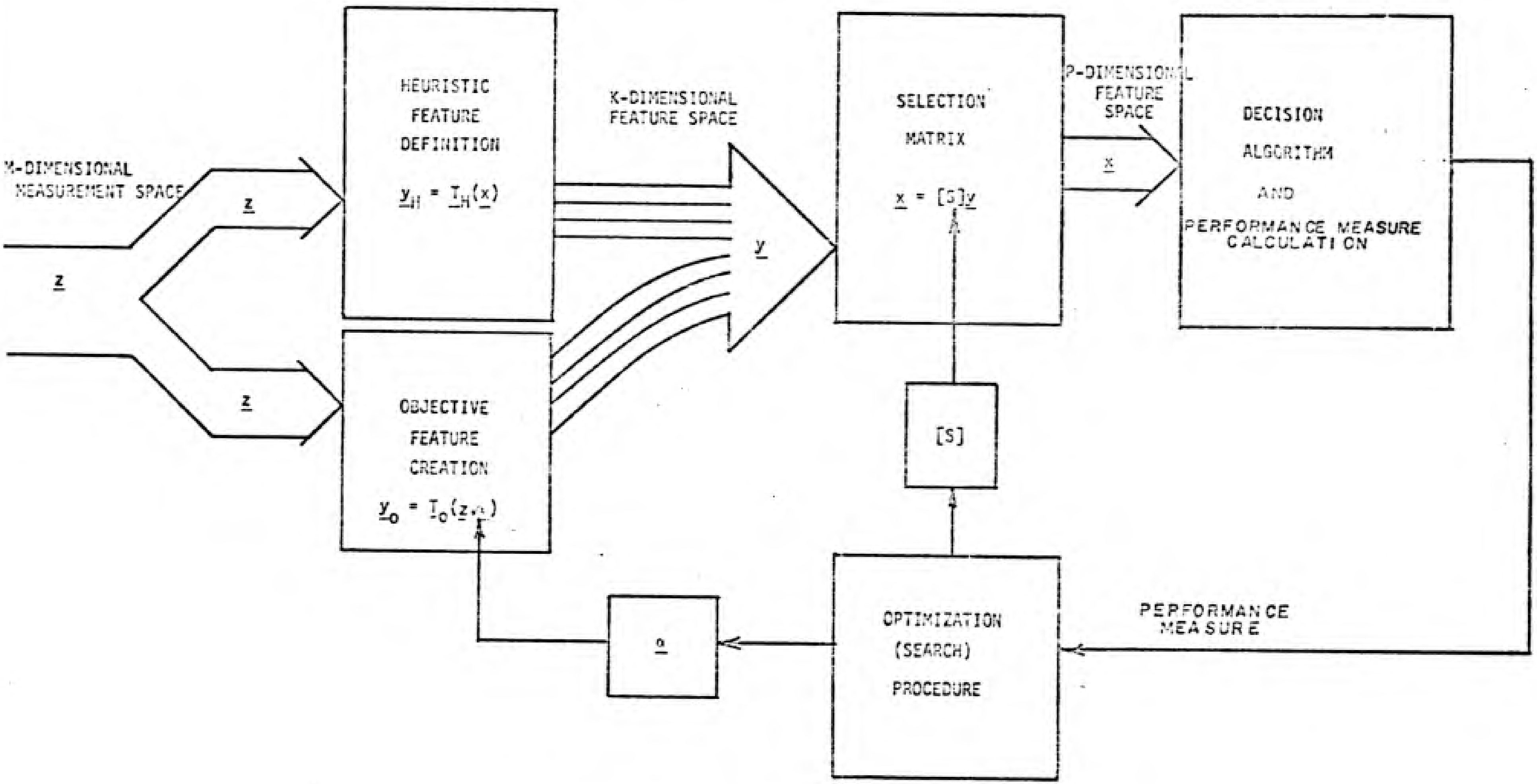


Figure 4.16 A general feature creation/selection procedure.

5.0 RANKING AND SELECTION OF FEATURES

This section describes the procedure employed in this study for systematically selecting a P-dimensional subset (P = 1,2,3,4) of the K-dimensional set of heuristic and objective features (K = 48) defined in Section 4.1 (see also figure 4.1). There are $\binom{K}{P}$ ways to select P features from a set of K features, thus, for example, with K = 48 and P = 4 there would be approximately 195,000 sets of features whose performance must be measured and ranked if an exhaustive search procedure were to be employed. The procedure followed, while not an exhaustive search, tends to lead to a good selection of features in a cost-effective manner.

We begin by ranking the features in terms of individual performance, then use those results to postulate and rank pairs, then triples and finally quadruples. The performance of the best quadruple is a good estimate of the discrimination performance achievable for the target and non-target classes analyzed. The performance measure employed in the ranking was the percent of correctly classified design set samples. The ranking is limited at this point to the performance of heuristic and objective features which are defined on a measurement space of 15 pulses (M = 30), with the exception of EFFAP3 which uses only 3 pulses of sensor data as input.

Table 5-1 indicates the ranking of 48 features;[1] Table 5-2 lists the 14 features used as candidates for subsequent development of combinations features. Note that these are not the 14 highest ranked single features; features were included which gave evidence of containing information valuable in a multivariate context (as for example during the stepwise regression analysis which led to the selection of the input features for the design of the REGR feature (Section 4.1.2)). We did not consider any of the features not in Table 5-2 in further ranking. Figures 5-1 (a-n) show the overlap of the probability densities and the nature of the distribution (estimated by Parzen estimators) for many of the more interesting features. Figure 5-2 shows confusion matrices and PD/PFA curves for five single features.

In obtaining pairs, one would like to avoid testing all two-out-of-fourteen (91) pairs; many can be avoided by elimination of the weaker of two highly correlated features from the analysis. Table 5-3 is the correlation matrix for the features. (Correlation of features with the class label is also included.) For example, we note that PP3M and PP4M have a correlation of -0.98, indicating that testing them as a pair is pointless.

The ranking of ten pairs is indicated in Table 5-4. In Figure 5-5 abc, scatter plots of a subset of the samples illustrate the bivariate distribution of the several pairs of features. Performance summaries for four pairs are provided in Figure 5-6.

---

[1]EFFAP6 and EFFAP15 were derived similarly to EFFAP3. Each used 3 pulses (6 input measurements) of information. The difference between them lies in the fact that while EFFAP3 used the first three pulses as inputs, EFFAP6 and EFFAP15 used 3 pulses which were more widely separated (pulses 1, 4, and 6 for EFFAP6 and pulses 1, 8, and 15 for EFFAP15). The EFFAP6 and EFFAP15 features were not included in the set of 14 investigated more fully because they were roughly equivalent to EFFAP3 and because the study was focused on features based on contiguous pulses.

FEATURE	PERCENT CORRECT CLASSIFICATION	RANK
REGR	95.04	1
PP4M	92.39	2
PP3M	91.81	3
MVAROP	91.55	4
PGM	88.45	5
EFFAP15	88.39	6
EFFAP3	87.00	7
TGM	86.14	8
EFFAP6	85.89	9
XANGL	81.79	10
FLOPSK	81.59	11
RVOGM	80.99	12
OPRAV	80.82	13
PGM/OGM	79.41	14
FLPPKR	78.45	15
OPKUR	78.33	16
MDOP	78.16	17
OPSKEW	77.94	18
VOGM	77.60	19
PPKUR	76.94	20
PPSKEW	76.82	21
OPRAT	76.55	22
XMPA	75.41	23
OPRAT2	75.31	24
MVARPP	72.87	25
RATVOVO	72.29	26
OGM	71.50	27
OP3M	71.32	28
OP4M	70.33	29
RVPGM	69.56	30
VARRAT	67.54	31
RATVDP	66.85	32
VTGM	64.19	33
OPSQR	61.47	34
MDPP	60.32	35
VPGM	59.80	36
VANGL	59.52	37
OPSQR2	57.34	38
RATVPVP	54.64	39
RATSKEW	53.24	40
OPSAV	50.24	41
VOPOW	49.82	42
OPOW	49.32	43
PPOW	48.96	44
RATKUR	48.95	45
TPOW	47.58	46
VTPOW	44.16	47
VPPOW	44.07	48

Table 5-1: Single Feature Ranking (15 pulses)

Feature	Percent Correct Classification	Rank
REGR	95.04	1
PP4M	92.39	2
PP3M	91.81	3
MVAROP	91.55	4
PGM	88.45	5
EFFAP3	87.00	7
TGM	86.14	8
XANGL	81.79	10
FLOPSK	81.59	11
OPRAV	80.82	13
FLPPKR	78.45	15
MDOP	78.16	17
VANGL	59.52	37
TPOW	47.58	46

Single Features Ranked at 15 Pulses (N)

Table 5.2: Selected 14 Single Features (15 pulses)

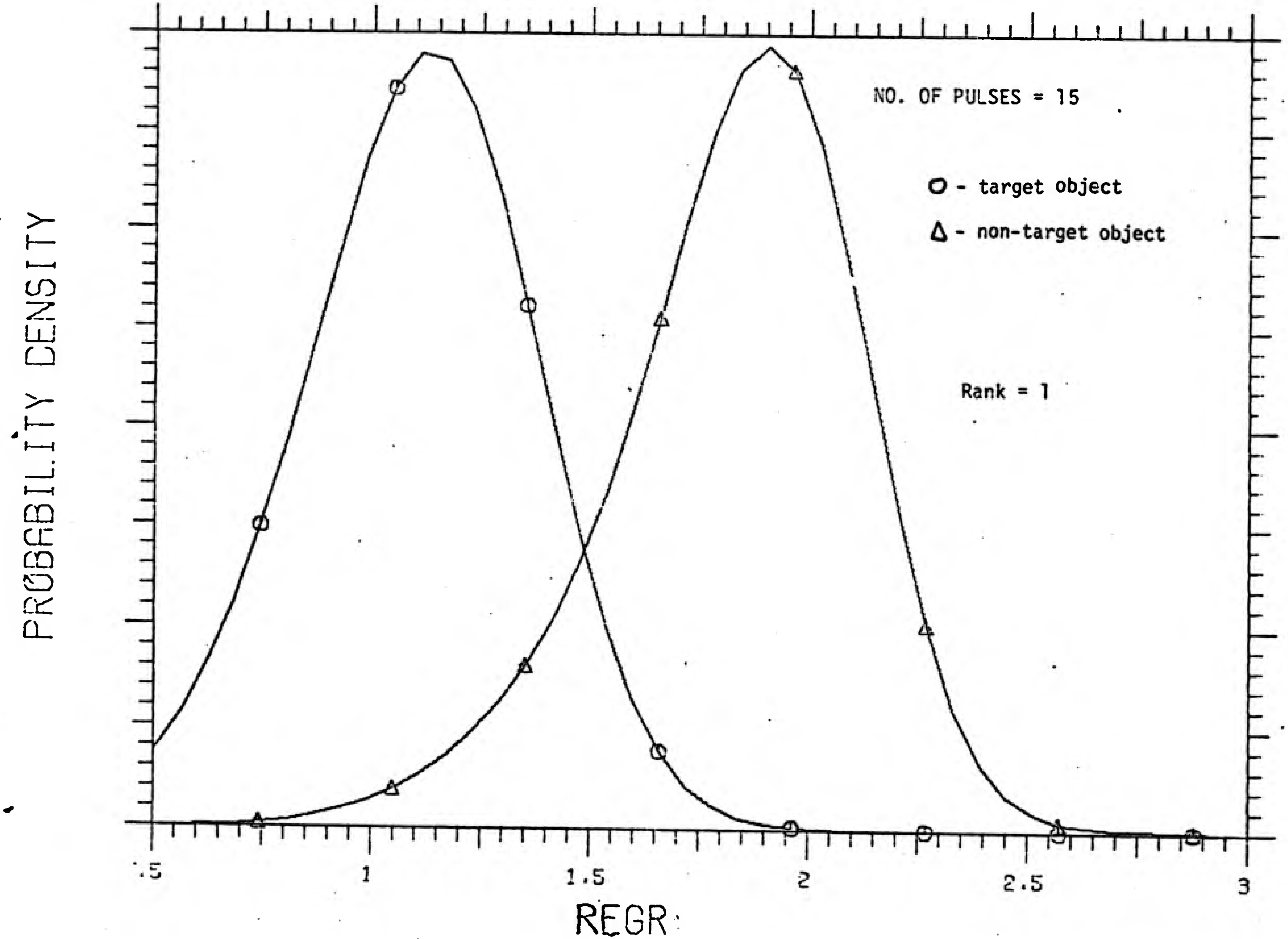

NO. OF PULSES = 15
O - target object
Δ - non-target object
Rank = 1
PROBABILITY DENSITY
.5
1
1.5
2
2.5
3
REGR

Figure 5-1a

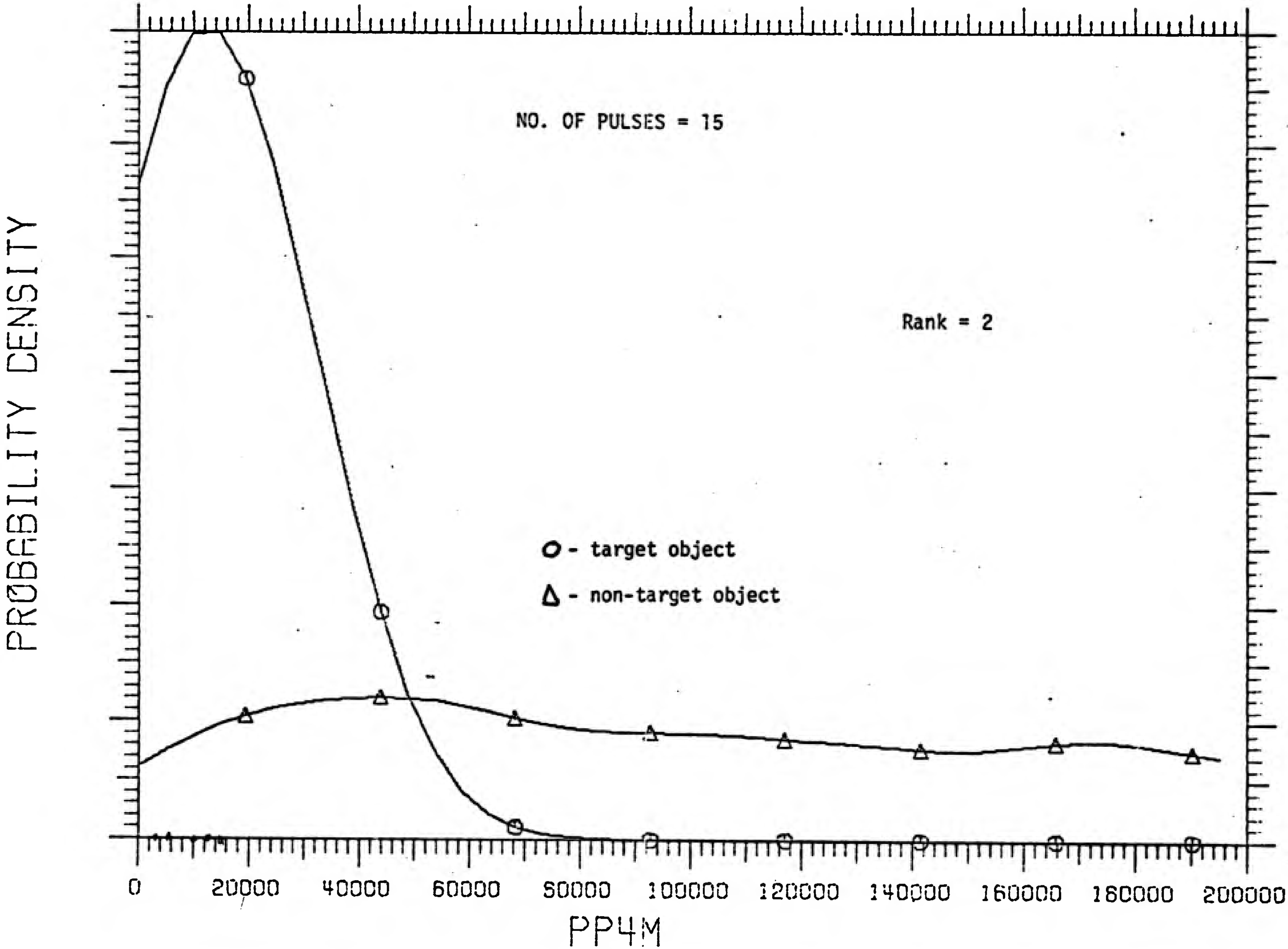

NO. OF PULSES = 15
Rank = 2
O - target object
Δ - non-target object
PROBABILITY DENSITY
0
20000
40000
60000
80000
100000
120000
140000
160000
160000
200000
PP4M

Figure 5-1b

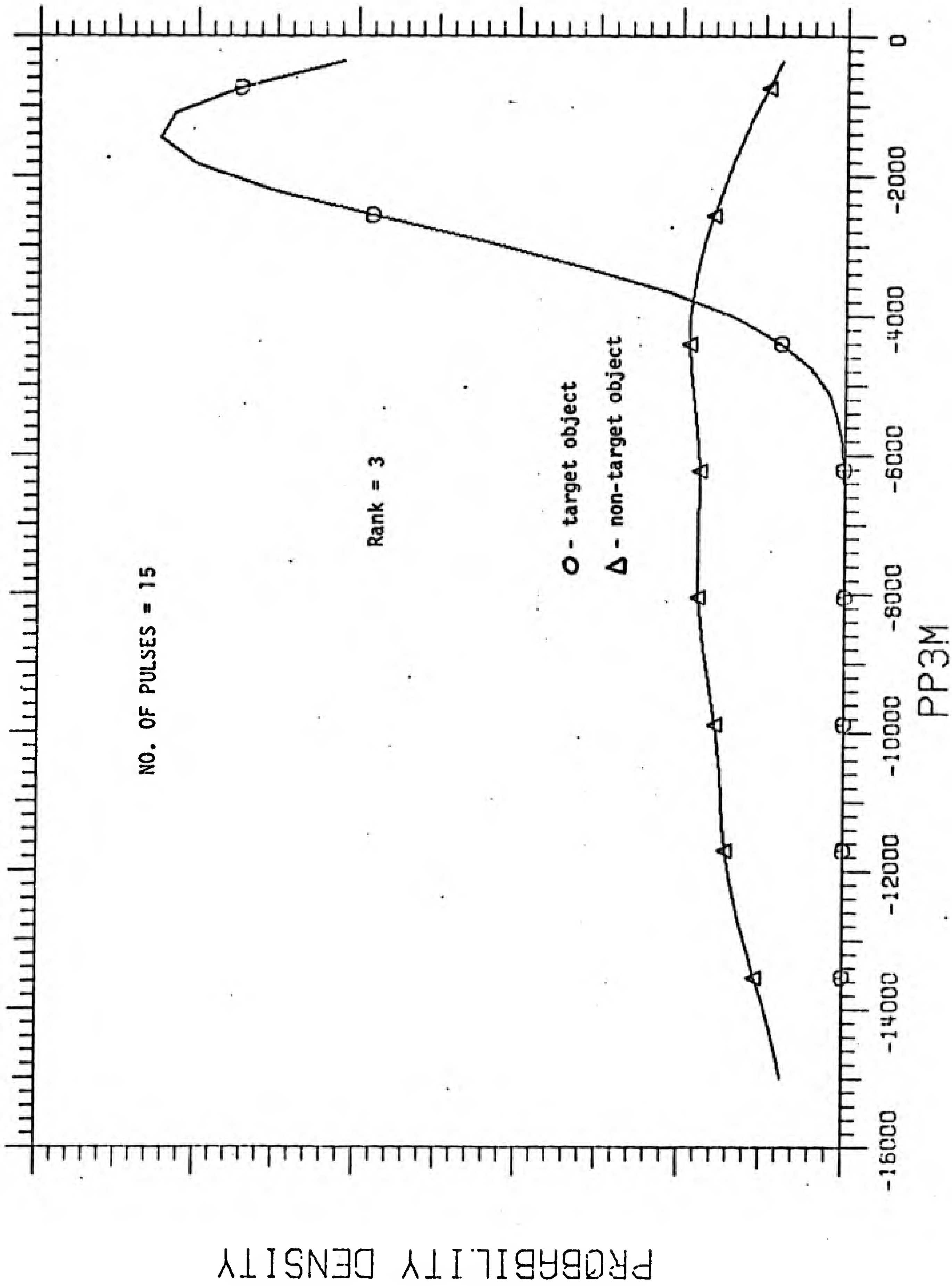

NO. OF PULSES = 15
Rank = 3
O - target object
Δ - non-target object
0
-2000
-4000
-6000
-8000
-10000
-12000
-14000
-16000
PP3M
PROBABILITY DENSITY

Figure 5-1c

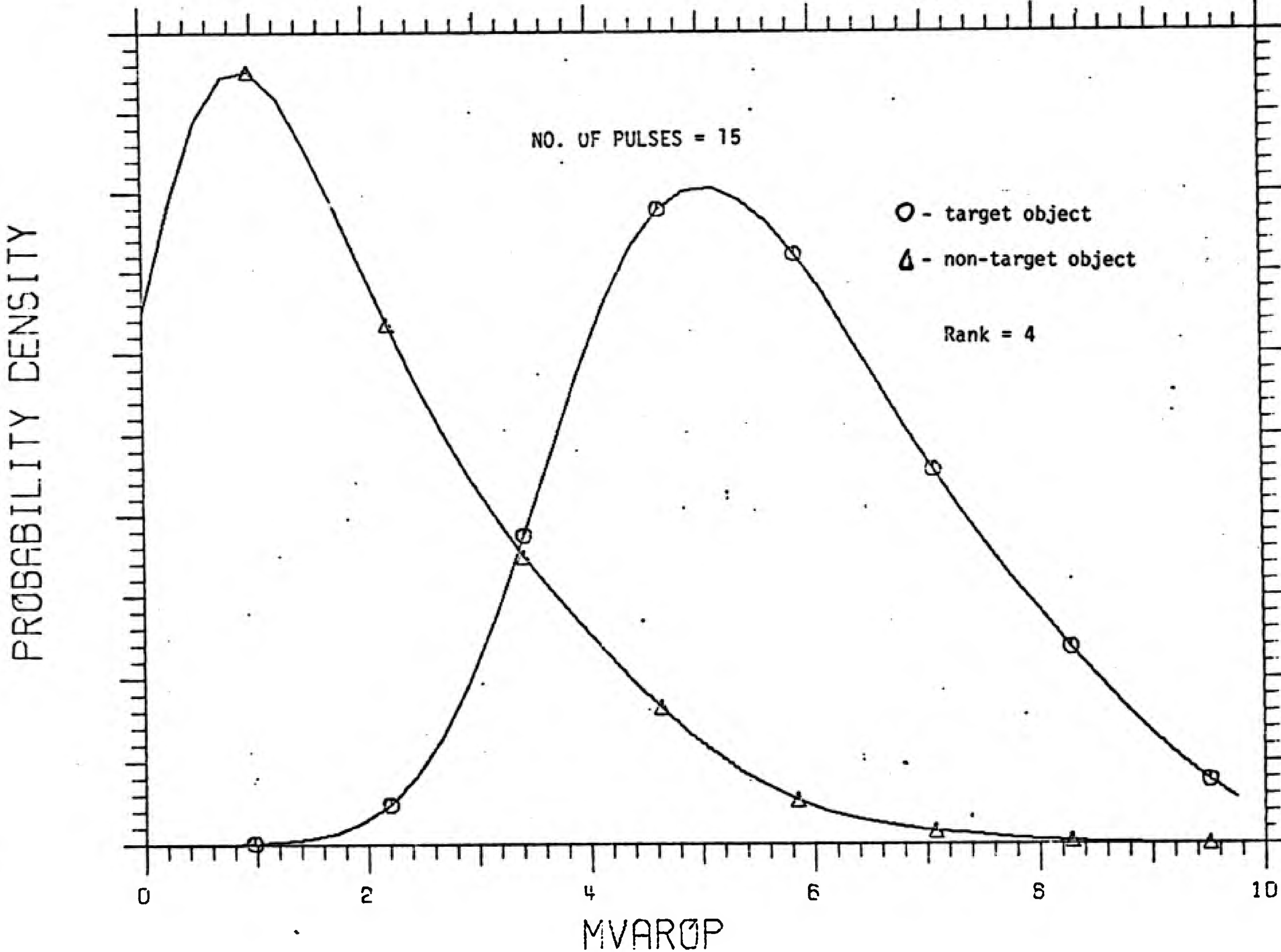

NO. OF PULSES = 15
O - target object
Δ - non-target object
Rank = 4
PROBABILITY DENSITY
0
2
4
6
8
10
MVARØP

Figure 5-1d

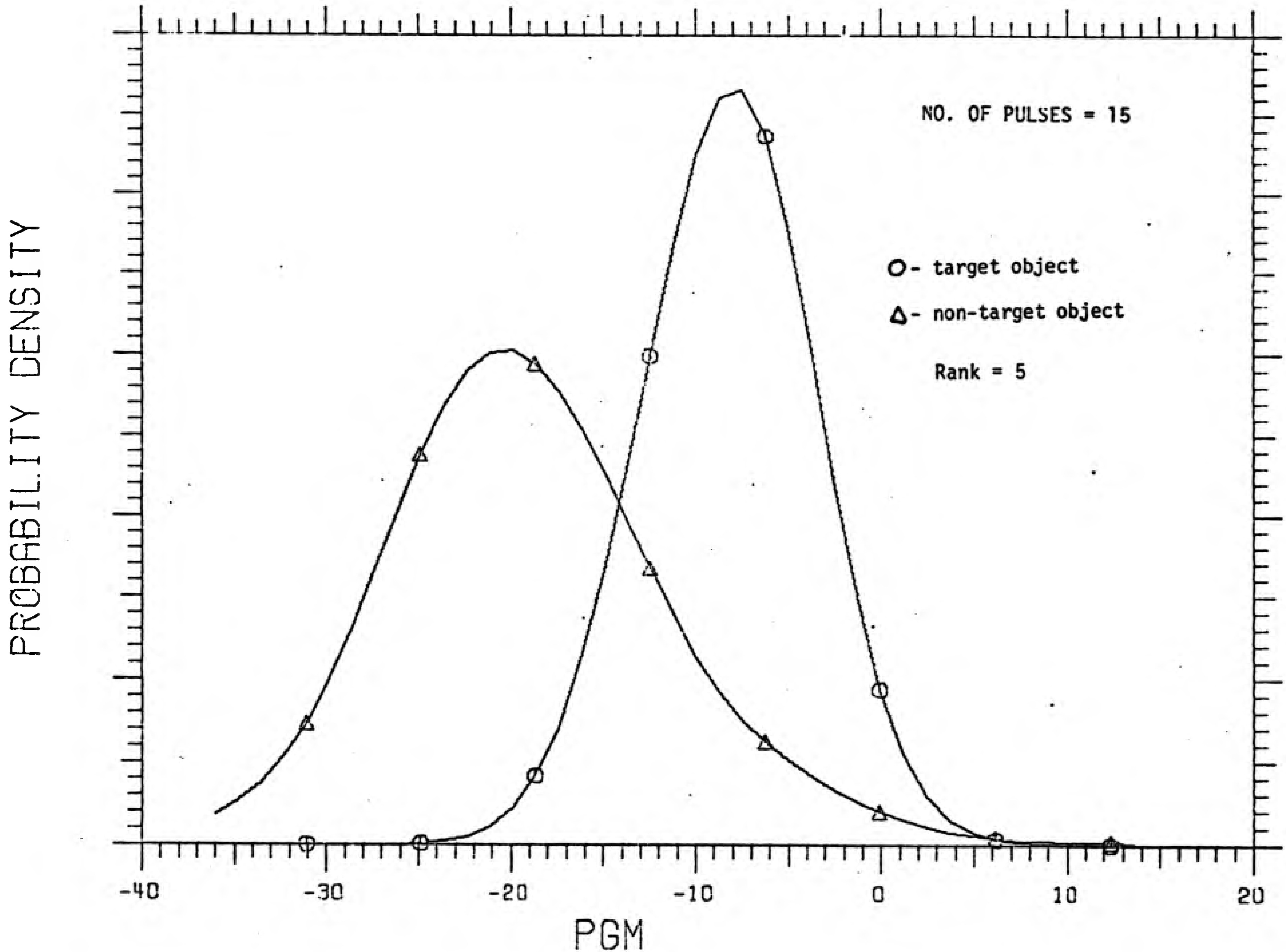

NO. OF PULSES = 15
O- target object
Δ- non-target object
Rank = 5
PROBABILITY DENSITY
-40
-30
-20
-10
0
10
20
PGM

Figure 5-1e

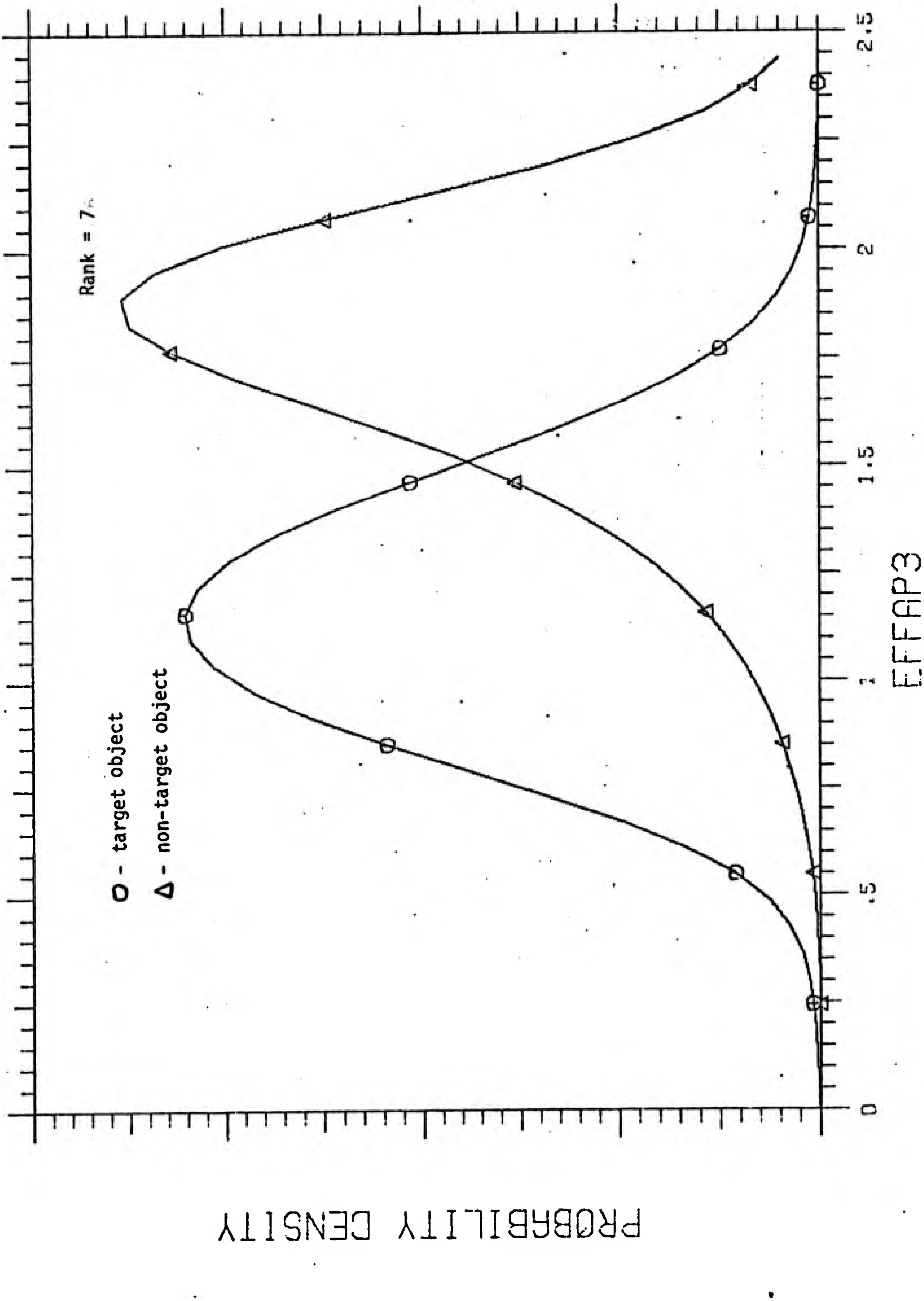

o - target object
Δ - non-target object
Rank = 7
0
.5
1
1.5
2
2.5
EFFAP3
PROBABILITY DENSITY

Figure 5-1f

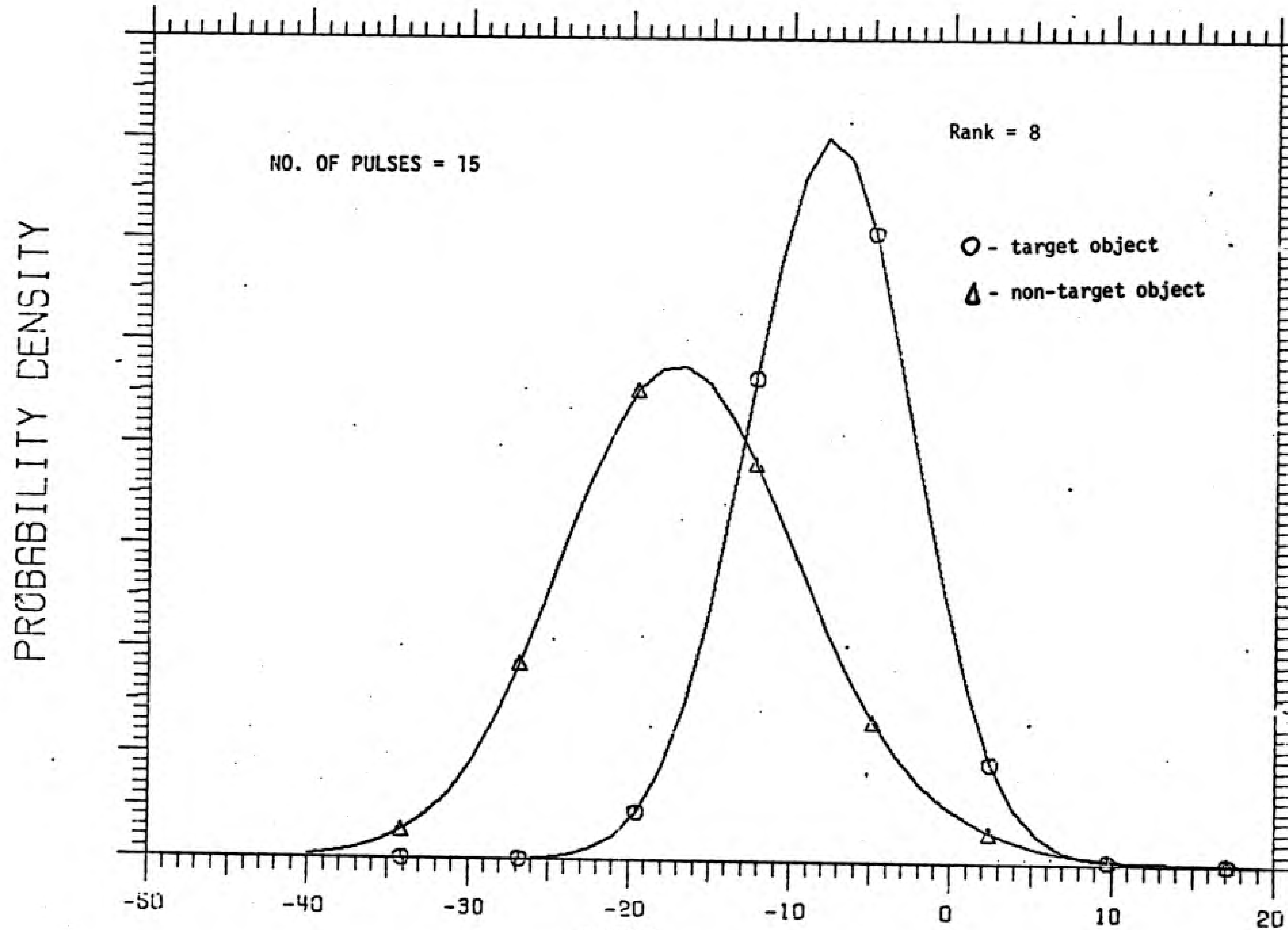
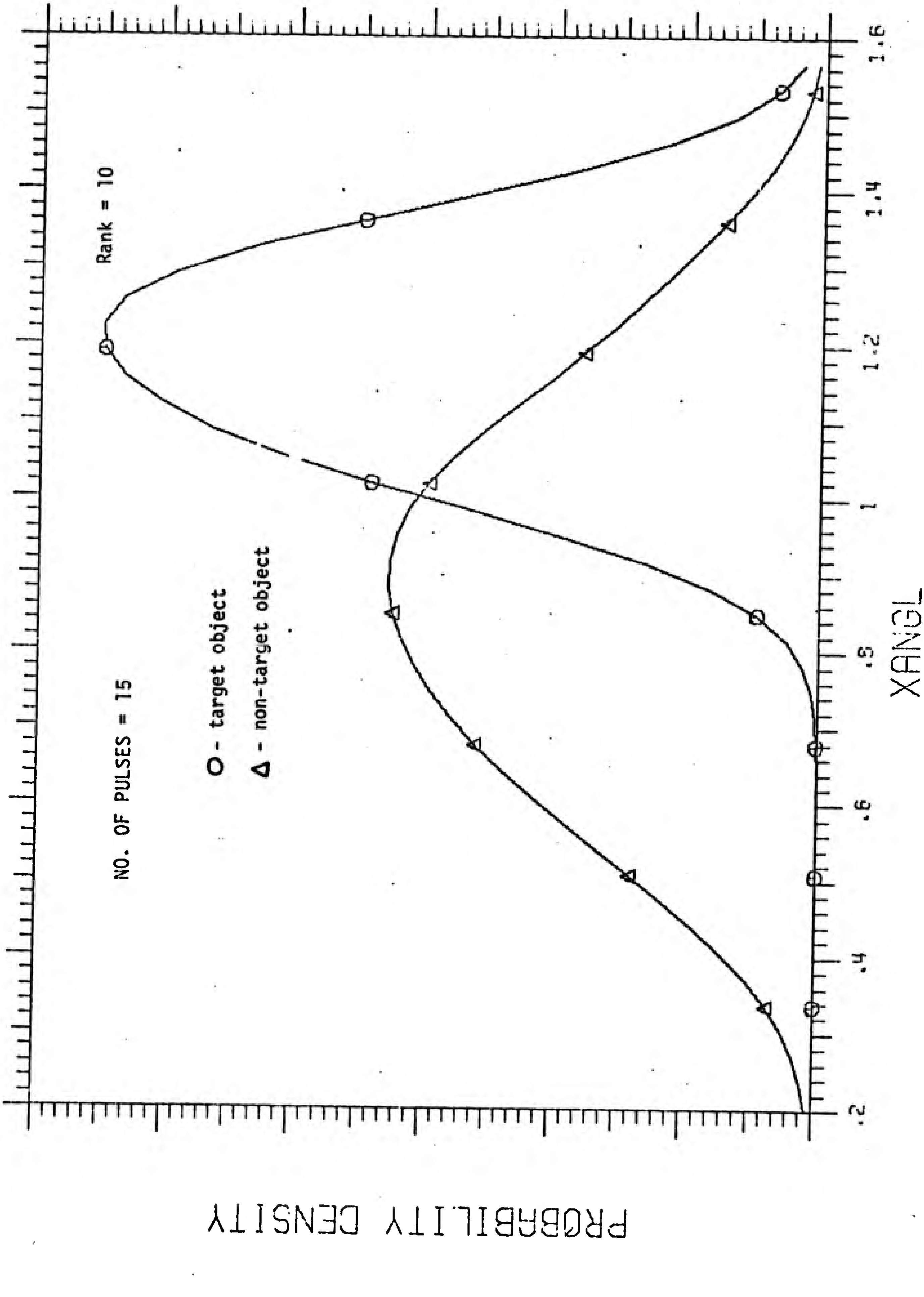

PROBABILITY DENSITY
NO. OF PULSES = 15
Rank = 8
O - target object
Δ - non-target object
-50
-40
-30
-20
-10
0
10
20
TGM

Figure 5-1g



Rank = 10
NO. OF PULSES = 15
O - target object
Δ - non-target object
.2
.4
.6
.8
1
1.2
1.4
1.6
XANGL
PROBABILITY DENSITY

Figure 5-1h

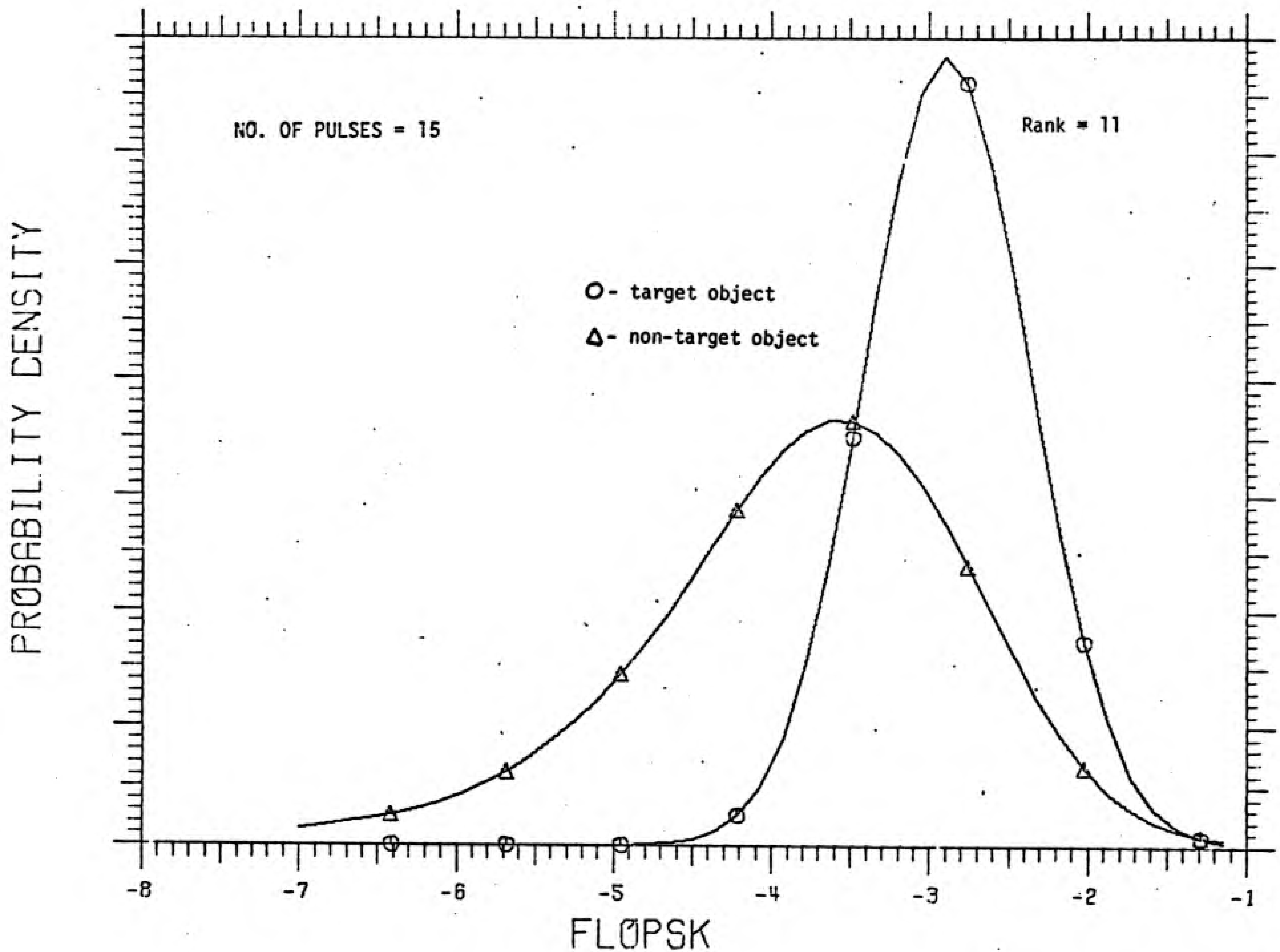

PROBABILITY DENSITY
NO. OF PULSES = 15
Rank = 11
O- target object
Δ- non-target object
-8
-7
-6
-5
-4
-3
-2
-1
FLØPSK

Figure 5-1i

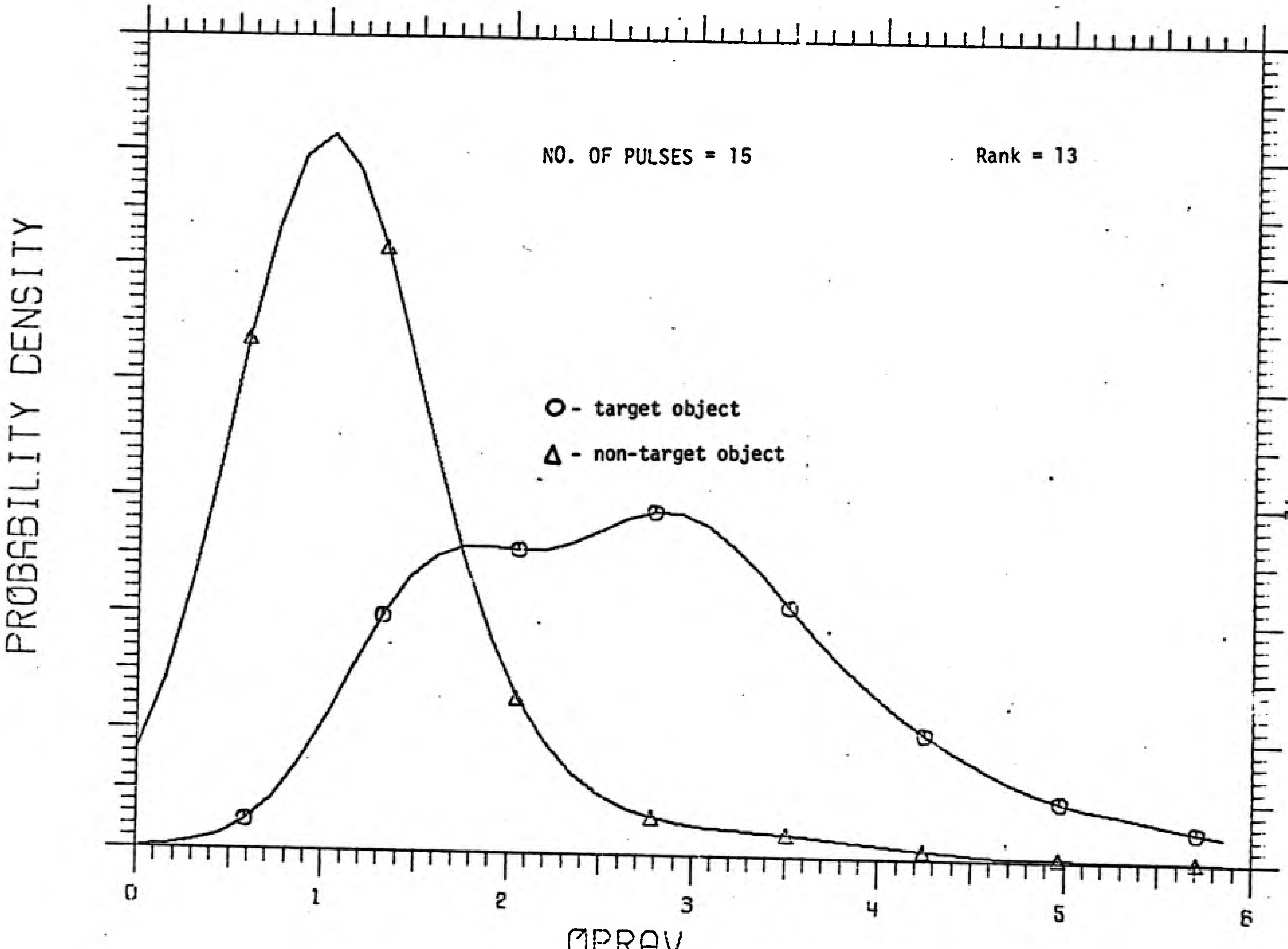

NO. OF PULSES = 15
Rank = 13
O - target object
Δ - non-target object
PROBABILITY DENSITY
0
1
2
3
4
5
6
ØPRAV

Figure 5-1j

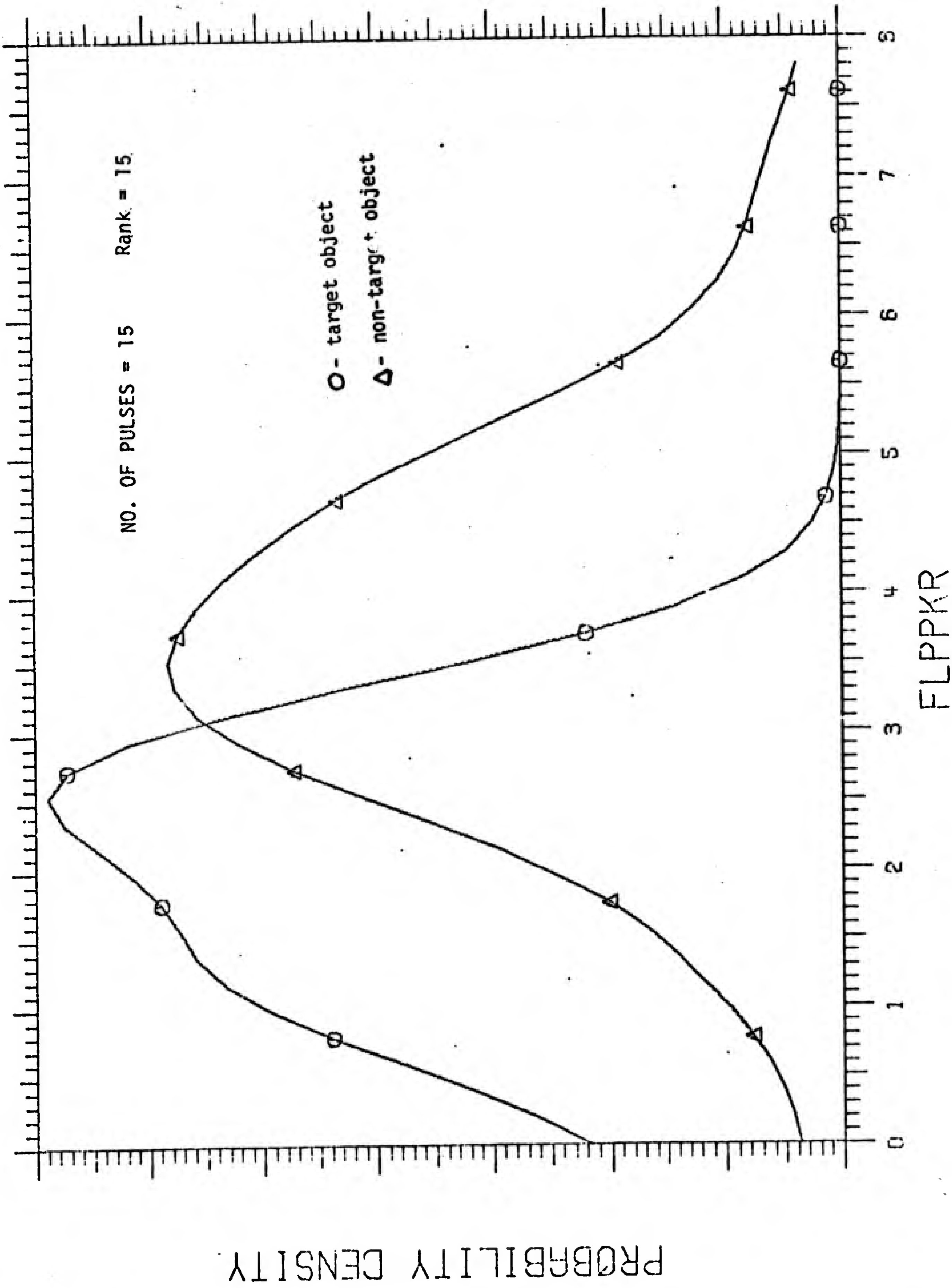

NO. OF PULSES = 15
Rank = 15
O- target object
Δ- non-targ object
FLPPKR
PROBABILITY DENSITY
0
1
2
3
4
5
6
7

Figure 5-1k

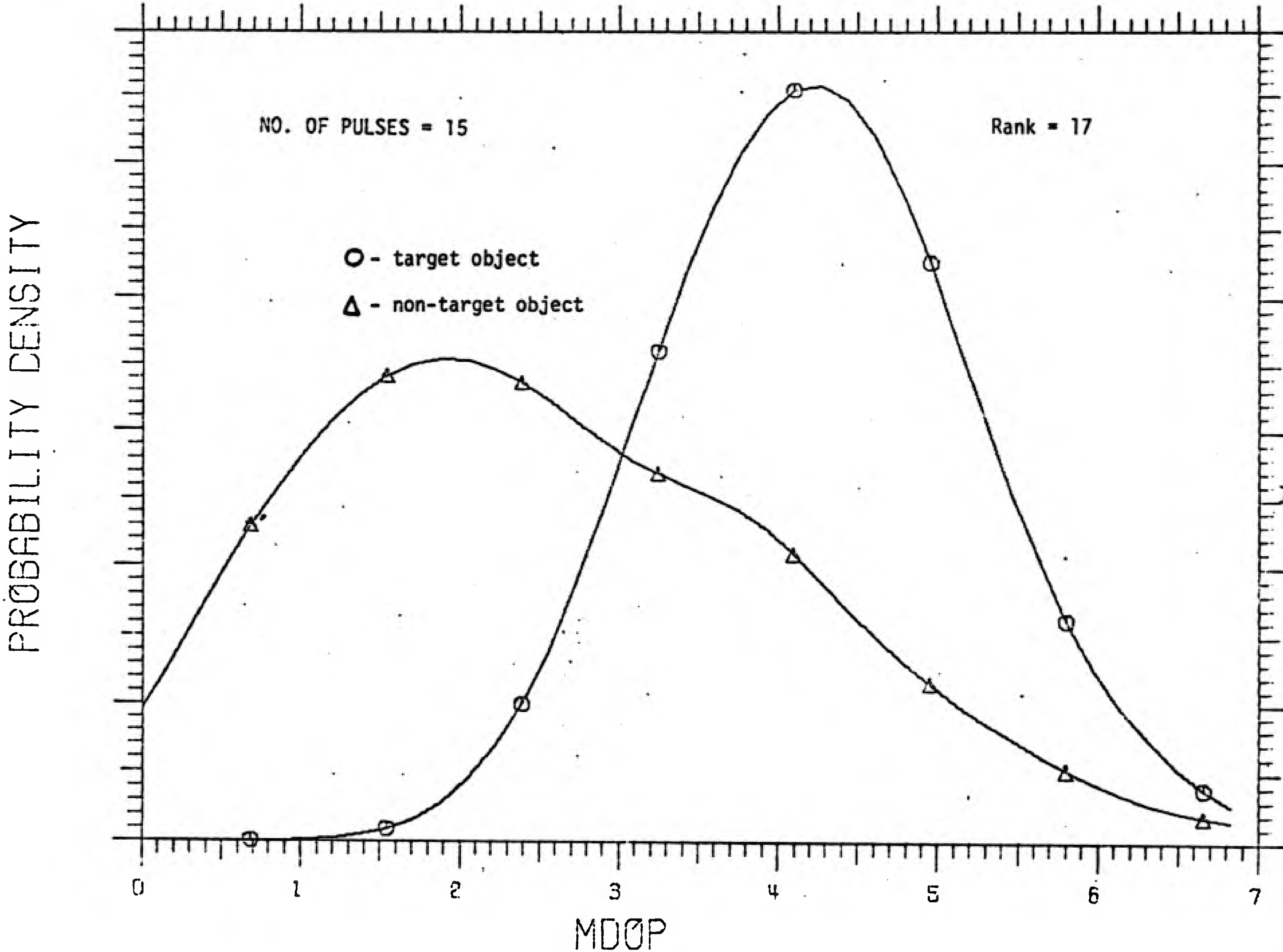

NO. OF PULSES = 15
Rank = 17
O - target object
Δ - non-target object
PROBABILITY DENSITY
0
1
2
3
4
5
6
7
MDOP

Figure 5-1ℓ

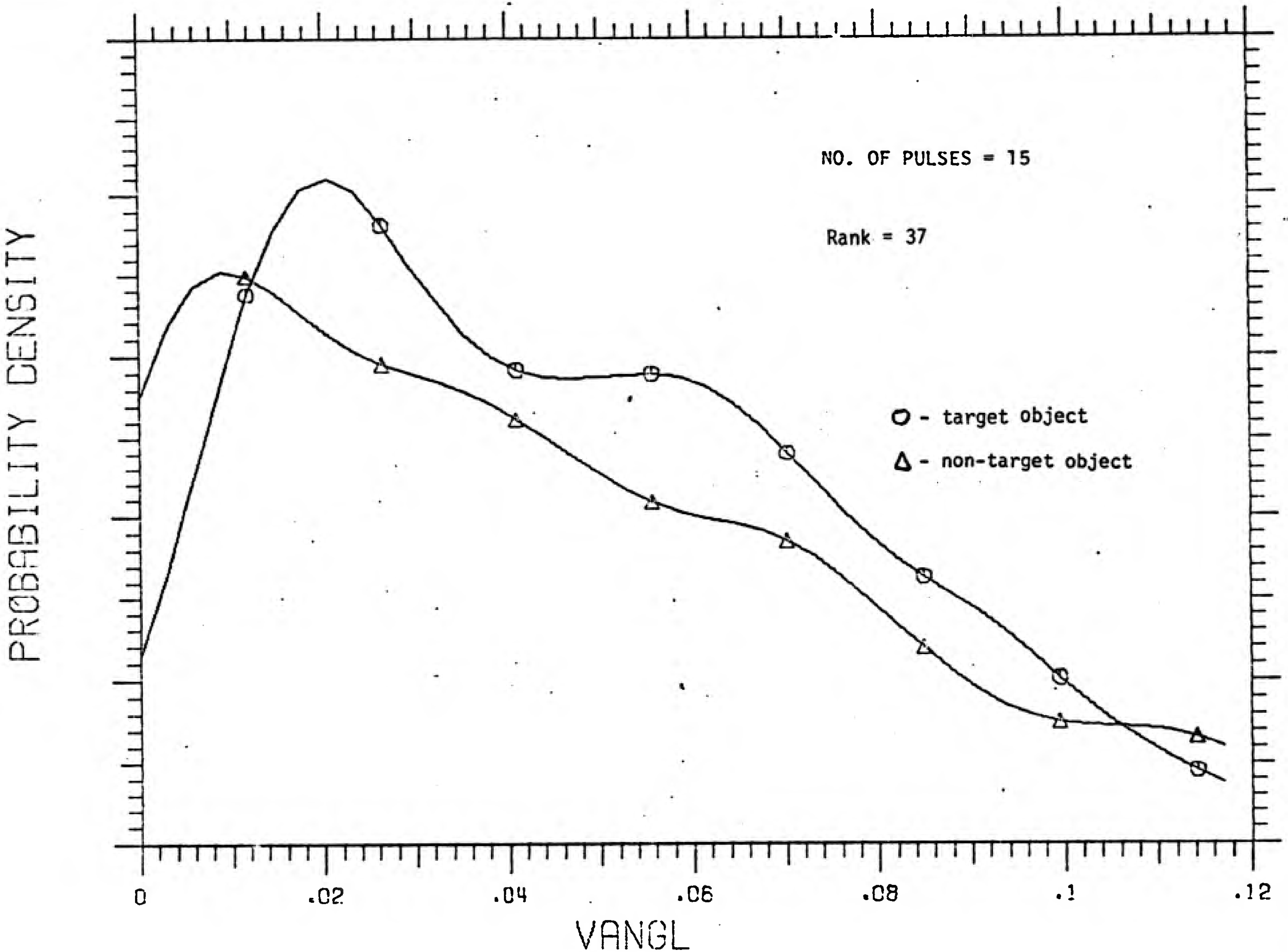

NO. OF PULSES = 15
Rank = 37
O - target object
Δ - non-target object
PROBABILITY DENSITY
0
.02
.04
.06
.08
.1
.12
VANGL

Figure 5-1m

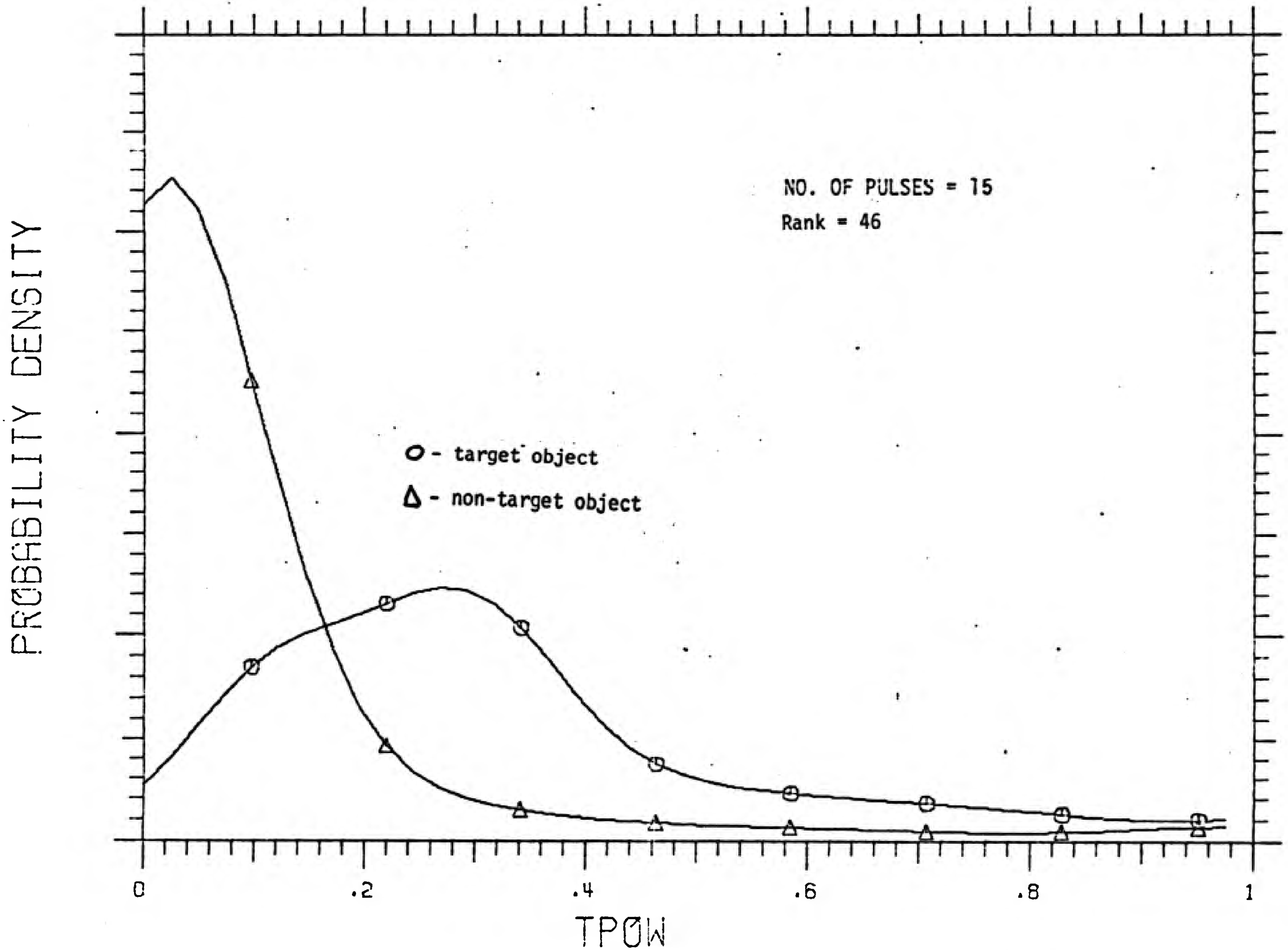

NO. OF PULSES = 15
Rank = 46
O - target object
Δ - non-target object
PROBABILITY DENSITY
0
.2
.4
.6
.8
1
TPØW

Figure 5-1n

CORRELATION COEFFECIENTS

	PPOW	VPPOW	OPOW	VOPOW	TPOW	VTPOW	PGM	VPGM	OGM	VOGM	TGM	VTGM
PPOW	1.00000	.84234	.36135	.06948	.99922	.83680	.41120	.45904	.28030	.22524	.42663	.51427
VPPOW	.84234	1.00000	.38978	.05995	.84529	.99927	.22815	.32261	.23079	.13262	.25207	.33567
OPOW	.36135	.38978	1.00000	.83515	.39778	.41887	.23672	.21002	.40648	.19832	.28778	.18654
VOPOW	.06948	.05995	.83515	1.00000	.10363	.08649	.05617	.03874	.11599	.09112	.07089	.04044
TPOW	.99922	.84529	.39778	.10363	1.00000	.84107	.41460	.46054	.29298	.23000	.43194	.51390
VTPOW	.83680	.99927	.41887	.08649	.84107	1.00000	.22703	.32186	.23467	.13420	.25186	.33336
PGM	.41120	.22815	.23672	.05617	.41460	.22703	1.00000	.37734	.56749	.51302	.97572	.46717
VPGM	.45904	.32261	.21002	.03874	.46054	.32186	.37734	1.00000	.26906	.36689	.41310	.94754
OGM	.28030	.23079	.40648	.11599	.29298	.23467	.56749	.26906	1.00000	.19912	.70593	.19642
VOGM	.22524	.13262	.19632	.09112	.23000	.13420	.51302	.36689	.19912	1.00000	.48176	.47336
TGM	.42663	.25207	.28778	.07089	.43194	.25186	.97572	.41310	.70593	.48176	1.00000	.46682
VTGM	.51427	.33567	.16654	.04044	.51390	.33336	.46717	.94754	.19642	.47336	.46682	1.00000
XANGL	.21834	.05270	-.04933	-.02723	.21276	.05850	.71175	.16547	-.15245	.44157	.55062	.33012
VANGL	.06805	.07724	.15745	.04958	.07360	.07937	.06984	.53763	.30733	.27475	.14503	.35435
XMPA	.33619	.17761	.02411	-..2416	.33182	.17430	.73891	.49285	-.02748	.40718	.61480	.58184
OPRAV	-.19011	-.36198	-.07588	-.00392	-.19018	-.35625	.22934	-.03811	.00801	-.01113	.20543	-.01427
OPSAV	-.03237	-.01004	-.80335	-.99686	-.06578	-.03568	-.04663	-.01738	-.10182	-.07562	-.05955	-.01662
OPRAT	-.16995	-.08778	-.02294	.00406	-.16819	-.08670	-.54445	-.28589	.08859	-.33539	-.38697	-.33713
OPSUR	.01152	.01609	.77369	.88140	.04401	.03955	-.01215	-.01106	.27252	.02551	.06155	-.03887
VARRAT	-.04155	-.02102	-.01676	-.00211	-.04167	-.02087	-.20713	-.10145	-.06778	-.05752	-.17302	-.08519
OPRAT2	-.07877	-.00427	.06273	.01075	-.07486	-.00179	-.39634	.06358	.23360	-.25000	-.21468	-.12338
OPSUR2	.16144	.27600	.80287	.63126	.19275	.30217	.03954	.14756	.41097	.04093	.14936	.04438
MOPP	.39493	.25894	.19888	.03117	.39700	.25605	.36627	.95321	.31307	.38704	.41081	.88503
MOOP	.21477	.11644	.14688	.05883	.21753	.11712	.53589	.35668	.14662	.95719	.49164	.47110
MVARPP	.26459	.15080	.16902	.02960	.28716	.15052	.57523	.66549	.46924	.43978	.61785	.63192
MVAROP	.17391	.05321	.11143	.05560	.17583	.05348	.69725	.27471	.30602	.74145	.66811	.36812
RATVOP	-.07220	-.06111	-.03857	.01064	-.07267	-.06051	.18130	-.29508	-.21383	.33288	.10181	-.19940
RATVPVP	-.20637	-.12963	-.09416	-.02762	-.20704	-.12919	-.05377	-.53454	-.08263	-.10775	-.06799	-.49025
RATVOVO	-.08342	-.06797	-.05195	-.02386	-.08426	-.06844	-.09910	-.15533	.15348	-.45968	-.03652	-.20804
RVPGM	.23642	.37010	.02532	-.03273	.23370	.36423	-.08627	.04961	-.04048	.09966	-.08654	.07298
RVOGM	-.42270	-.49202	-.72726	-.34269	-.44654	-.51440	-.45224	-.53423	-.48683	-.51354	-.48929	-.33070
PP3M	.23664	.13102	.15630	.03231	.24141	.13041	.89737	.19940	.56566	.37995	.87520	.26796
OP3M	.17550	.12929	.25105	.05314	.18528	.13049	.53658	.20287	.93997	.09625	.66677	.13560
PP4M	-.15097	-.07058	-.11920	-.02262	-.15358	-.07038	-.60747	-.13130	-.53756	-.29930	-.78657	-.18810
OP4M	-.15084	-.10666	-.20818	-.03988	-.15721	-.10739	-.51184	-.17931	-.88414	-.05203	-.63185	-.11847
PPSKEW	.03477	.01742	.02943	.00438	.03546	.01735	.16507	.10794	.12790	.13232	.17521	.09683
OPSKEW	.04150	.02092	.03397	.00530	.04227	.02084	.19216	.10452	.11939	.18134	.19311	.10625
PPKUR	-.02744	-.01372	-.02342	-.00346	-.02799	-.01367	-.13715	-.08602	-.10512	-.10786	-.14427	-.07706
OPKUR	-.02973	-.01490	-.02499	-.00375	-.03031	-.01485	-.15150	-.07506	-.09489	-.13176	-.15065	-.07824
RATSKEW	-.02171	-.03755	-.02557	-.00525	-.02244	-.03709	-.01330	.06588	-.07026	-.09813	-.02211	.05910
RATKUR	-.00588	.00294	-.01284	-.00083	-.00653	.00270	-.03595	.00954	-.06486	-.01479	-.03967	.00421
FLPPKR	-.39631	-.22331	-.24323	-.06661	-.40022	-.22301	-.79088	-.70277	-.50921	-.54430	-.80537	-.73149
FLOPSK	.28997	.20924	.31876	.11041	.29878	.21292	.67446	.40049	.52959	.76969	.69296	.43356
CLASS	-.08770	.04346	.00945	.03125	-.08589	.04511	-.69483	-.10763	-.25134	-.49444	-.64479	-.22366

Table 5-3a

CORRELATION COEFFECIENTS

	XANGL	VANGL	XMPA	OPRAV	OPSAV	OPRAT	OPSQR	VARRAT	OPRAT2	OPSQR2	MOPP	MOOP
PPOW	.21834	.06805	.53619	-.19011	-.03237	-.16995	.01152	-.04155	-.07877	.16144	.39493	.21477
VPPUW	.06270	.07724	.17761	-.36198	-.01004	-.08778	.01609	-.02102	-.00427	.27600	.25894	.11644
OPOW	-.04935	.15745	.02411	-.07388	-.80335	-.02294	.77369	-.01876	.06273	.80287	.19888	.14688
VOPUW	-.02723	.04958	-.02416	-.00392	-.99686	.00406	.88140	-.00211	.01075	.63126	.03117	.05883
TPOW	.21276	.07360	.53182	-.19018	-.06578	-.16819	.04401	-.04167	-.07486	.19275	.39700	.21753
VTPOW	.05850	.07937	.17430	-.35825	-.03566	-.08670	.03935	-.02087	-.00179	.30217	.25505	.11712
PGM	.71175	.06984	.73891	.22934	-.04663	-.54445	-.01815	-.20713	-.59634	.03954	.36627	.53569
VPGM	.16547	.55763	.49285	-.03611	-.01738	-.28569	-.01106	-.10145	.06358	.14756	.95321	.35868
OGM	-.15246	.30735	-.02748	.00801	-.10182	.08859	.27252	-.06778	.23360	.41097	.31507	.14662
VOGM	.44157	.27475	.40718	-.01113	-.07562	-.53539	.02551	-.05752	-.25000	.04093	.38704	.95719
TGM	.55062	.14503	.61480	.20545	-.05955	-.38697	.06155	-.17302	-.21468	.14936	.41081	.49164
VTGM	.33012	.35435	.58184	-.01427	-.01862	-.33715	-.03887	-.08519	-.12338	.04438	.88503	.47110
XANGL	1.00000	-.17785	.90522	.24494	.02681	-.78493	-.26486	-.21011	-.72215	-.31295	.12556	.51103
VANGL	-.17785	1.00000	.08441	-.05605	-.04582	-.11201	.05337	-.02020	.30024	.27930	.61664	.26970
XMPA	.90522	.08441	1.00000	.23036	.02945	-.82439	-.26665	-.25017	-.59548	-.21031	.44228	.47345
OPRAV	.24494	-.05605	.23036	1.00000	-.04299	-.15169	-.04386	-.03866	-.13919	-.05759	-.00536	.05835
OPSAV	.02681	-.04582	.02945	-.04299	1.00000	-.00710	-.87967	.00134	-.00844	-.60708	-.01485	-.04801
OPRAT	-.78493	-.11201	-.82439	-.15169	-.00710	1.00000	.29031	.36220	.75098	.23206	-.26864	-.38591
OPSQR	-.26486	.05337	-.26665	-.04386	-.87967	.29031	1.00000	.05693	.28504	.75855	.00124	-.02394
VARRAT	-.21011	-.02020	-.25017	-.03866	.00134	.36220	.05693	1.00000	.21826	.05292	-.13350	-.04001
OPRAT2	-.72215	.30024	-.59548	-.13919	-.00844	.75098	.28504	.21826	1.00000	.44643	.08694	-.28100
OPSQR2	-.31295	.27930	-.21031	-.05759	-.60708	.23206	.75855	.05292	.44643	1.00000	.14591	.00613
MOPP	.12556	.61664	.44228	-.00536	-.01485	-.26864	.00124	-.13350	.08694	.14591	1.00000	.38517
MOOP	.51103	.26970	.47345	.05835	-.04801	-.38591	-.02394	-.04001	-.28100	.00613	.38517	1.00000
MVARPP	.25739	.52816	.47023	.13410	-.02353	-.30664	.03676	-.13895	.02529	.19067	.69916	.46211
MVAROP	.55612	.21253	.53991	.19537	-.05162	-.41154	-.01991	-.11783	-.29427	.00773	.29814	.76116
RATVOP	.40123	-.24476	.20296	.07445	-.01298	-.20896	-.06882	.13688	-.30051	-.12425	-.36817	.35904
RATVPVP	.04418	-.32322	-.17736	-.01106	.02175	.04975	-.01933	.17986	-.12745	-.07658	-.63962	-.11612
RATVOVO	-.27322	-.16176	-.23875	-.00005	.01869	.26105	.07592	-.06735	.19952	.05295	-.17134	-.55374
RVPGM	-.05769	-.04547	-.08176	-.95757	.08400	.05334	-.02215	.01673	.02427	-.03589	.00286	.03912
RVOGM	-.14054	-.27215	-.22471	.01431	.29907	.18875	-.30073	.04345	.04367	-.59992	-.34100	-.49026
PP3M	.61337	.02416	.60664	.17220	-.02731	-.50107	.00134	-.25448	-.37829	.04081	.18837	.39563
OP3M	-.14205	.24064	-.03350	.02802	-.04631	.09938	.19185	-.07625	.21720	.27751	.23882	.05415
PP4M	-.53567	-.00675	-.52132	-.15553	.02014	.45190	-.00924	.25580	.34842	-.03502	-.11999	-.30832
OP4M	.12366	-.20511	.02322	-.02990	.03443	-.08567	-.16204	.08137	-.19277	-.23299	-.20817	-.01376
PPSKEW	.08408	.12005	.11518	.02767	-.00366	-.07520	.01479	-.36552	.00685	.02980	.18446	.16285
OPSKEW	.13652	.12328	.14136	.03429	-.00442	-.10778	.01445	.00754	-.01506	.03327	.14206	.26136
PPKUR	-.07820	-.09692	-.09799	-.02246	.00285	.06725	-.01165	.37042	-.00321	-.02373	-.14995	-.13501
OPKUR	-.10965	-.09350	-.11301	-.02527	.00313	.09451	-.01277	-.00764	.01476	-.02540	-.10840	-.19481
RATSKEW	.03259	.01689	.06449	-.01482	.00452	-.03878	-.02066	-.01738	-.02138	-.02529	.07997	-.11524
RATKUR	-.00036	.00367	-.00206	.00409	.00062	-.00125	-.00329	-.00245	.00119	-.00613	.00857	-.02957
FLPPKR	-.48276	-.36852	-.66377	-.19341	.05512	.43726	-.01069	.22147	.16661	-.10607	-.76340	-.57358
FLOPSK	.36435	.39915	.40896	.09562	-.09685	-.32966	.10157	-.05448	-.11978	.22914	.45864	.83146
CLASS	-.61566	.02561	-.56348	-.27187	-.02645	.40356	.13203	.10674	.38027	.15968	-.11758	-.52860

Table 5-3b

CORRELATION COEFFECIENTS

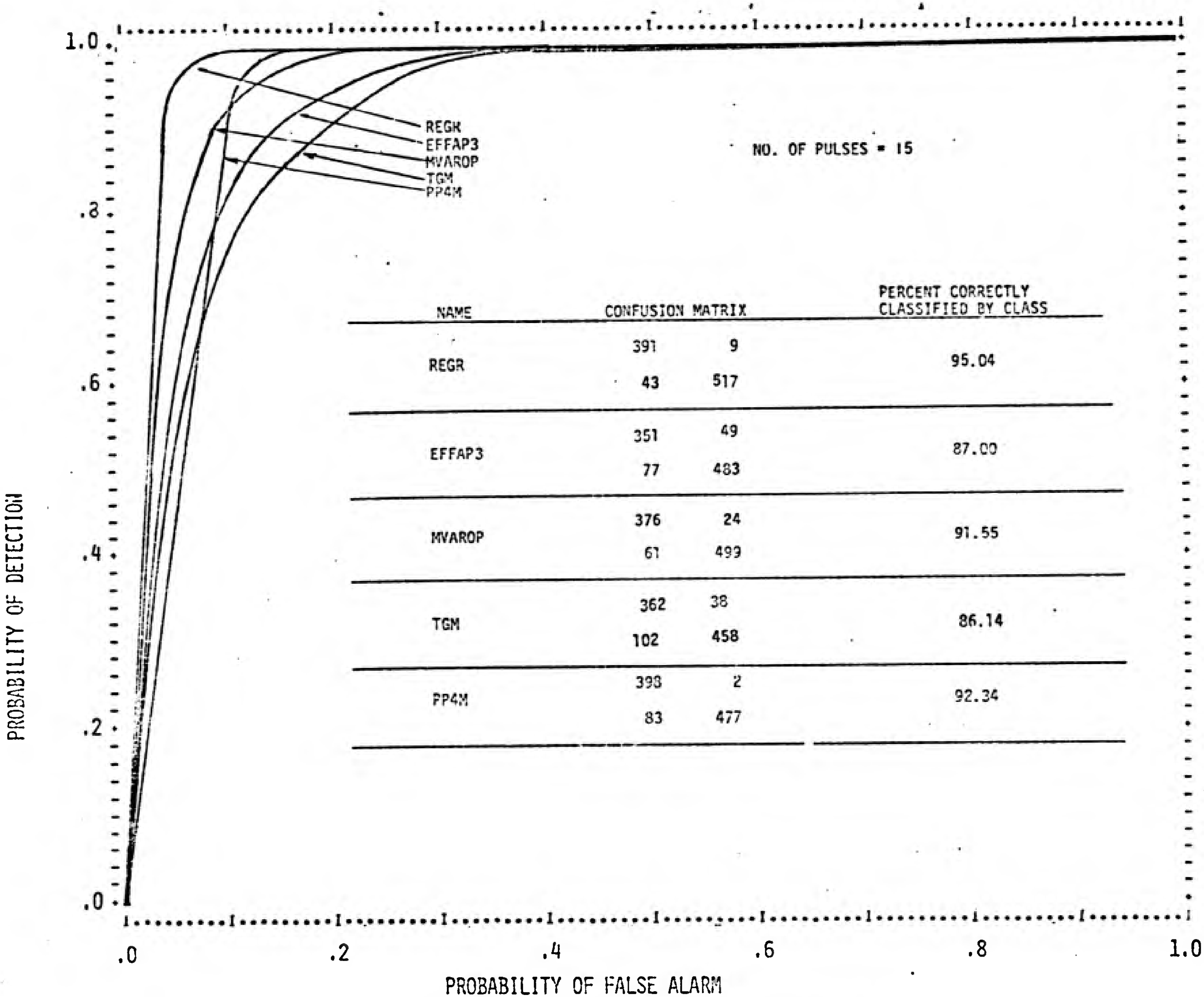
	MVARPP	MVAROP	RATVOP	RATVPVP	RATVOVO	RVPGM	RVOGM	PP3M	OP3M	PP4M	OP4M	PPSKEW
PPOW	.28459	.17391	-.07220	-.20637	-.08342	.23642	-.42270	.23864	.17550	-.15097	-.15084	.03477
VPPOW	.15080	.05321	-.06111	-.12963	-.06797	.37010	-.49202	.13102	.12929	-.07058	-.10666	.01742
OPOW	.16902	.11143	-.03857	-.09418	-.05195	.02532	-.72726	.15630	.25105	-.11920	-.20818	.02943
VOPOW	.02960	.05560	.01064	-.02762	-.02386	-.03273	-.34269	.03231	.05314	-.02262	-.03988	.00438
TPOW	.28716	.17583	-.07267	-.20704	-.08428	.23370	-.44664	.24141	.18328	-.15358	-.15721	.03546
VTPOW	.15052	.05348	-.06061	-.12919	-.06844	.36425	-.51440	.13041	.13049	-.07038	-.10739	.01735
PGM	.57523	.69726	.18130	-.05377	-.09910	-.08627	-.45224	.89737	.53658	-.80747	-.51184	.16507
VPGM	.66549	.27471	-.29508	-.53454	-.15533	.04961	-.33423	.19940	.20287	-.13130	-.17931	.10794
OGM	.46924	.30602	-.21383	-.08263	.15348	-.04048	-.48683	.56566	.93997	-.53756	-.88414	.12790
VOGM	.43978	.74145	.33288	-.10775	-.45968	.09966	-.51334	.37995	.09625	-.29930	-.05203	.13232
TGM	.61785	.66811	.10181	-.06799	-.03652	-.08554	-.46929	.87520	.66677	-.78857	-.63185	.17521
VTGM	.63192	.36812	-.19940	-.49025	-.20604	.07298	-.33670	.26796	.13660	-.18810	-.11647	.09683
XANGL	.25739	.55812	.40123	.04418	-.27322	-.05769	-.14054	.61337	-.14205	-.53587	.12366	.08908
VANGL	.52816	.21253	-.24476	-.32322	-.16176	-.04547	-.27215	.02916	.24064	-.00675	-.20511	.12005
XMPA	.47023	.53991	.20296	-.17736	-.23875	-.08176	-.22471	.60564	-.03350	-.52132	.02322	.11518
OPRAV	.13410	.19337	.07445	-.01106	.00005	-.95757	.01431	.17220	.02802	-.15553	-.02990	.02767
OPSAV	-.02353	-.05162	-.01298	.02175	.01669	.08400	.29907	-.02731	-.04631	.02014	.03443	-.00356
OPRAT	-.30664	-.41154	-.20696	.04975	.26105	.05334	.18875	-.50107	.09938	.45190	-.08567	-.07620
OPSQR	.03676	-.01991	-.06682	-.01933	.07592	-.02215	-.30073	.00134	.19185	-.00924	-.15204	.01479
VARRAT	-.13895	-.11783	.13688	.17986	-.06735	.01673	.04345	-.25448	-.07825	.25580	.08137	-.38552
OPRAT2	.02529	-.29427	-.30051	-.12743	.19952	.02427	.04387	-.37829	.21720	.34842	-.19277	.00685
OPSQR2	.19067	.00773	-.12425	-.07658	.05295	-.03589	-.59992	.04081	.27751	-.03602	-.23299	.02980
MOPP	.69916	.29814	-.36817	-.63962	-.17134	.00286	-.34100	.18837	.23882	-.11999	-.20817	.18446
MOOP	.46211	.76116	.35904	-.11612	-.55374	.03912	-.49025	.39563	.05415	-.30832	-.01376	.16235
MVARPP	1.00000	.63558	-.26723	-.28648	-.03376	-.11117	-.38125	.46046	.43111	-.38942	-.40030	.17370
MVAROP	.63568	1.00000	.32233	.03116	-.15178	-.08217	-.40414	.58578	.28940	-.50049	-.26790	.14326
RATVOP	-.26723	.32233	1.00000	.44122	-.23808	.01749	-.06679	.17673	-.27076	-.15165	.29306	-.17448
RATVPVP	-.28648	.03116	.44122	1.00000	.14264	.03576	.12038	.04289	-.03411	-.07162	.01482	-.46582
RATVOVO	-.03376	-.15178	-.23808	.14264	1.00000	-.02737	.24869	-.02581	.21813	-.00993	-.22945	-.04889
RVPGM	-.11117	-.08217	.01749	.03876	-.02737	1.00000	-.00708	-.06471	-.04606	.06630	.04203	-.01657
RVOGM	-.38125	-.40414	-.06679	.12036	.24869	-.00708	1.00000	-.34083	-.32400	.27569	.27202	-.09453
PP3M	.46046	.58578	.17673	.04289	-.02581	-.06471	-.34083	1.00000	.60771	-.98220	-.60757	.16809
OP3M	.43111	.28940	-.27076	-.03411	.21813	-.04606	-.32400	.60771	1.00000	-.60900	-.98869	.12284
PP4M	-.38942	-.50049	-.15165	-.07162	-.00993	.06630	.27569	-.98220	-.60900	1.00000	.62154	-.15068
OP4M	-.40030	-.26790	.29306	.01482	-.22945	.04203	.27202	-.60757	-.98869	.62154	1.00000	-.11113
PPSKEW	.17370	.14326	-.17448	-.46582	-.04889	-.01657	-.09453	.16809	.12284	-.15068	-.11113	1.00000
OPSKEW	.18001	.18296	.10147	-.05565	-.42953	-.01744	-.13116	.20259	.09298	-.18645	-.07236	.11290
PPKUR	-.13961	-.11501	.15172	.41846	.03738	.01323	.07686	-.14375	-.09911	.12992	.08953	-.99083
OPKUR	-.13442	-.13372	-.08117	.03249	.34220	.01285	.09665	-.17328	-.07506	.16505	.05852	-.07233
RATSKEW	.01466	-.06370	-.09756	-.08774	.20027	.00188	.06782	-.02994	-.04850	.02979	.03442	.01422
RATKUR	.00437	-.01913	-.03510	-.01088	.07406	.00265	.02601	-.03510	-.05506	.03195	.04546	.00194
FLPPKR	-.77167	-.64137	.16086	.49835	.15792	.09625	.45793	-.60087	-.44897	.50026	.41372	-.33472
FLOPSK	.61534	.71638	.20195	-.15041	-.53353	-.04343	-.64004	.55424	.40985	-.46592	-.34742	.23484
CLASS	-.50668	-.80481	-.24028	-.12320	.05581	.13895	.24437	-.57462	-.28006	.49238	.27534	-.09583

Table 5-3c

CORRECTION COEFFECIENTS

	UPSKEW	PPKUR	OPKUR	RATSKEW	RATKUR	FLPPKR	FLOPSK	CLASS
PPOW	.04150	-.02744	-.02973	-.02171	-.00588	-.39631	.28997	-.08770
VPPOW	.02092	-.01372	-.01490	-.03755	.00294	-.22351	.20924	.04346
OPOW	.03397	-.02342	-.02499	-.02557	-.01284	-.24323	.31876	.00945
VUPOW	.00530	-.00346	-.00375	-.00525	-.00083	-.06661	.11041	.03125
TPOW	.04227	-.02799	-.03031	-.02244	-.00653	-.40022	.29378	-.05589
VTPOW	.02084	-.01367	-.01485	-.03709	.00270	-.22301	.21292	.04511
PGM	.19216	-.13713	-.15150	-.01330	-.03595	-.79088	.67446	-.69483
VPGM	.10452	-.06602	-.07806	.06588	.00954	-.70277	.40049	-.10763
OGM	.11939	-.10312	-.09489	-.07026	-.06486	-.50921	.52959	-.25134
VUGM	.18134	-.10786	-.13176	-.09813	-.01479	-.54430	.76969	-.49444
TGM	.19311	-.14427	-.15065	-.02211	-.03967	-.80537	.69295	-.64479
VTGM	.10625	-.07706	-.07624	.05910	.00421	-.73149	.43356	-.22366
XANGL	.13652	-.07820	-.10965	.03259	-.00036	-.43276	.36933	-.61566
VANGL	.12525	-.09692	-.09350	.01689	.00367	-.36652	.39915	.02561
XMPA	.14138	-.09799	-.11301	.06449	-.00206	-.66377	.40896	-.56348
OPRAV	.03429	-.02245	-.02527	-.01482	.00409	-.19341	.09562	-.27187
OPSAV	-.00442	.00288	.00313	.00452	.00062	.05512	-.09685	-.02643
OPRAT	-.10778	.06725	.09451	-.03878	-.00125	.43726	-.32966	.40356
OPSQR	.01445	-.01165	-.01277	-.02056	-.00329	-.01069	.10157	.13203
VARRAT	.00754	.37042	-.00764	-.01738	-.00245	.22147	-.05448	.10374
OPRAT2	-.01506	-.00321	.01476	-.02138	.00119	.16561	-.11978	.38027
OPSQR2	.03327	-.02373	-.02540	-.02529	-.00613	-.10507	.22914	.15968
MDPP	.14206	-.14995	-.10640	.07997	.00857	-.76340	.45864	-.11758
MDOP	.26136	-.13501	-.19481	-.11524	-.02957	-.57358	.83146	-.52860
MVARPP	.18001	-.13961	-.13442	.01466	.00437	-.77167	.61534	-.50568
MVAROP	.18296	-.11501	-.13372	-.05370	-.01913	-.64137	.71638	-.80481
RATVOP	.10147	.15172	-.08117	-.09756	-.03510	.16085	.20195	-.24026
RATVPVP	-.05565	.41846	.03249	-.08774	-.01088	.49835	-.15041	-.12320
RATVOVO	-.42953	.03738	.34220	.20027	.07406	.15792	-.53353	.05581
RVPGM	-.01744	.01323	.01285	.00188	.00265	.09625	-.04343	.13895
RVOGM	-.13116	.07686	.09665	.06782	.02601	.45193	-.64004	.24437
PP3M	.20259	-.14375	-.17328	-.02994	-.03510	-.60087	.55424	-.57462
OP3M	.09298	-.09911	-.07506	-.04850	-.05506	-.44097	.40985	-.28006
PP4M	-.18645	.12992	.16505	.02979	.03195	.50026	-.46592	.49236
OP4M	-.07236	.08953	.05852	.03442	.04546	.41372	-.34742	.27534
PPSKEW	.11290	-.99083	-.07233	.01422	.00194	-.33472	.23484	-.09583
OPSKEW	1.00000	-.08178	-.97989	-.19051	-.09995	-.23116	.43906	-.11478
PPKUR	-.08178	1.00000	.04512	-.01163	-.00162	.28235	-.19559	.07579
OPKUR	-.97989	.04512	1.00000	.14454	.07507	.17749	-.34621	.08225
RATSKEW	-.19051	-.01163	.14454	1.00000	.24550	-.03970	-.16932	.00995
RATKUR	-.09995	-.00162	.07507	.24550	1.00000	.03522	-.08295	-.01077
FLPPKR	-.23116	.28235	.17749	-.03970	.03522	1.00000	-.70699	.55076
FLOPSK	.43906	-.19559	-.34621	-.16932	-.08295	-.70699	1.00000	-.51298
CLASS	-.11478	.07579	.08225	.00995	-.01077	.55076	-.51298	1.00000

Table 5-3d



NAME	CONFUSION MATRIX		PERCENT CORRECTLY CLASSIFIED BY CLASS
REGR	391	9	95.04
	43	517	
EFFAP3	351	49	87.00
	77	483	
MVAROP	376	24	91.55
	61	499	
TGM	362	38	86.14
	102	458	
PP4M	398	2	92.34
	83	477	

Figure 5-2 : Single feature performance.

Feature Pairs		Percent Correct Classification	Rank
TGM	REGR	96.04	1
XANGL	REGR	95.98	2
MVAROP	PP4M	95.87	3
PP4M	REGR	95.82	4
EFFAP3	REGR	95.79	5
PP4M	FLOPSK	94.56	6
TGM	XANGL	94.45	7
XANGL	MVAROP	94.32	8
TGM	MVAROP	94.09	9
PP4M	EFFAP3	94.14	10

Table 5-4: Feature Pairs Ranked at 15 Pulses (N)

8 = Target class

1-7 = Non-target classes

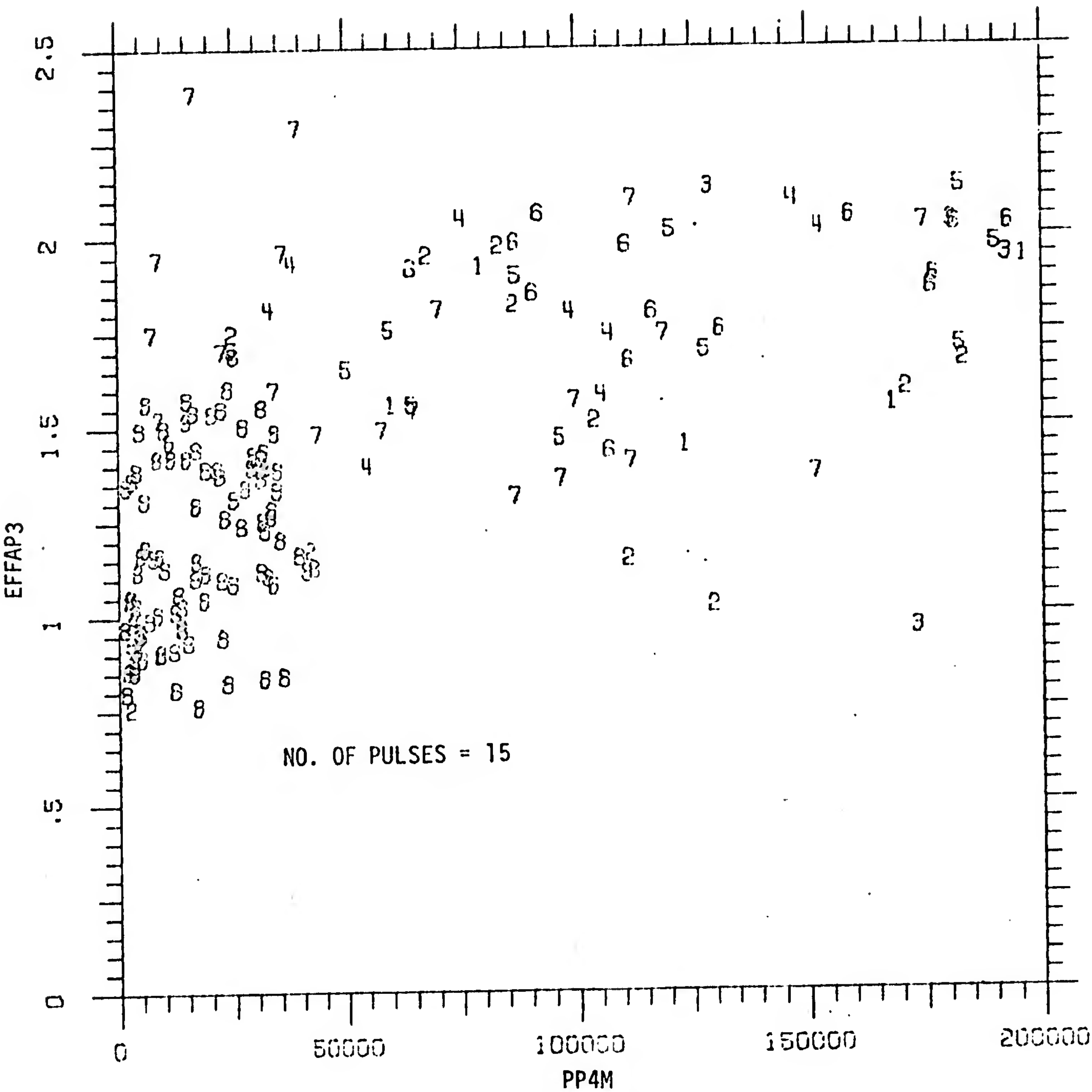


Figure 5-3a

8 = Target class

1-7 = Non-target classes

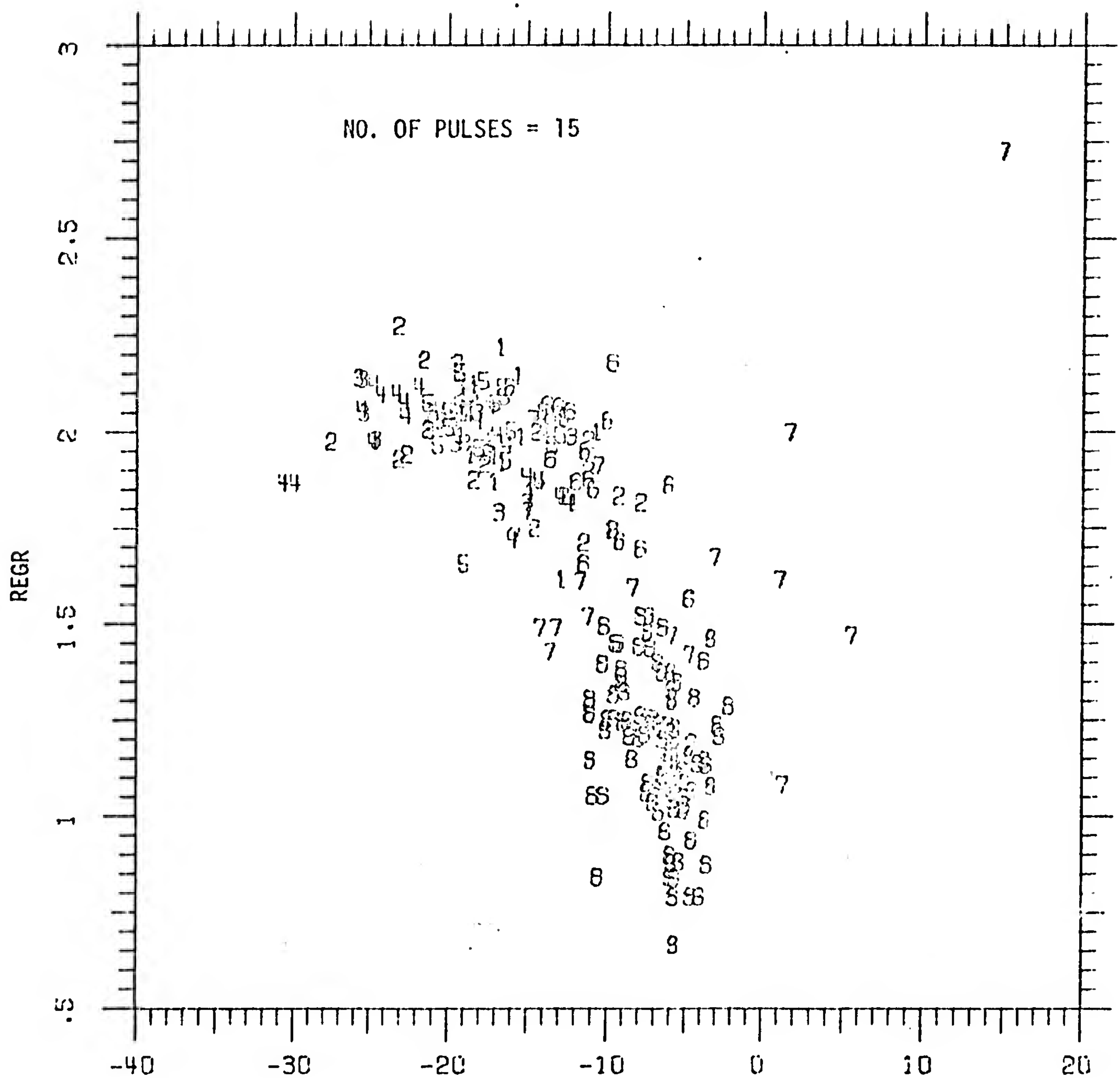


Figure 5-3b

TGM

8 = Target class
1-7 = Non-target classes

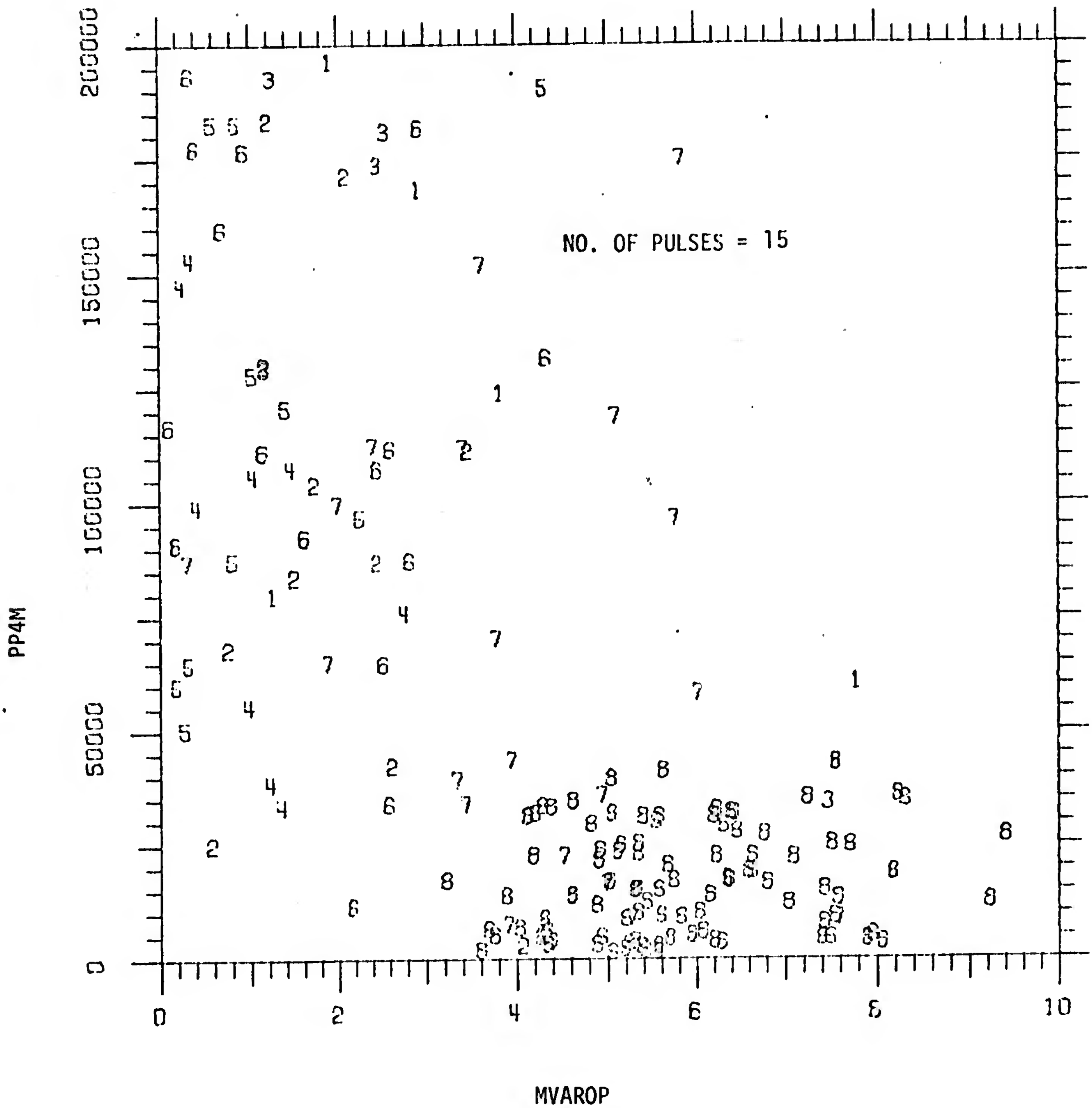
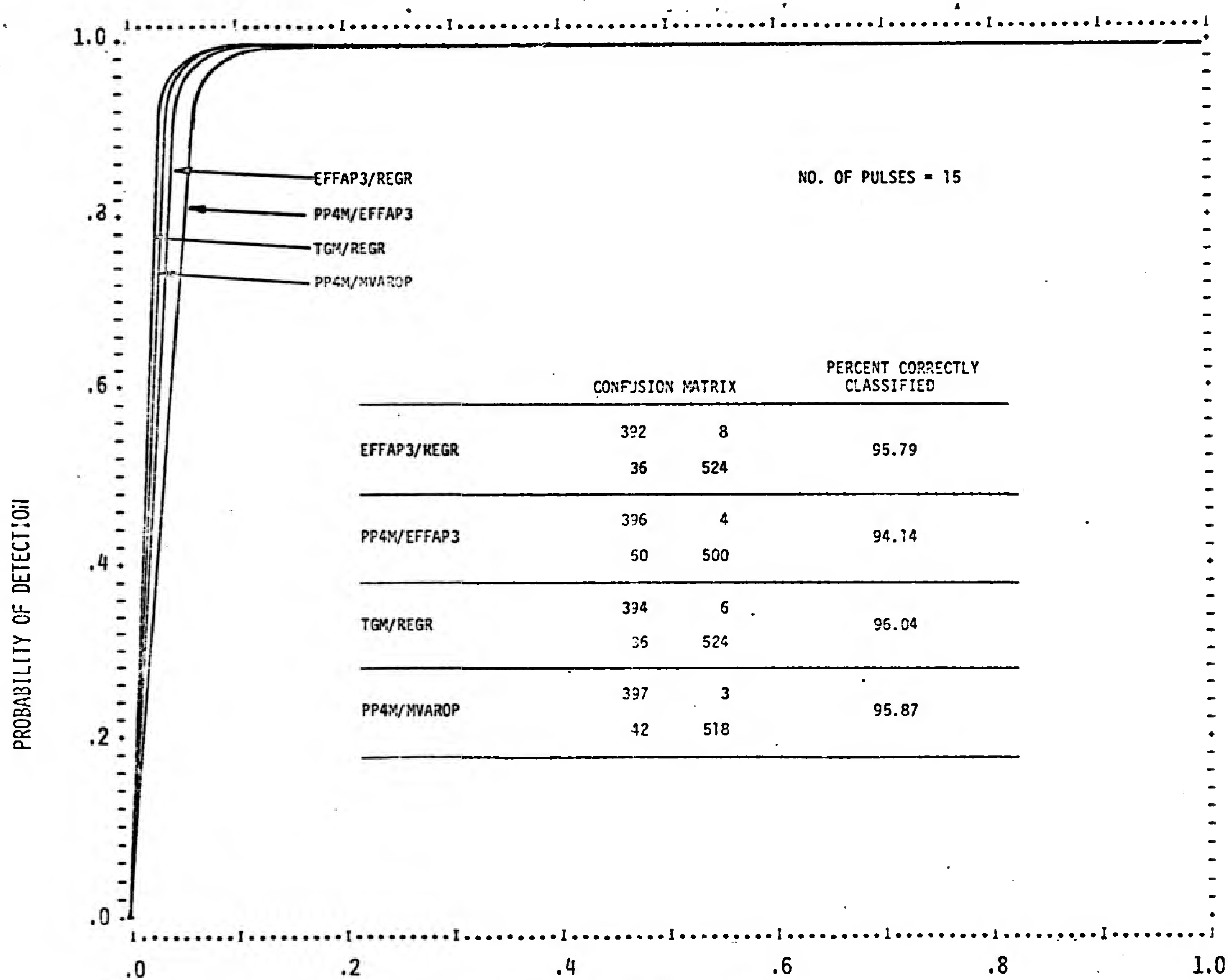


Figure 5-3c



	CONFUSION MATRIX		PERCENT CORRECTLY CLASSIFIED
EFFAP3/REGR	392	8	95.79
	36	524	
PP4M/EFFAP3	396	4	94.14
	50	500	
TGM/REGR	394	6	96.04
	36	524	
PP4M/MVAROP	397	3	95.87
	42	518	

Figure 5-4 Performance of pairs of features.

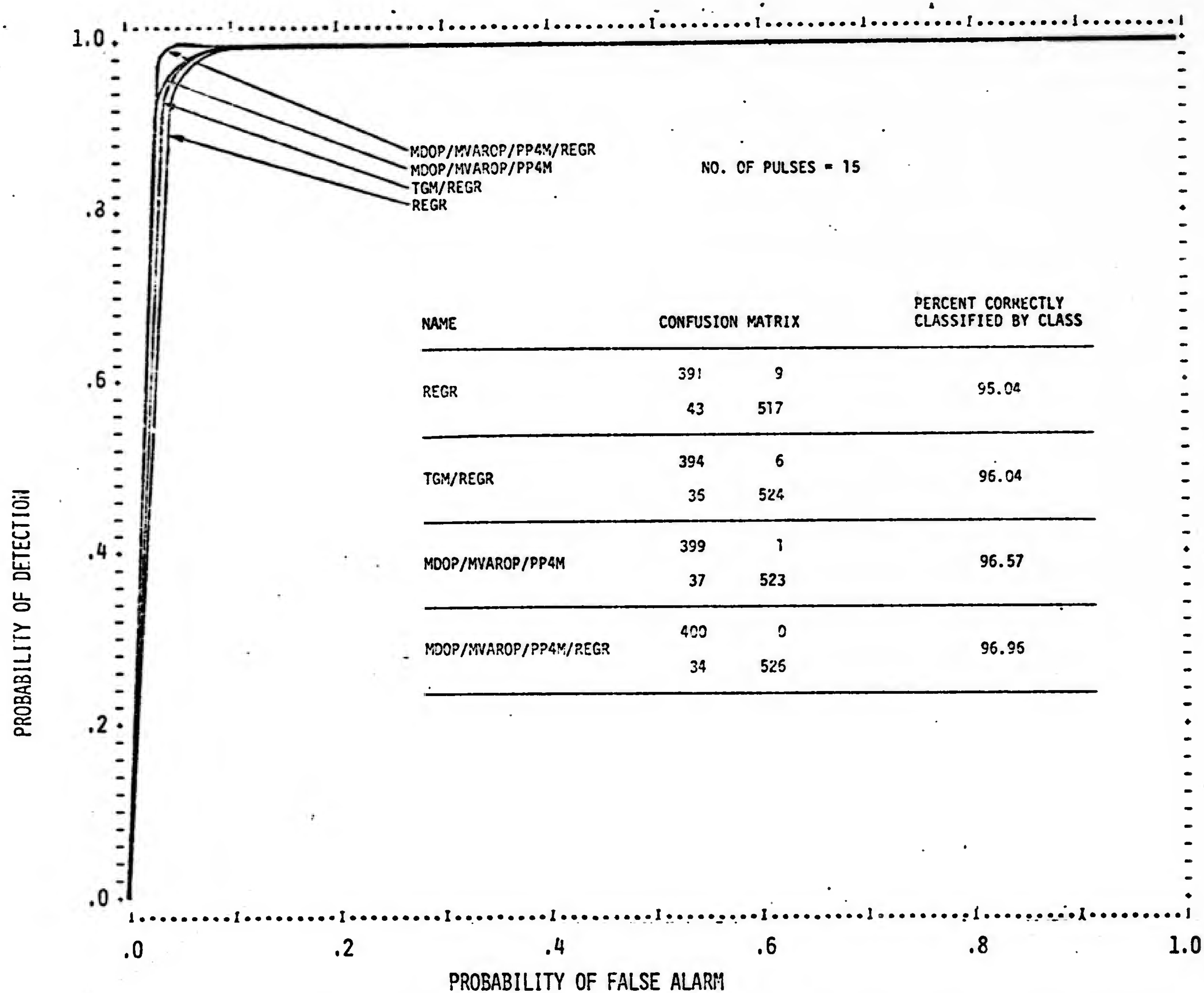
Triples are ranked in Table 5-5 and quadruples in Table 5-6. Figure 5-7 compares the performance of the selected single feature, pair, triple, and quadruple. The selection of the triple was not entirely based on the rank as the top 5 sets had similar values of percent correctly classified. Note that the quadruple achieves a 1.00 probability of detection for a probability of false alarm of about .05.

Feature Triples			Percent Correct Classification	Rank
TGM	REGR	EFFAP3	96.66	1
MVAROP	PP4M	EFFAP3	96.51	2
PP4M	EFFAP3	REGR	96.45	3
MVAROP	PP4M	MDOP	96.40	4
MVAROP	PP4M	FLOPSK	96.36	5
TGM	REGR	MDOP	96.30	6
TGM	REGR	XANGL	96.25	7
MVAROP	PP4M	XANGL	96.22	8
TGM	REGR	FLOPSK	96.07	9
MVAROP	PP4M	REGR	95.91	10
MVAROP	PP4M	TGM	95.91	11
PP4M	EFFAP3	FLOPSK	95.79	12
TGM	REGR	OPRAV	95.75	13
PP4M	EFFAP3	XANGL	95.73	14
TGM	REGR	MVAROP	95.32	15
PP4M	EFFAP3	MDOP	94.96	16
PP4M	EFFAP3	TGM	94.80	17
PP4M	EFFAP3	OPRAV	93.96	18

Table 5-5: Feature Triples Ranked at 15 Pulses (N)

Feature Quadruples				Percent Correct Classification	Rank
MDOP	MVAROP	PP4M	REGR	96.96	1
MVAROP	PP4M	EFFAP3	MDOP	96.91	2
MVAROP	PP4M	EFFAP3	TGM	96.84	3
MVAROP	PP4M	EFFAP3	REGR	96.67	4
MDOP	MVAROP	PP4M	TGM	96.42	5

Table 5-6: Feature Quadruples Ranked at 15 Pulses (N)



NAME	CONFUSION MATRIX		PERCENT CORRECTLY CLASSIFIED BY CLASS
REGR	391 43	9 517	95.04
TGM/REGR	394 35	6 524	96.04
MDOP/MVAROP/PP4M	399 37	1 523	96.57
MDOP/MVAROP/PP4M/REGR	400 34	0 526	96.95

Figure 5-5 : Comparison of the selected single feature, pair, triple, and quadruple.

6.0 VALIDATION OF PERFORMANCE ESTIMATES

The methodology utilized in comparing the performance of features singly and in groups is designed for the stability of the results upon extension to data not used in the creation of the decision rule. It is the purpose of this section to validate that this objective was met.

We shall refer to the set of signatures used to design the discrimination rule and used in obtaining the results of previous sections as the "design set" and to the set of target and non-target signatures set aside for an independent test as the "test set." As discussed in Section 4.1.2 the decision algorithm employed for feature performance evaluation involved use of a Bayes rule and an estimation of the class conditional probability densities from the design set of samples (via a use of the normal form and an estimation of the mean vector and covariance matrix.) The decision rule so derived was used unchanged on the test samples to determine the validity and stability of the performance estimates. It should be emphasized that the test set is used to evaluate features and discrimination algorithms which were obtained using the design set; the test set is not used in the creation of the decision rule which it tests.

Figures 6-1 through 6-8 are comparisons of the characteristics obtained on the design and test set for several of the single features, pairs of features, a triple of features, and a quadruple of features. In some cases, the performance on the test set was <u>better</u> than on the design set; in all cases, the difference in performance was easily attributable to normal statistical fluctuations. Any overestimation of the performance of the features was small.

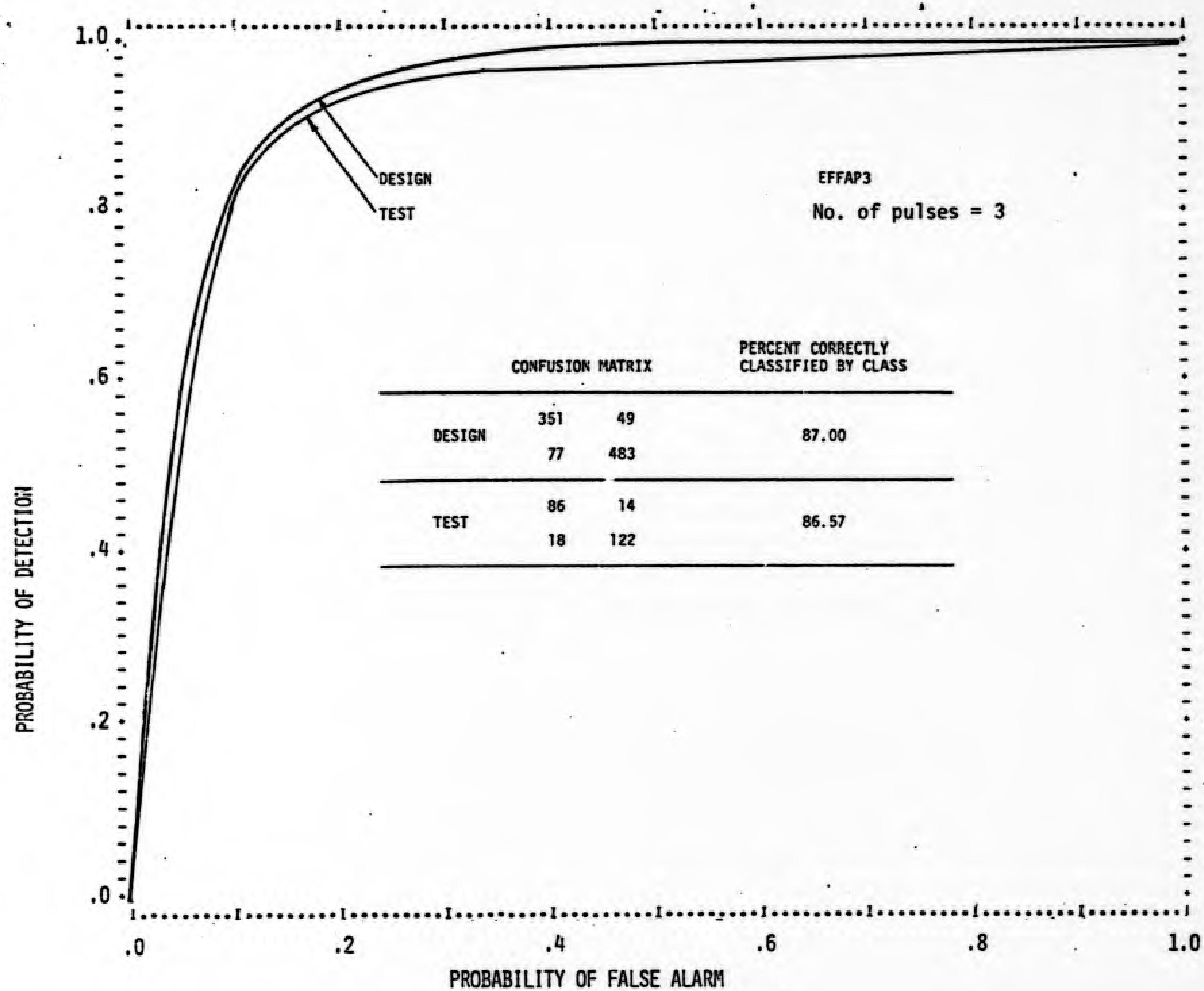

1.0
.8
.6
.4
.2
.0
PROBABILITY OF DETECTION
.0
.2
.4
.6
.8
1.0
PROBABILITY OF FALSE ALARM
DESIGN
TEST
EFFAP3
No. of pulses = 3
CONFUSION MATRIX
PERCENT CORRECTLY CLASSIFIED BY CLASS
DESIGN
351 49
77 483
87.00
TEST
86 14
18 122
86.57

Figure 6-1

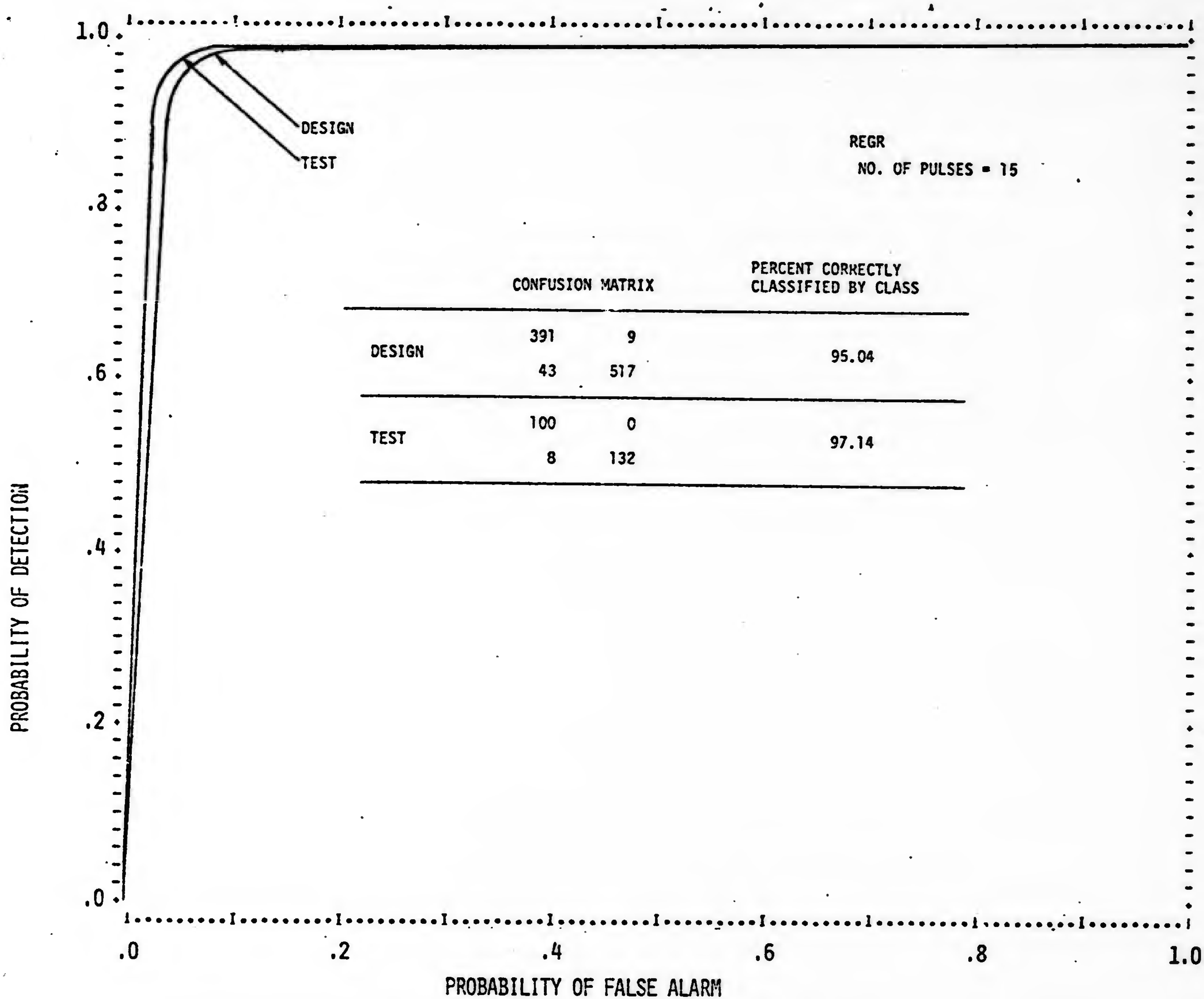

DESIGN
TEST
REGR
NO. OF PULSES = 15
CONFUSION MATRIX
PERCENT CORRECTLY CLASSIFIED BY CLASS
DESIGN
391 9
43 517
95.04
TEST
100 0
8 132
97.14
PROBABILITY OF DETECTION
PROBABILITY OF FALSE ALARM
1.0
.8
.6
.4
.2
.0
.0
.2
.4
.6
.8
1.0

Figure 6-2

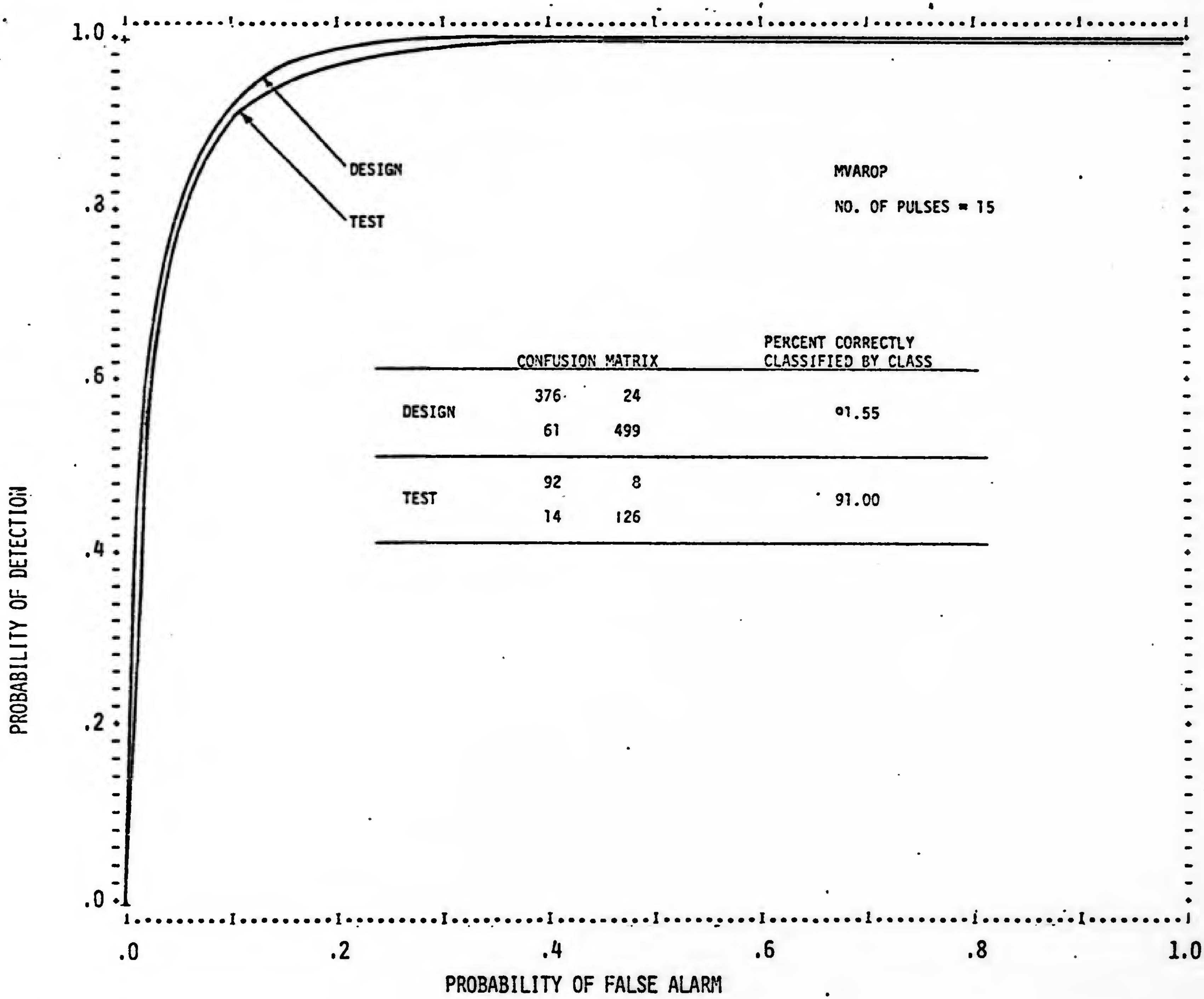

MVAROP
NO. OF PULSES = 15
DESIGN
TEST
CONFUSION MATRIX
PERCENT CORRECTLY CLASSIFIED BY CLASS
DESIGN
376 24
61 499
91.55
TEST
92 8
14 126
91.00
PROBABILITY OF DETECTION
PROBABILITY OF FALSE ALARM
1.0
.8
.6
.4
.2
.0

Figure 6-3

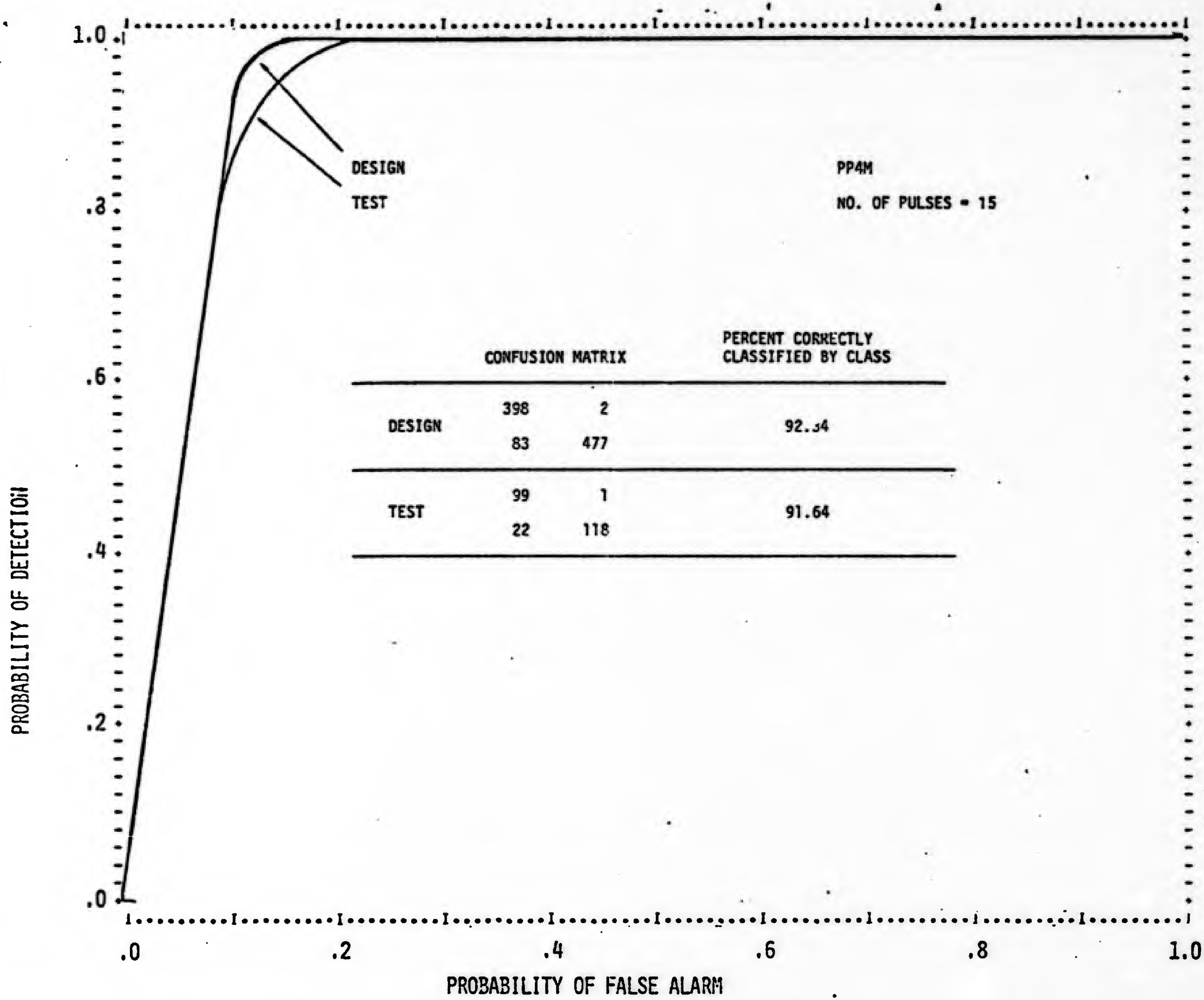

1.0
.8
.6
.4
.2
.0
PROBABILITY OF DETECTION
.0
.2
.4
.6
.8
1.0
PROBABILITY OF FALSE ALARM
DESIGN
TEST
PP4M
NO. OF PULSES = 15
CONFUSION MATRIX
PERCENT CORRECTLY CLASSIFIED BY CLASS
DESIGN
398 2
83 477
92.54
TEST
99 1
22 118
91.64

Figure 6-4

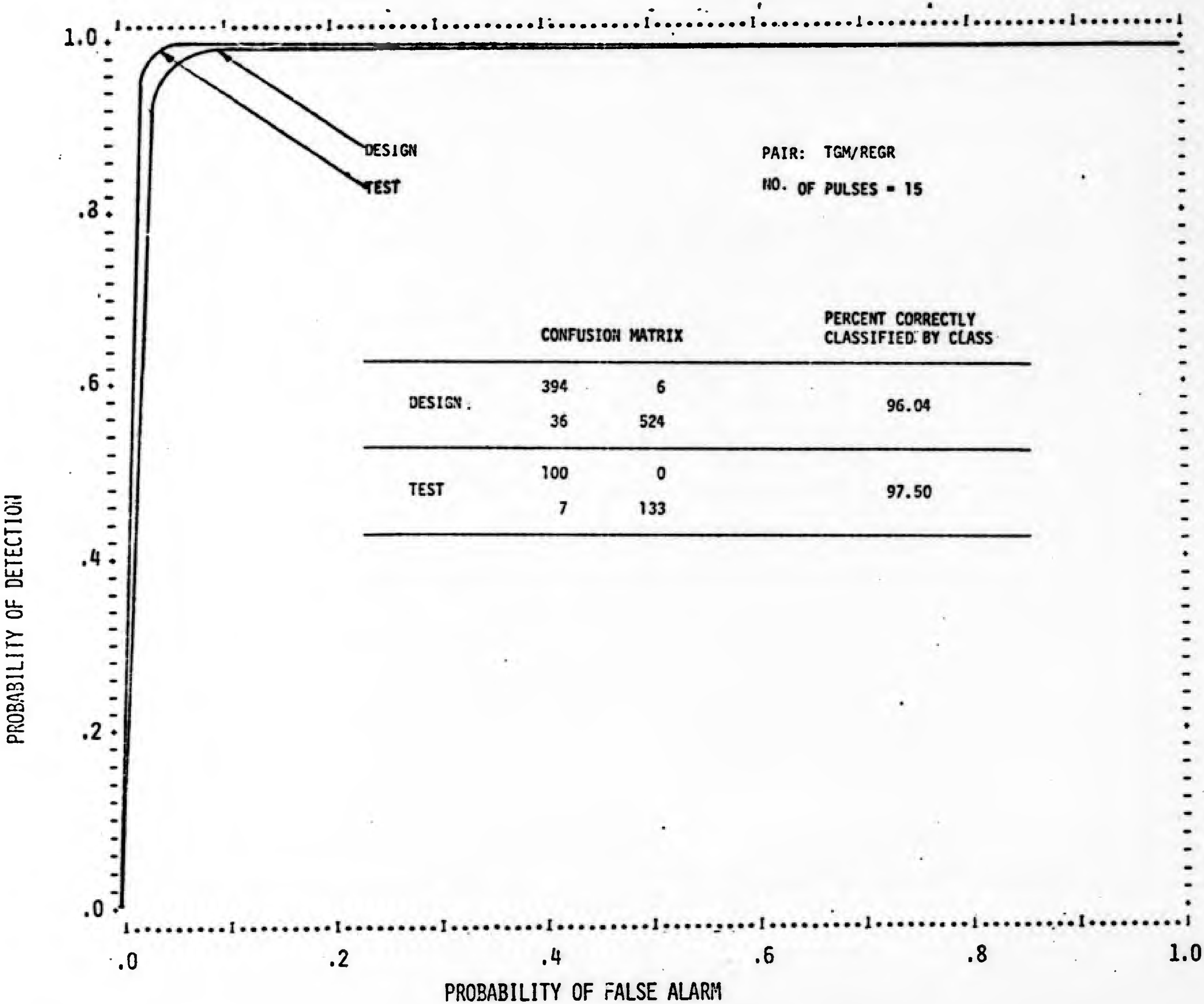

1.0
.8
.6
.4
.2
.0
PROBABILITY OF DETECTION
DESIGN
TEST
PAIR: TGM/REGR
NO. OF PULSES = 15
CONFUSION MATRIX
PERCENT CORRECTLY CLASSIFIED BY CLASS
DESIGN
394 6
36 524
96.04
TEST
100 0
7 133
97.50
.0
.2
.4
.6
.8
1.0
PROBABILITY OF FALSE ALARM

Figure 6-5

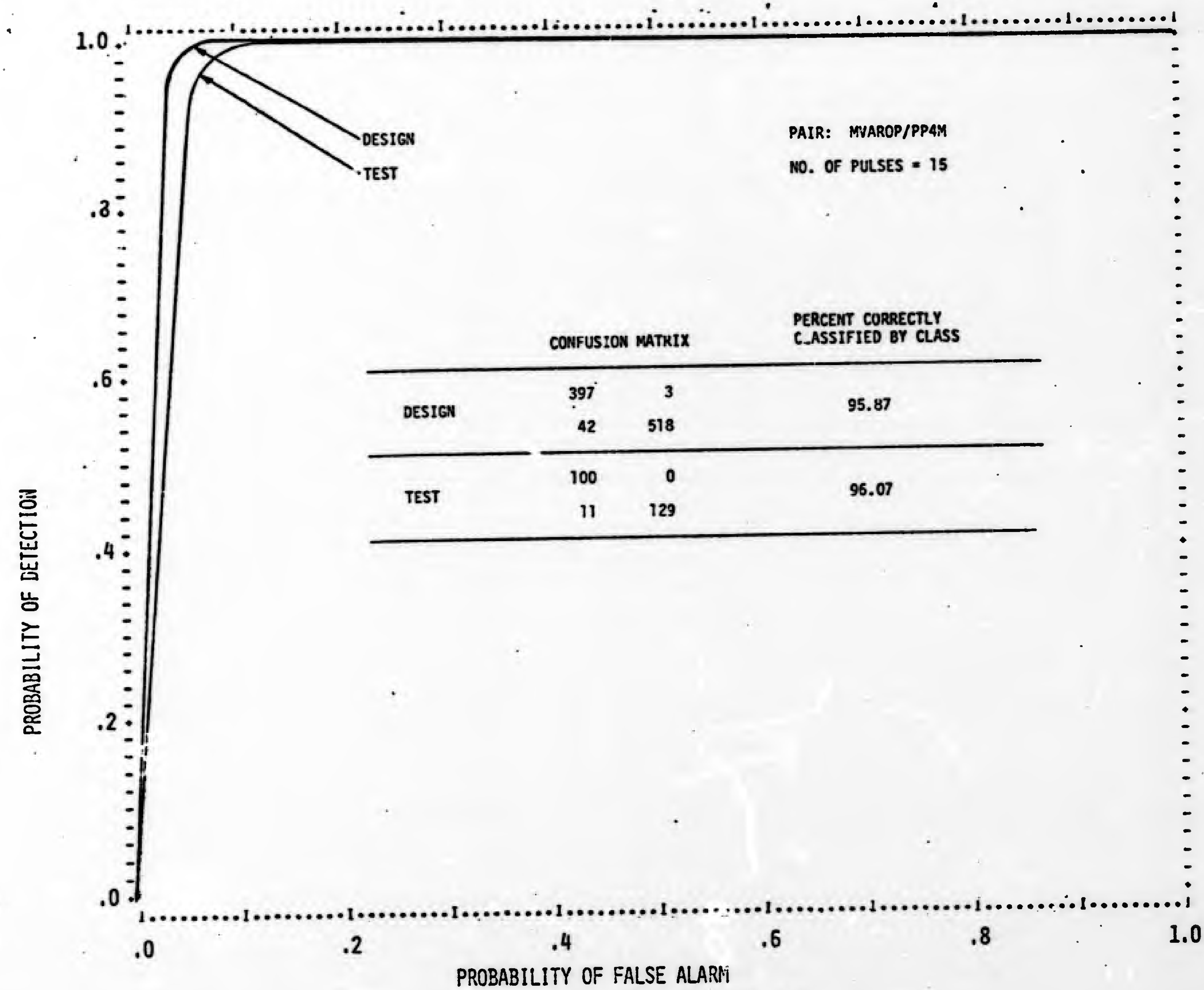
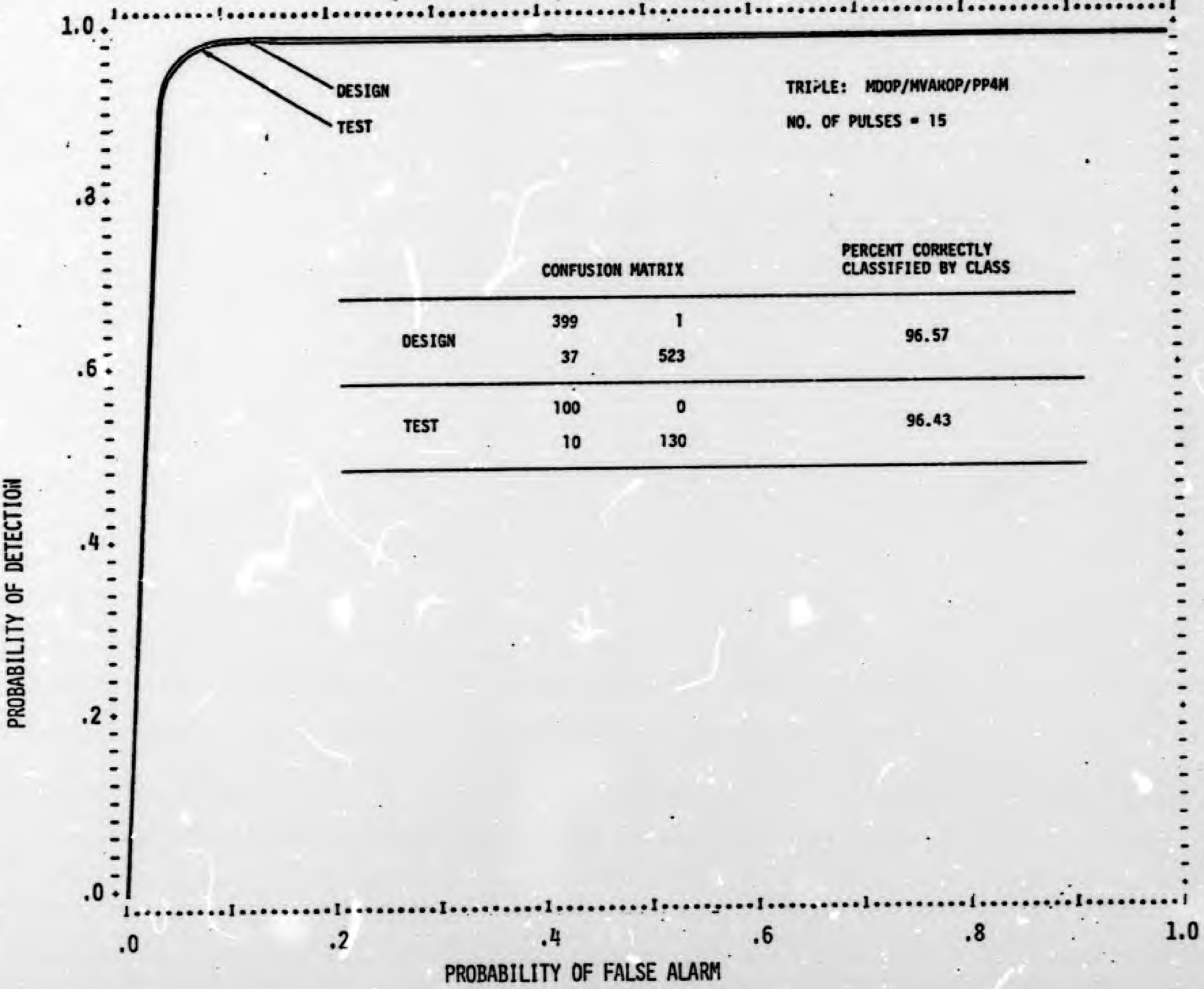

1.0
.8
.6
.4
.2
.0
PROBABILITY OF DETECTION
.0
.2
.4
.6
.8
1.0
PROBABILITY OF FALSE ALARM
DESIGN
TEST
PAIR: MVAROP/PP4M
NO. OF PULSES = 15
CONFUSION MATRIX
PERCENT CORRECTLY CLASSIFIED BY CLASS
DESIGN 397 3 / 42 518 95.87
TEST 100 0 / 11 129 96.07

Figure 6-6



	CONFUSION MATRIX		PERCENT CORRECTLY CLASSIFIED BY CLASS
DESIGN	399	1	96.57
	37	523	
TEST	100	0	96.43
	10	130	

Figure 6-7

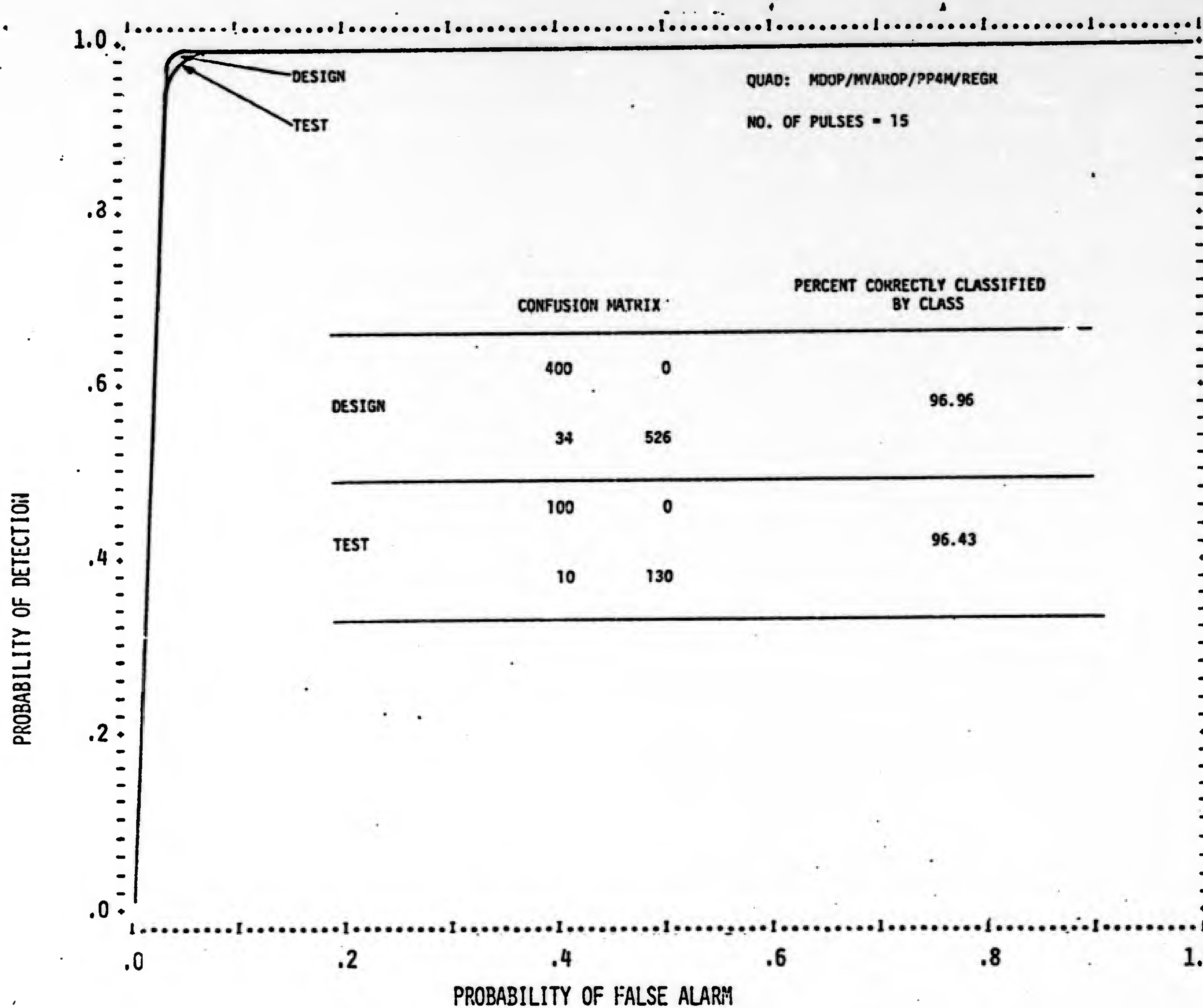

1.0
.8
.6
.4
.2
.0
PROBABILITY OF DETECTION
.0
.2
.4
.6
.8
1.0
PROBABILITY OF FALSE ALARM
DESIGN
TEST
QUAD: MDOP/MVAROP/PP4M/REGR
NO. OF PULSES = 15
CONFUSION MATRIX
PERCENT CORRECTLY CLASSIFIED BY CLASS
DESIGN
400 0
34 526
96.96
TEST
100 0
10 130
96.43

Figure 6-8

7.0 IMPACT OF THE NUMBER OF PULSES ON DISCRIMINATION

The analysis and feature ranking of the previous sections have largely emphasized the case of 15 pulses. One would certainly expect the discrimination performance to be lower with fewer pulses. Since, with few exceptions, the heuristic features are defined such as to be applicable to any number of pulses (e.g., one can average the RCS's of three pulses as well as fifteen pulses), we can study the same features at varying numbers of pulses (hereafter N). This question is of obvious importance in a system context where the number of pulses available for discrimination is a limited resource.

The degradation of performance with the number of pulses is illustrated for the single feature PP4M in figures 7-1 and 7-2. Figure 7-1 shows the change in form of a non-parametric estimate of the class conditional probability densities for changing numbers of pulses (N=3,6,15,200); figure 7-2 shows performance measures corresponding to these changes in the density. Figures 7-3 and 7-4 provide the same information for MVAROP; figures 7-5 and 7-6, for REGR. Changes in the densities for several other single features as the number of pulses (N) increased are plotted in figures 7-7 through 7-15 for whatever insight they may provide.

Turning to pairs of features, we find that looking at overlap of the probability densities is more difficult. Figures 7-16 and 7-17 provide a pictorial representation of how the overlap of bivariate densities decreases (using a simple histogram estimate) in going from N=15 to N=200; in these figures, the high peak is the target class. Since it is

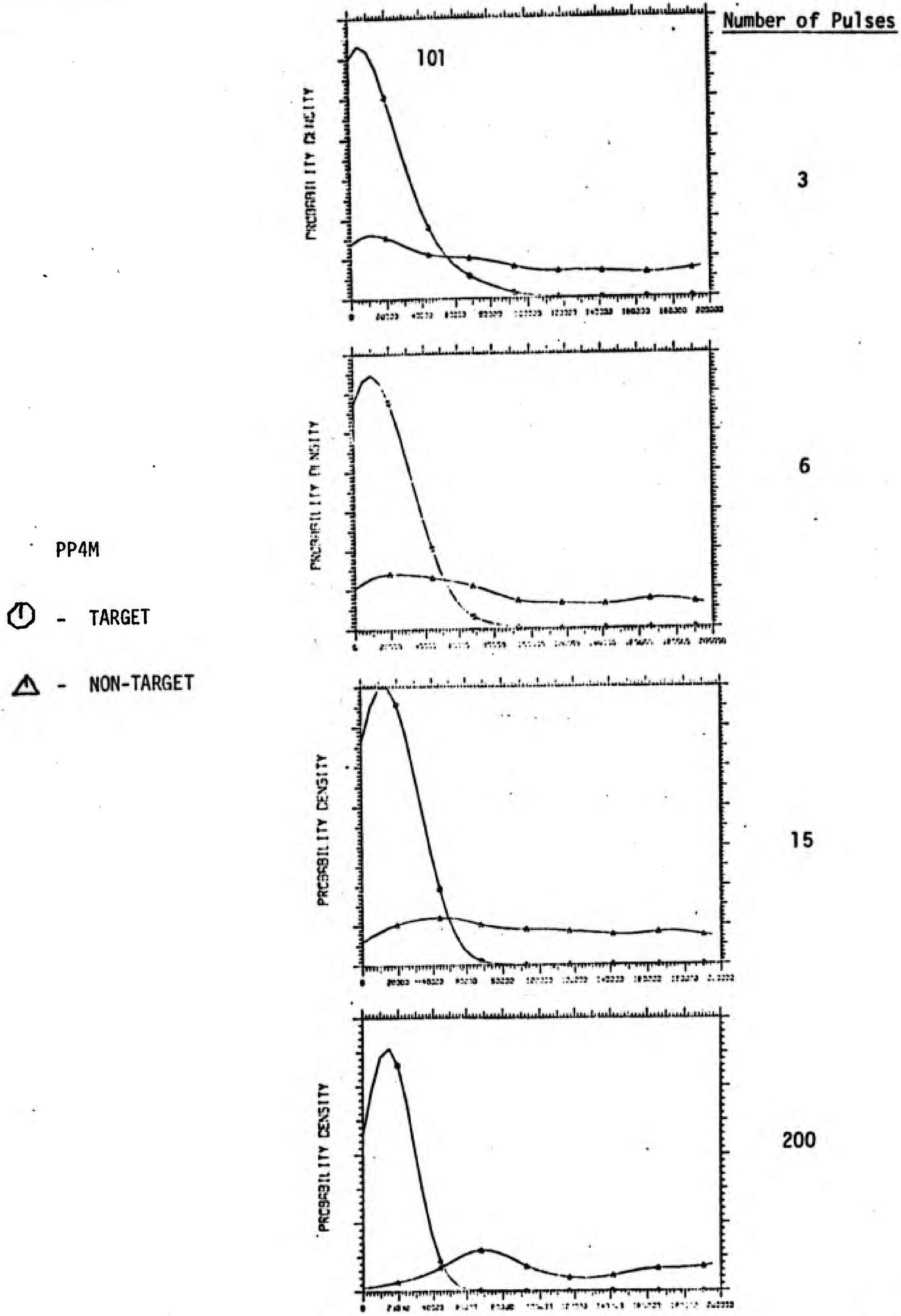
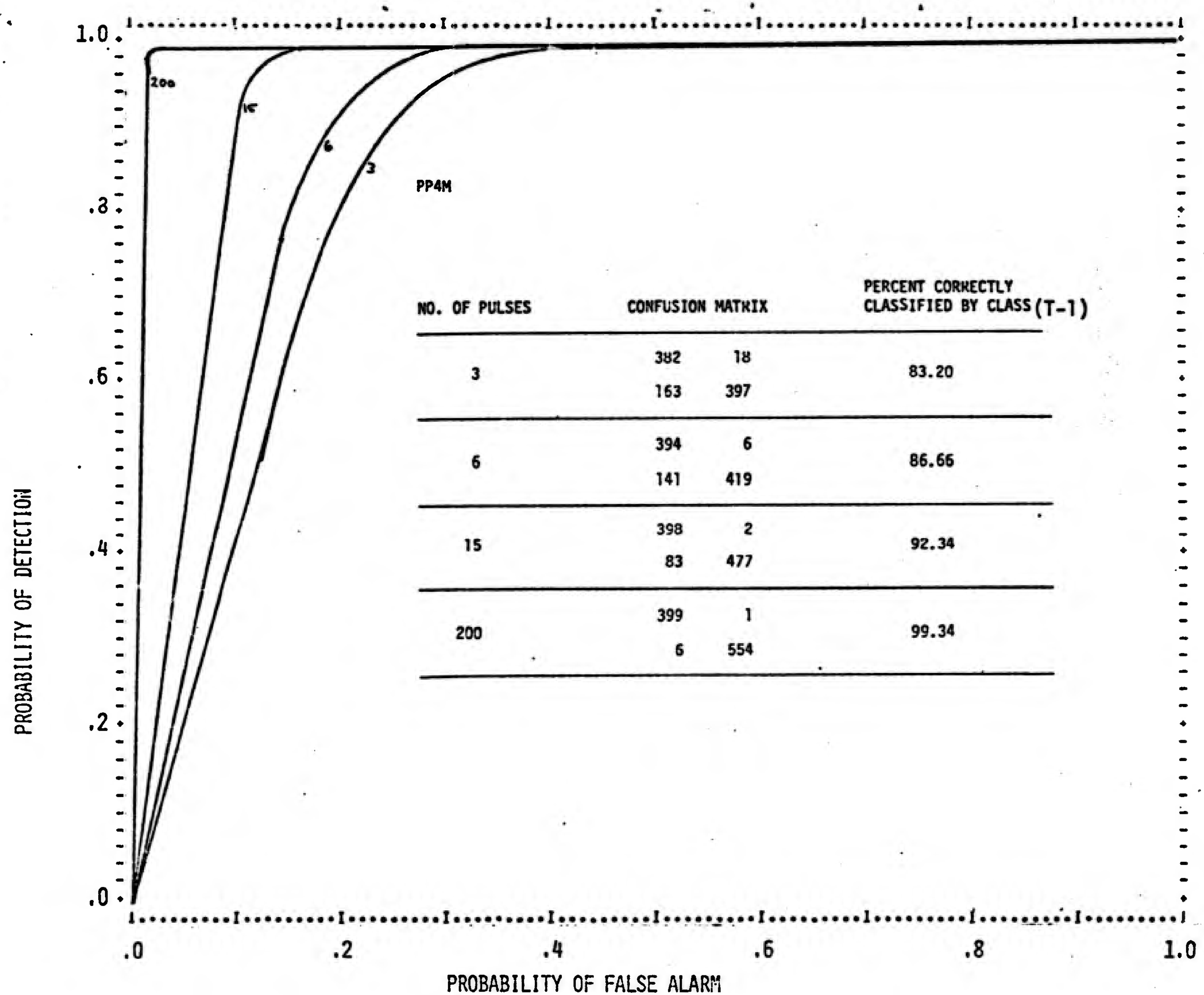


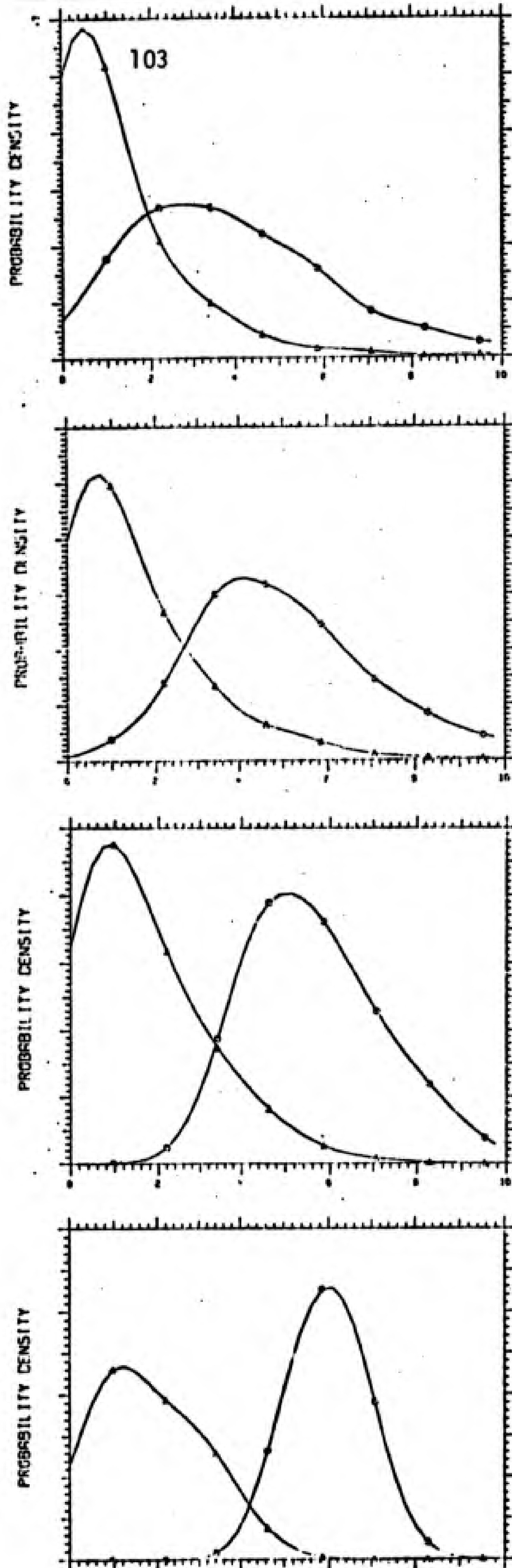
Figure 7-1



NO. OF PULSES	CONFUSION MATRIX		PERCENT CORRECTLY CLASSIFIED BY CLASS (T-1)
3	382	18	83.20
	163	397	
6	394	6	86.66
	141	419	
15	398	2	92.34
	83	477	
200	399	1	99.34
	6	554	

Figure 7-2

Number of Pulses



3

6

15

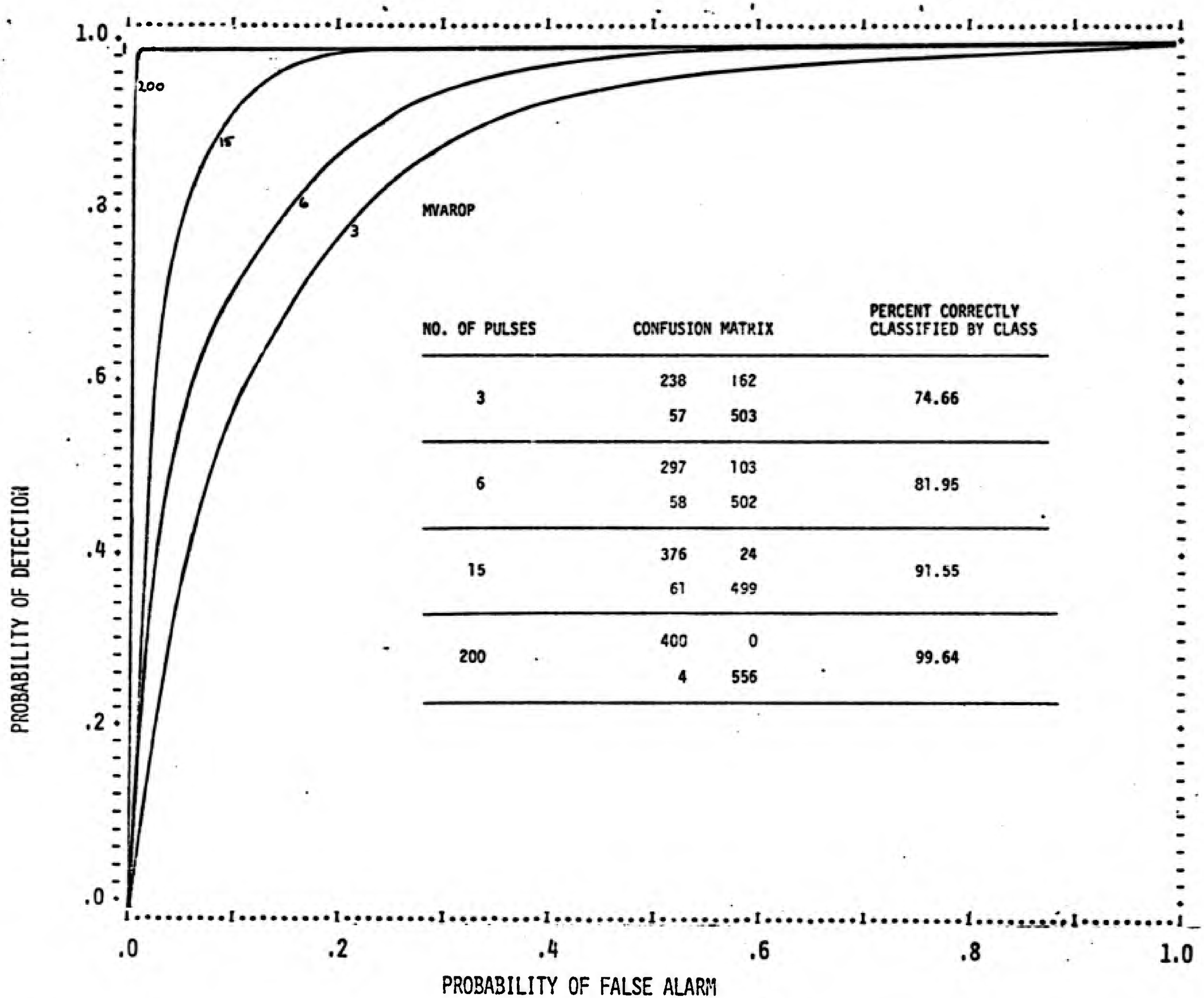
200

MVAROP

○ - TARGET

△ - NON-TARGET

Figure 7-3



NO. OF PULSES	CONFUSION MATRIX		PERCENT CORRECTLY CLASSIFIED BY CLASS
3	238	162	74.66
	57	503	
6	297	103	81.95
	58	502	
15	376	24	91.55
	61	499	
200	400	0	99.64
	4	556	

Figure 7-4

Number of Pulses

REGR

⏀ - TARGET

△ - NON-TARGET

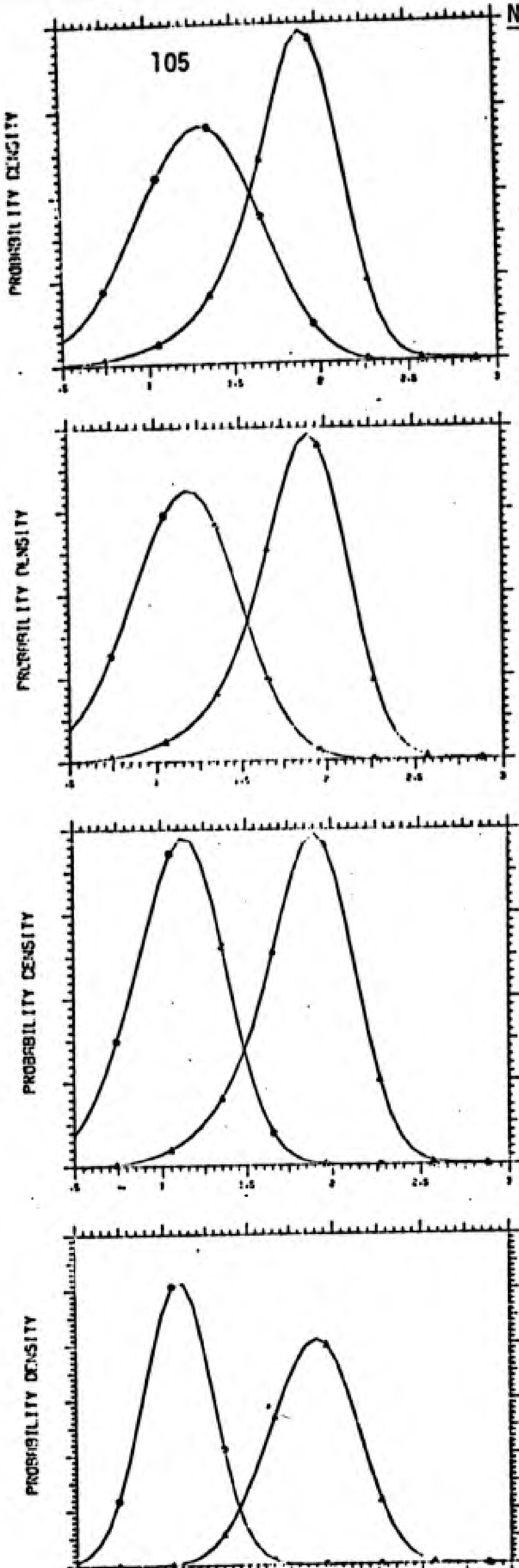
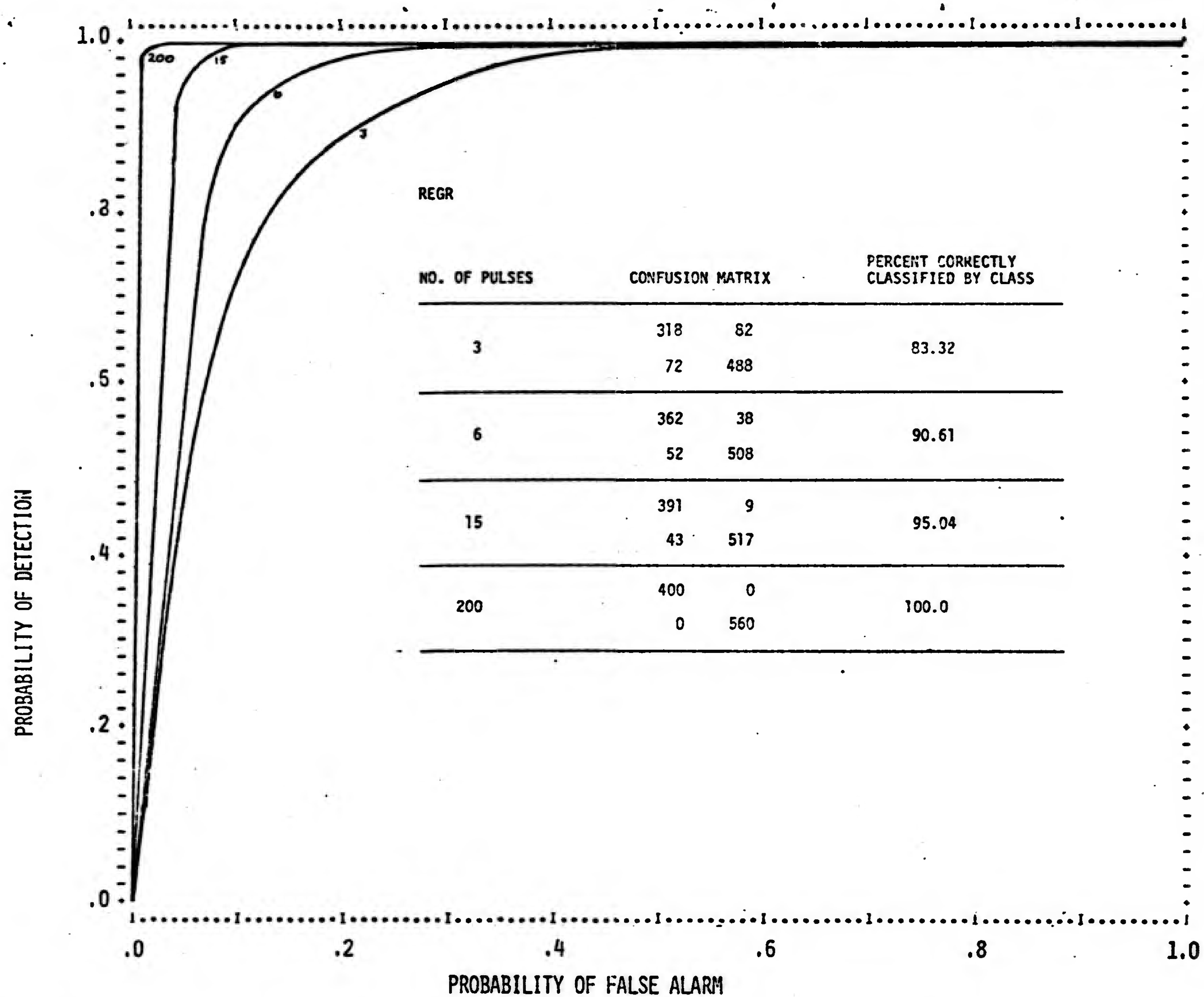


Figure 7-5



NO. OF PULSES	CONFUSION MATRIX		PERCENT CORRECTLY CLASSIFIED BY CLASS
3	318	82	83.32
	72	488	
6	362	38	90.61
	52	508	
15	391	9	95.04
	43	517	
200	400	0	100.0
	0	560	

Figure 7-6

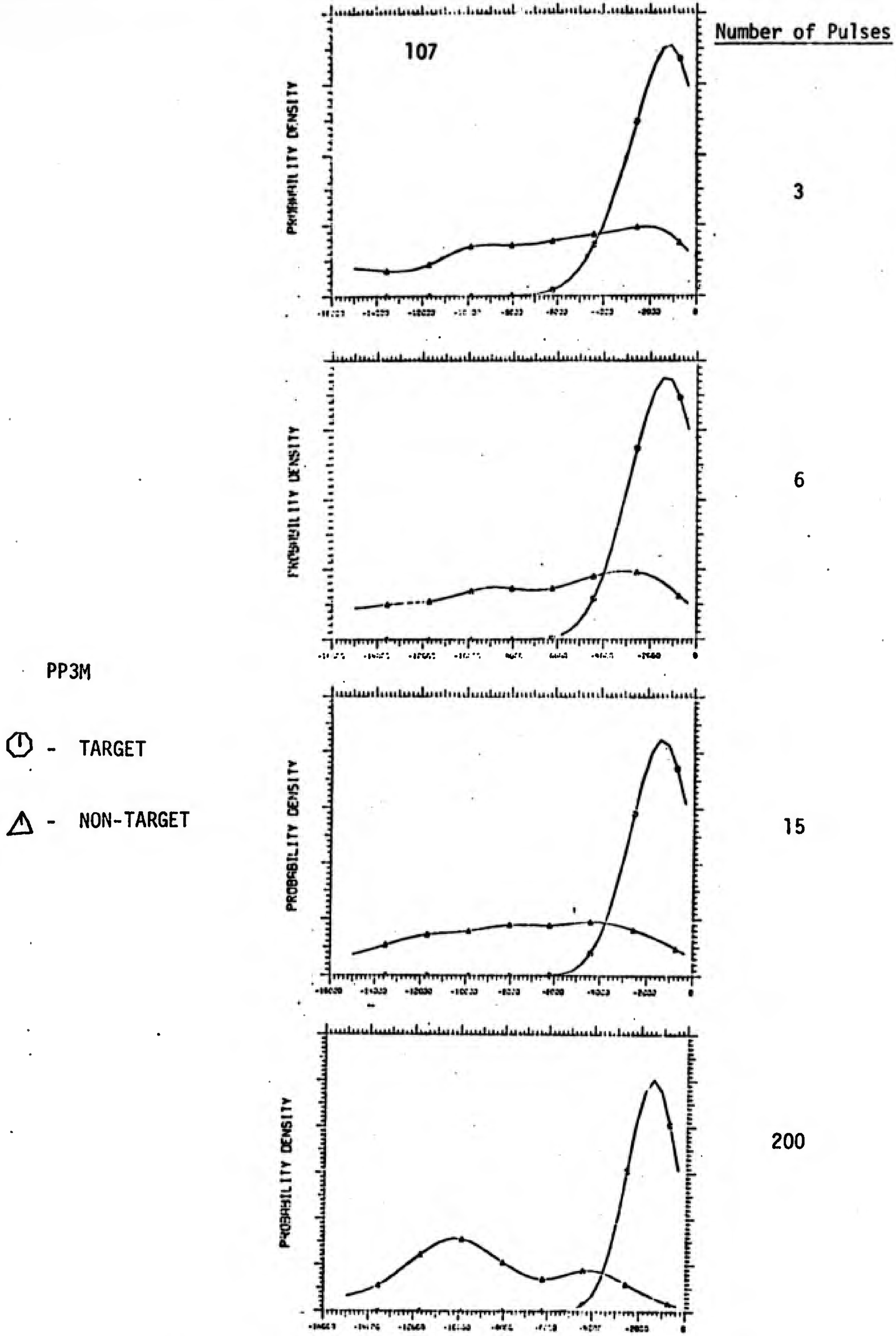

Number of Pulses
107
PROBABILITY DENSITY
3
6
15
200
PP3M
TARGET
NON-TARGET

Figure 7-7

PGM

Ⓣ - TARGET

△ - NON-TARGET

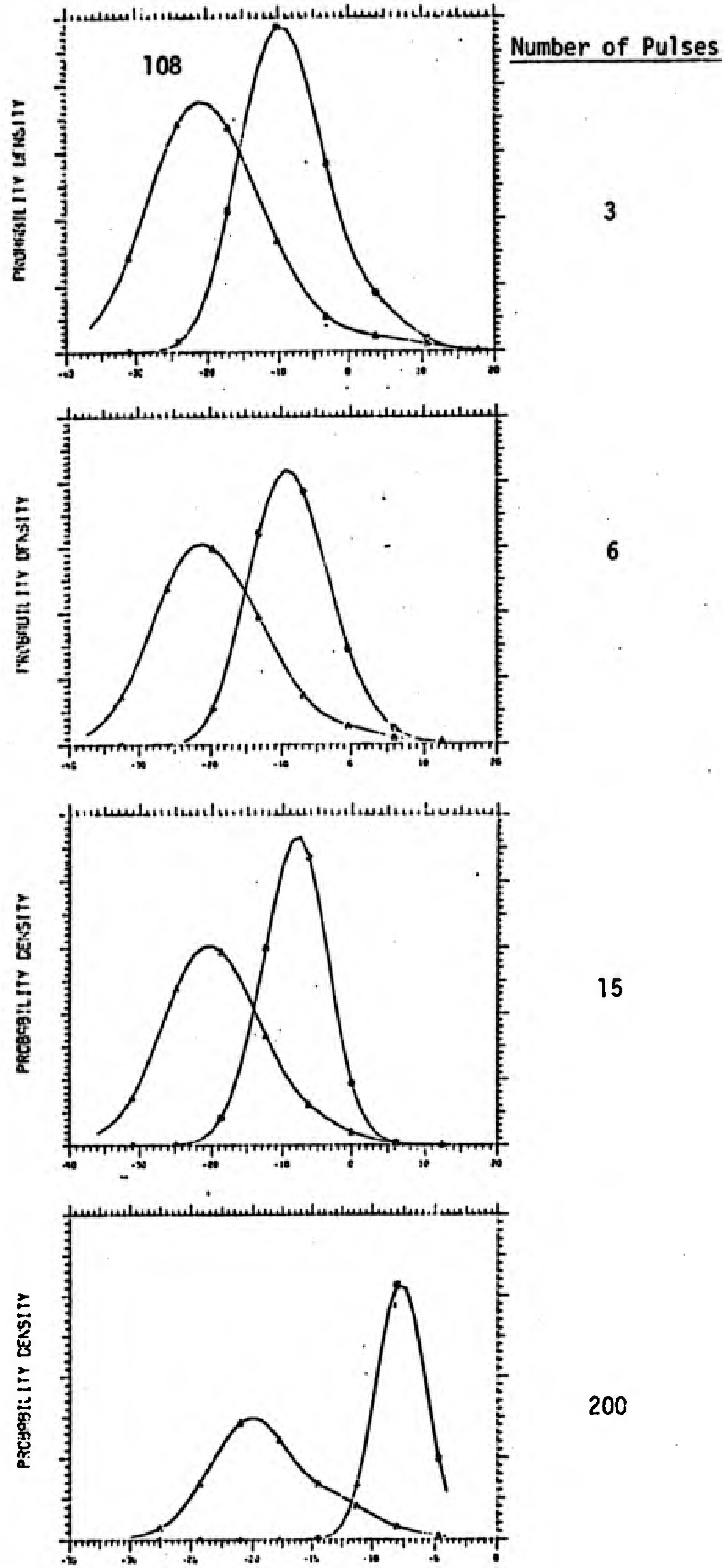


Figure 7-8

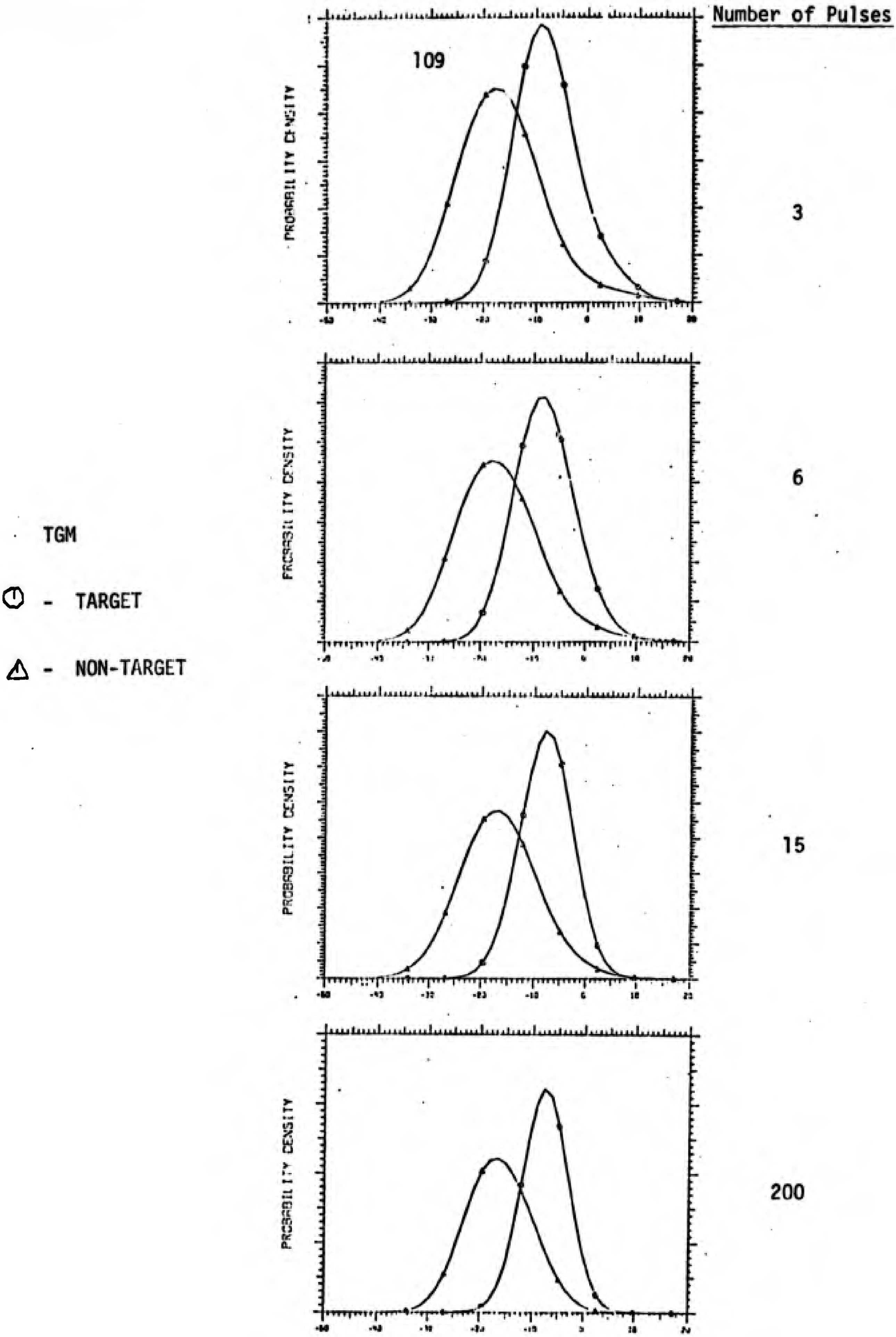

Number of Pulses
109
PROBABILITY DENSITY
3
6
15
200
TGM
- TARGET
- NON-TARGET

Figure 7-9

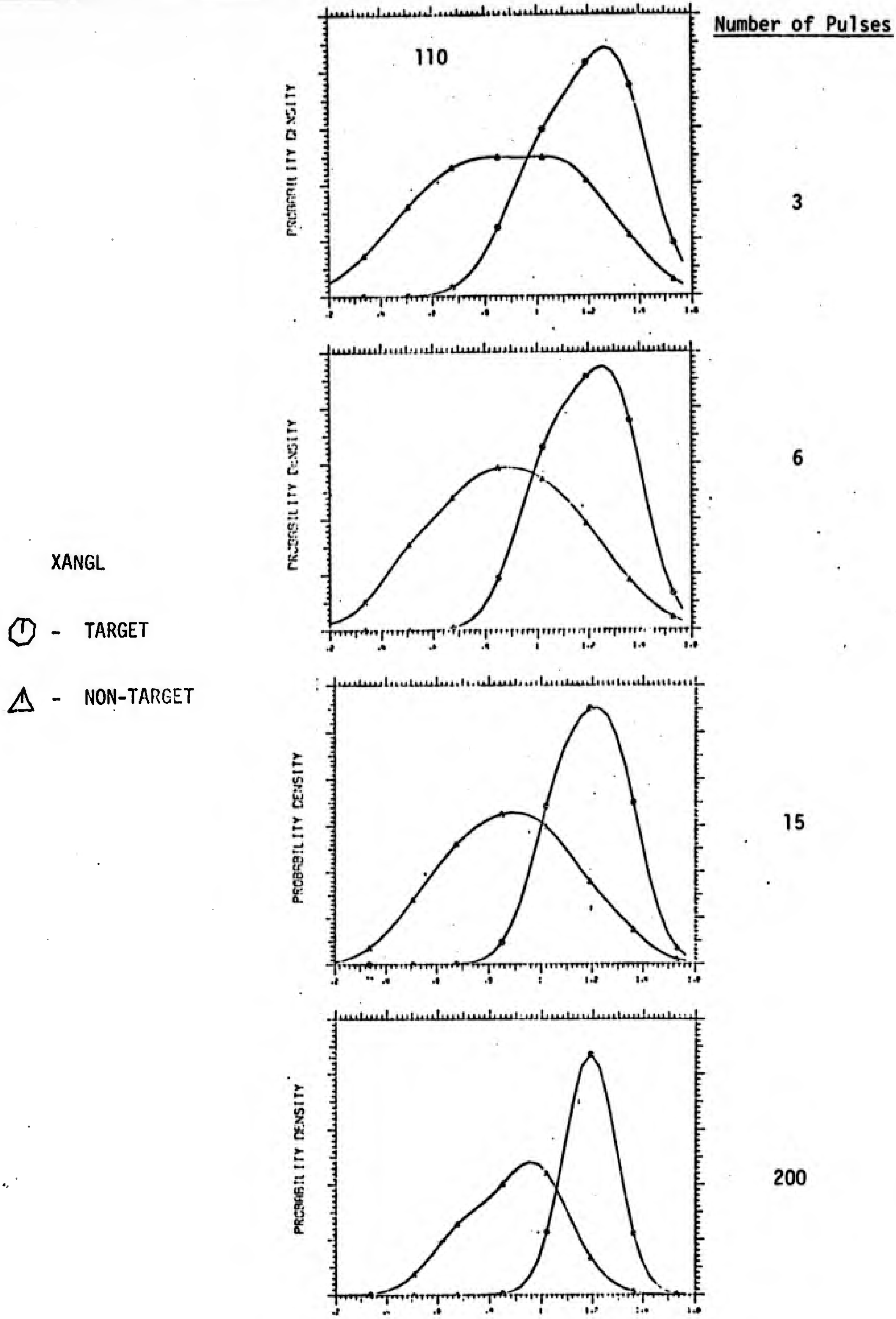
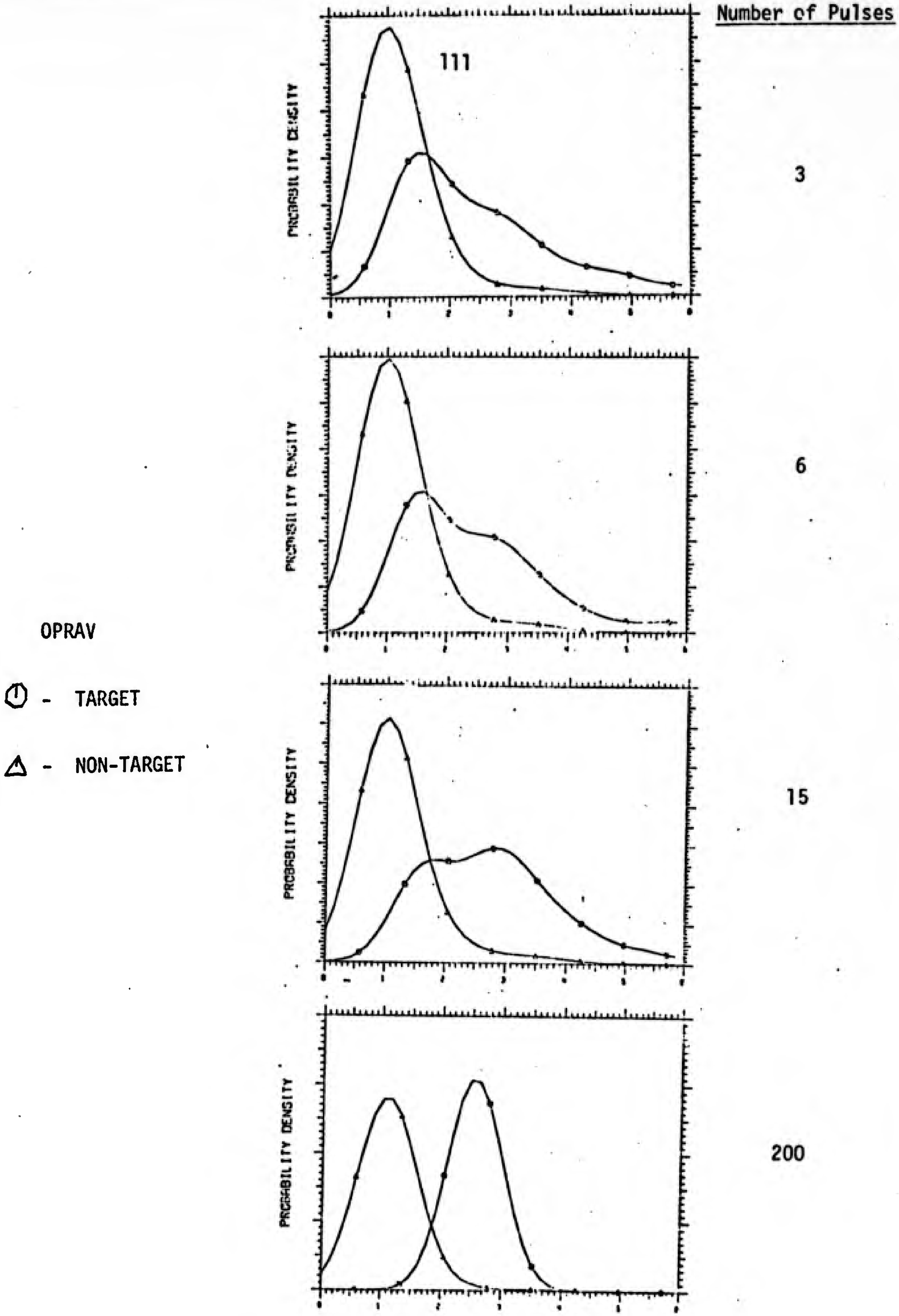

Number of Pulses
110
PROBABILITY DENSITY
3
6
15
200
XANGL
- TARGET
- NON-TARGET

Figure 7-10



Number of Pulses
111
PROBABILITY DENSITY
3
6
15
200
OPRAV
- TARGET
- NON-TARGET

Figure 7-11

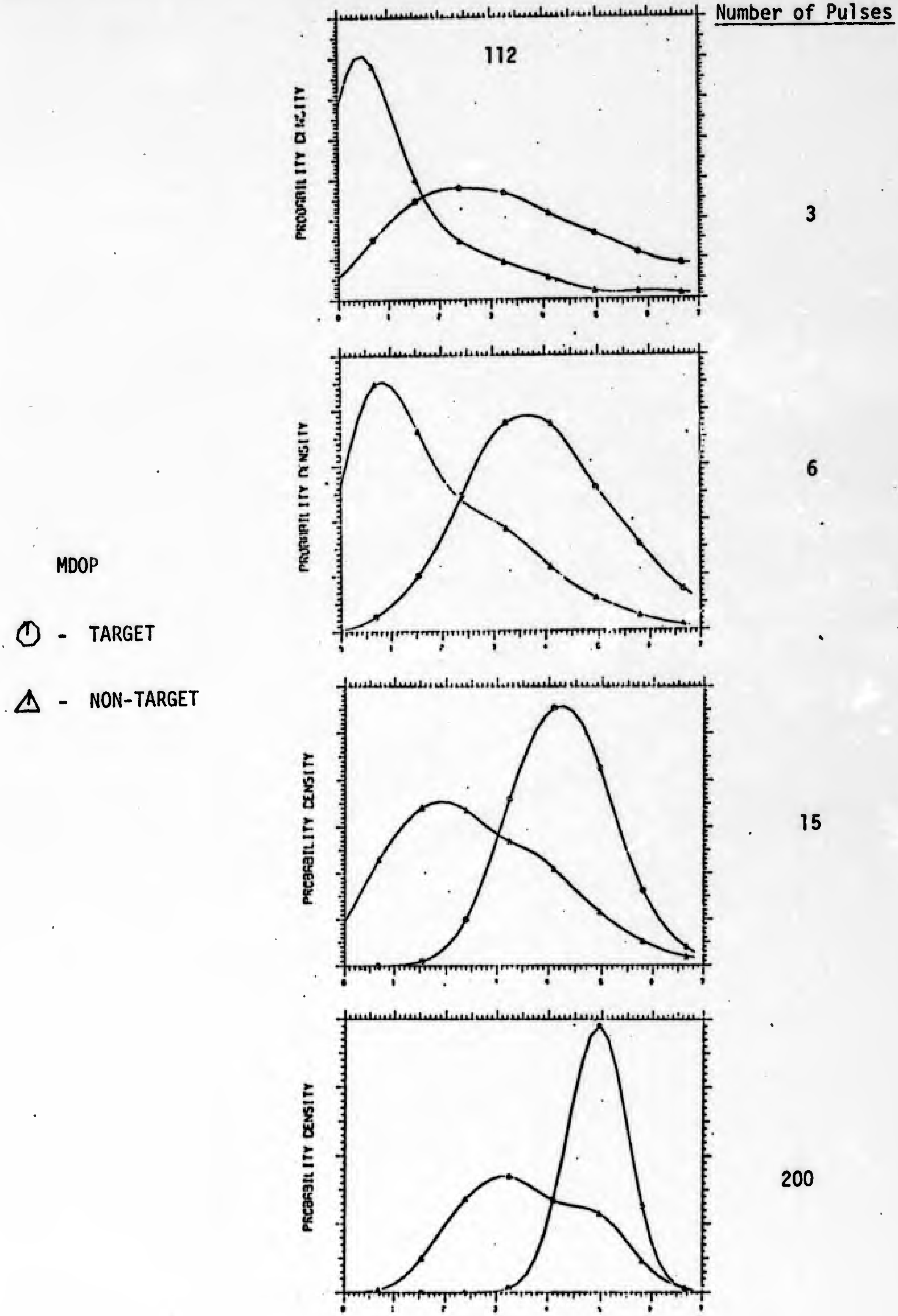

Number of Pulses
112
3
6
15
200
PROBABILITY DENSITY
MDOP
- TARGET
- NON-TARGET

Figure 7-12

Number of Pulses

VANGLE

(T) - TARGET

△ - NON-TARGET

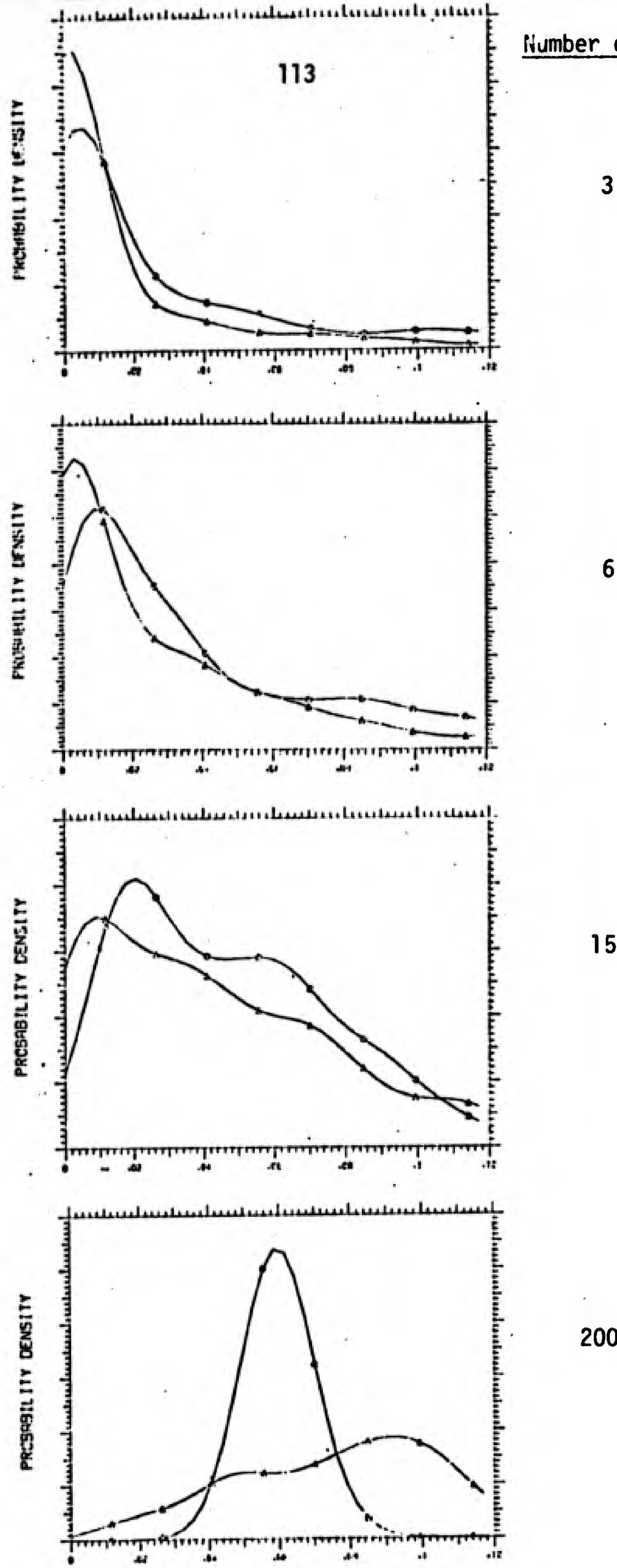


Figure 7-13

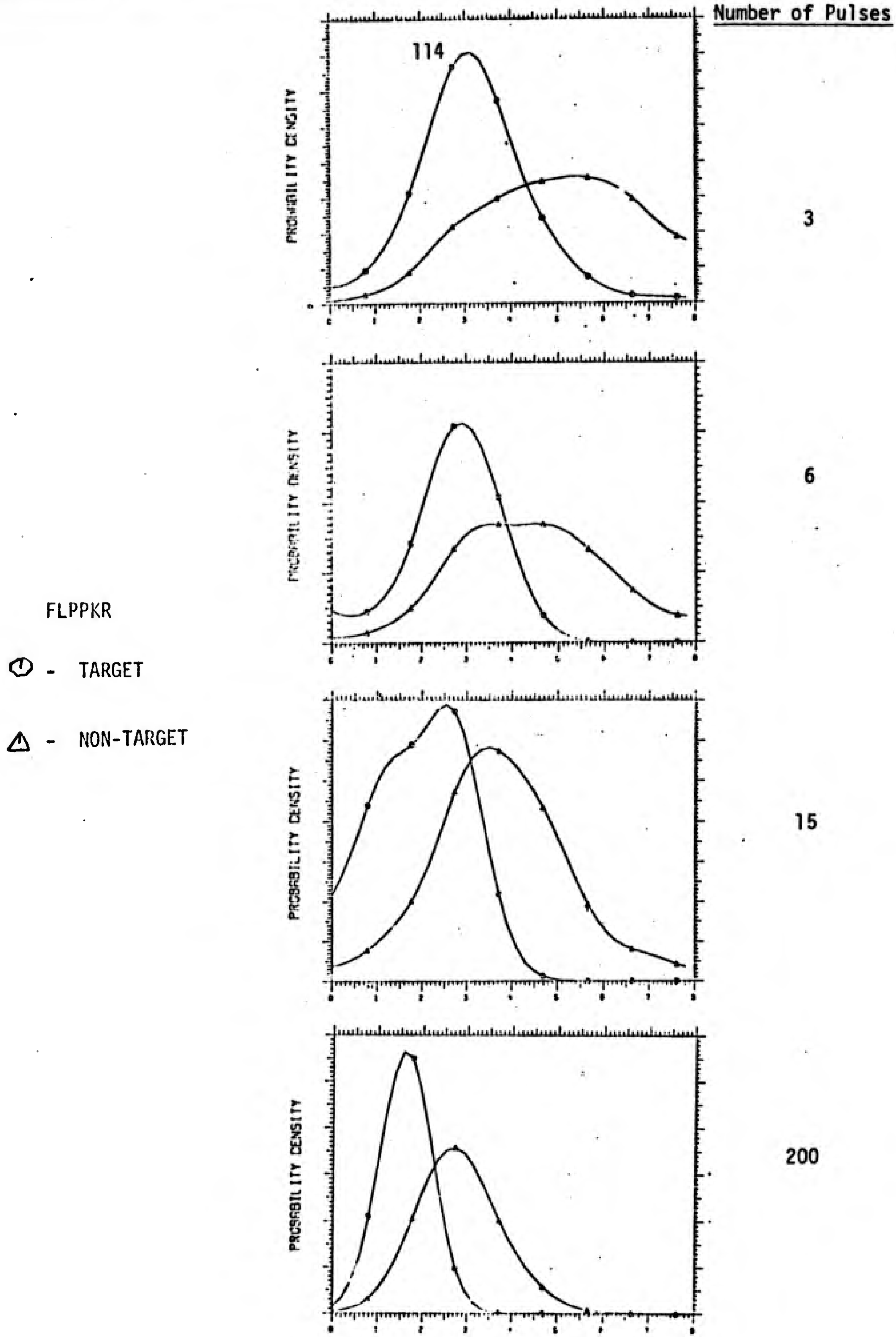
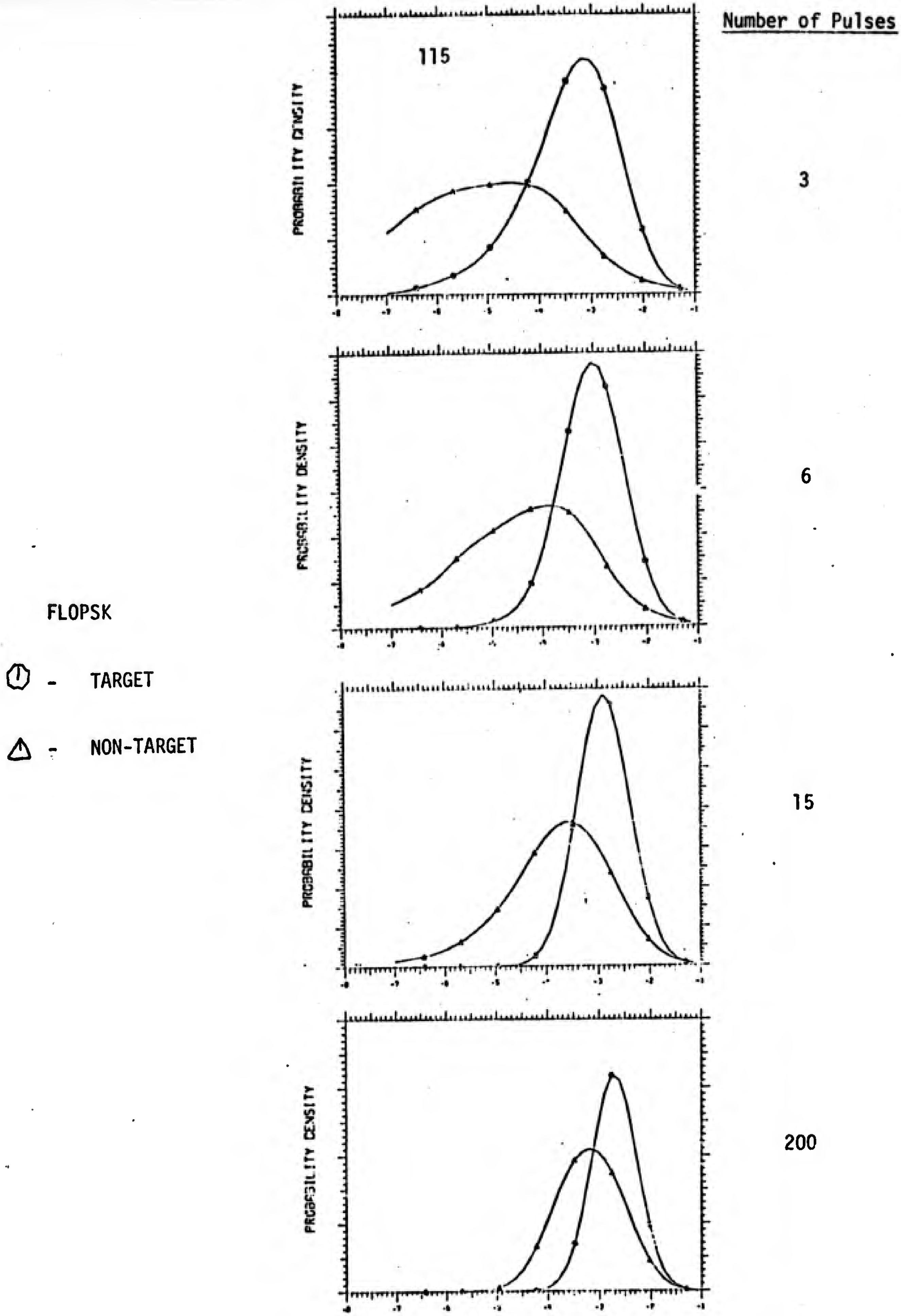

Number of Pulses
114
PROBABILITY DENSITY
3
6
15
200
FLPPKR
- TARGET
- NON-TARGET

Figure 7-14



Number of Pulses
115
PROBABILITY DENSITY
3
6
15
200
FLOPSK
TARGET
NON-TARGET

Figure 7-15

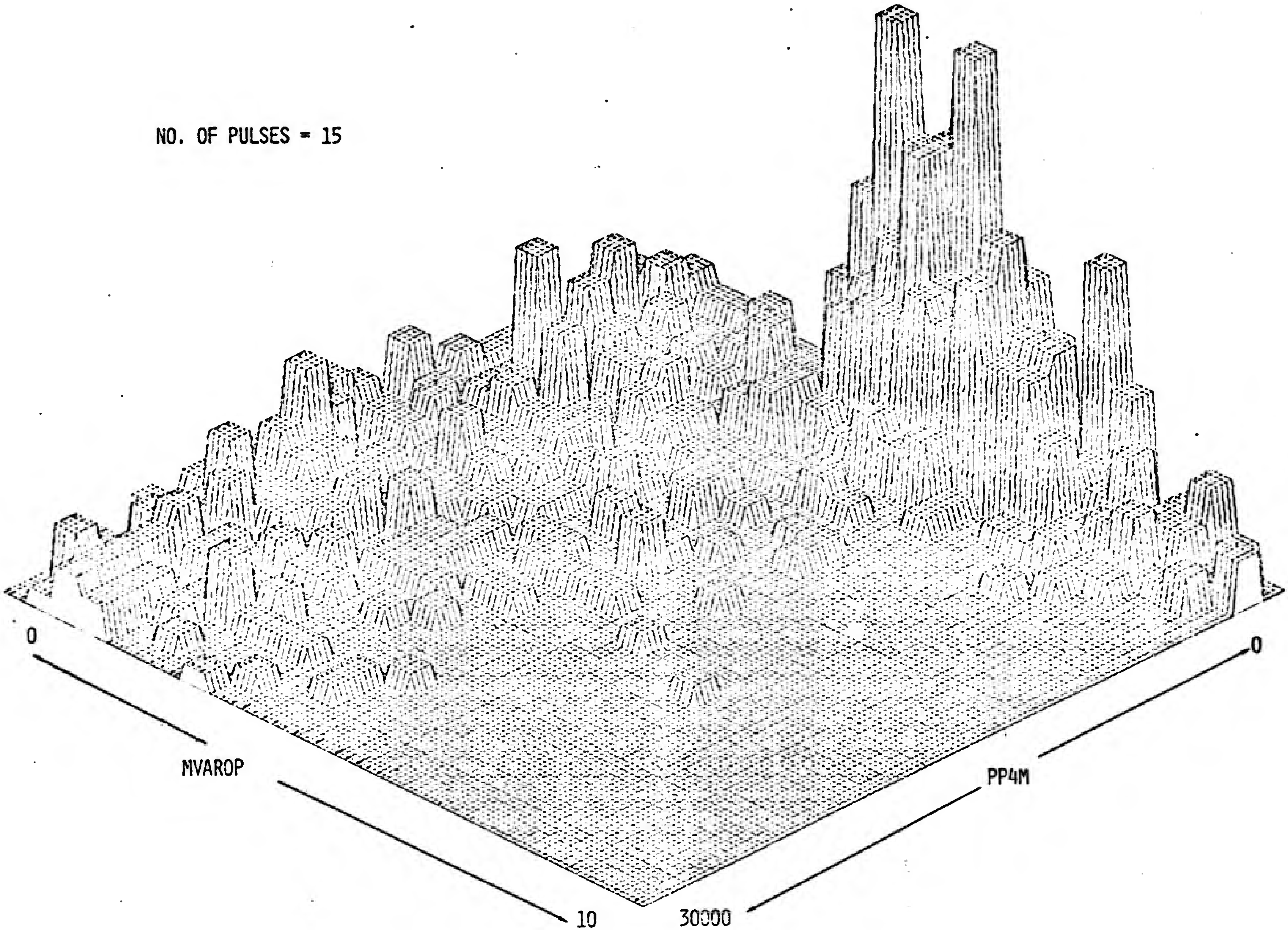


Figure 7-16

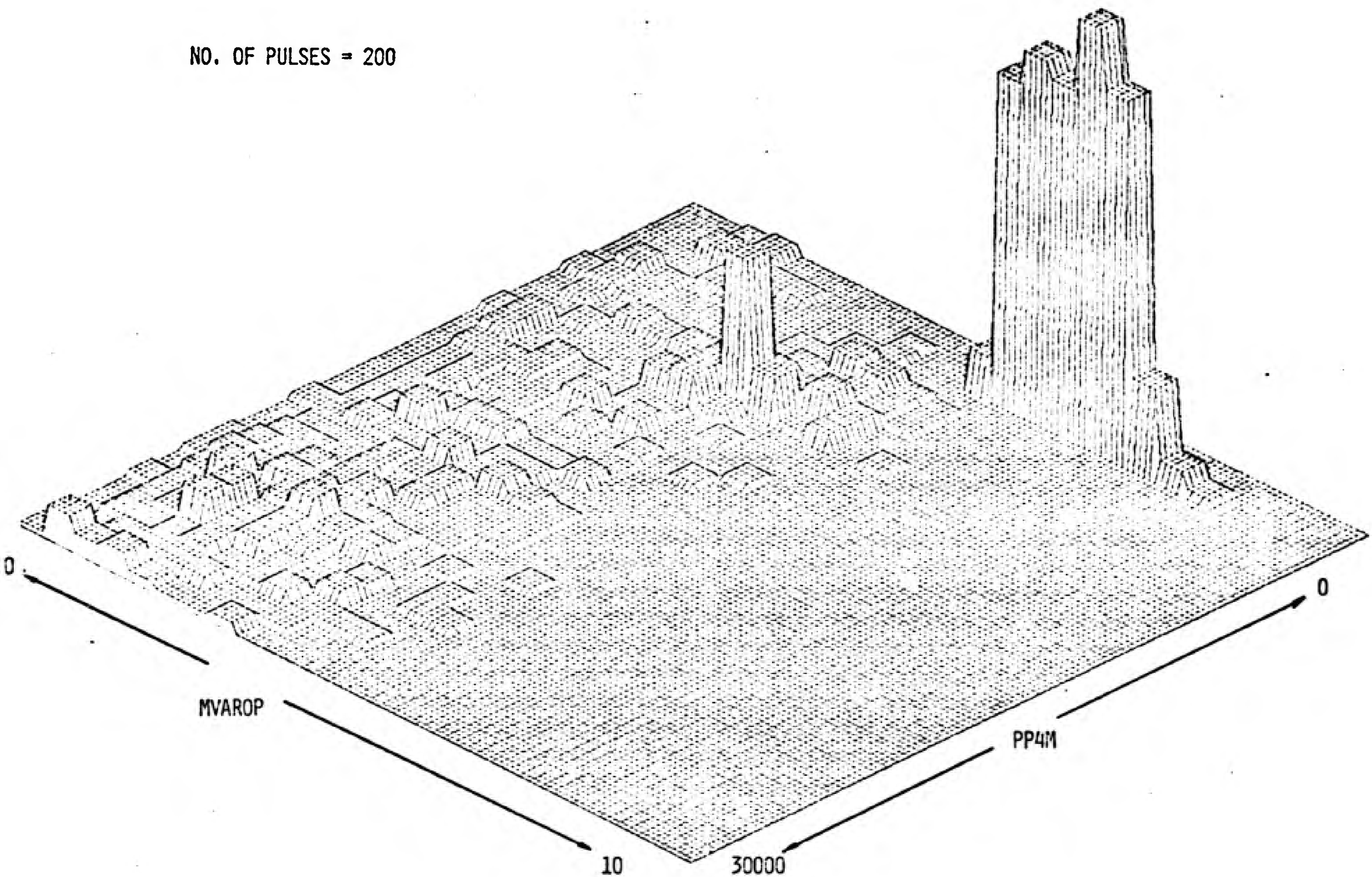


Figure 7-17

difficult to label the classes in such a diagram, we turn to scatter diagrams such as in figure 7-18, where the change in distribution with N of the pair of features TGM/REGR is indicated. Figure 7-19 provides performance curves as a function of N for that pair. Figures 7-20 through 7-22 provide similar information for the pairs MVAROP/PP4M and PP4M/EFFAP3. (Note that EFFAP3 in all cases uses only three pulses.) Figures 7-23 and 7-24 show performance curves for the triple MDOP/MVAROP/PP4M and the quadruple MDOP/MVAROP/PP4M/REGR.

Recall that the feature ranking was performed for N=15; the best feature sets for 15 pulses are not necessarily the best for N=3. Examination of the previous figures in this section shows that some features degrade more rapidly than others.

Figure 7-25 shows performance estimates for four single features at N=3 and N=15; comparison indicates a shift in relative performance. Figure 7-26 similarly displays the relative performance of pairs of features at N=3 and N=15. Feature ranking should optimally be repeated at lower numbers of pulses to avoid underestimating potential performance in those cases.

Although there is little reason to expect the validation of Section 6.0 for N=15 not to hold up for N=3, the performance on the design and test sets for the lower number of pulses is indicated in figures 7-27 through 7-30 for a single feature, a pair, a triple, and a quadruple. The validation holds in general.

X = TGM

Y = REGR

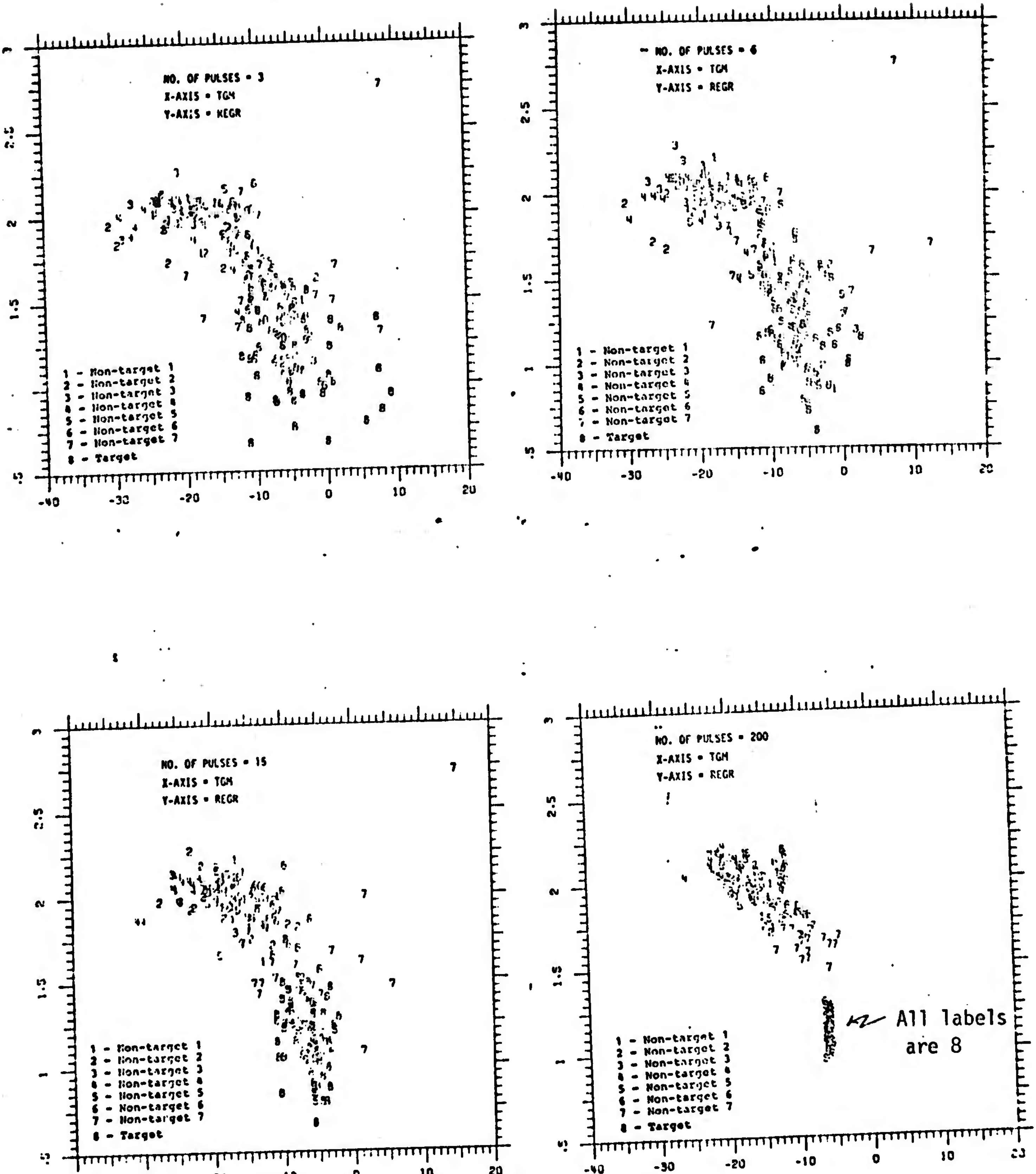
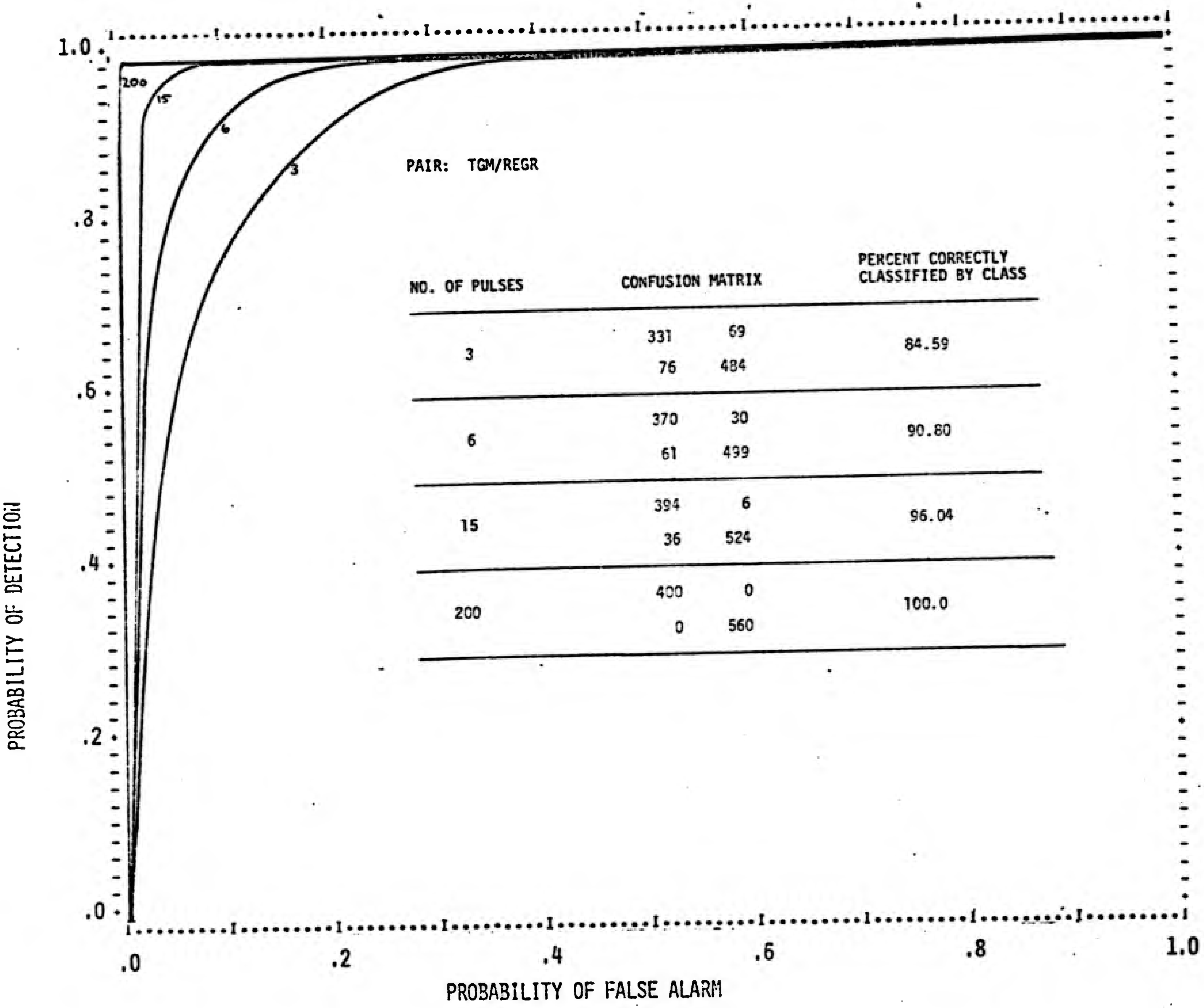


Figure 7-18



NO. OF PULSES	CONFUSION MATRIX		PERCENT CORRECTLY CLASSIFIED BY CLASS
3	331	69	84.59
	76	484	
6	370	30	90.80
	61	499	
15	394	6	96.04
	36	524	
200	400	0	100.0
	0	560	

Figure 7-19

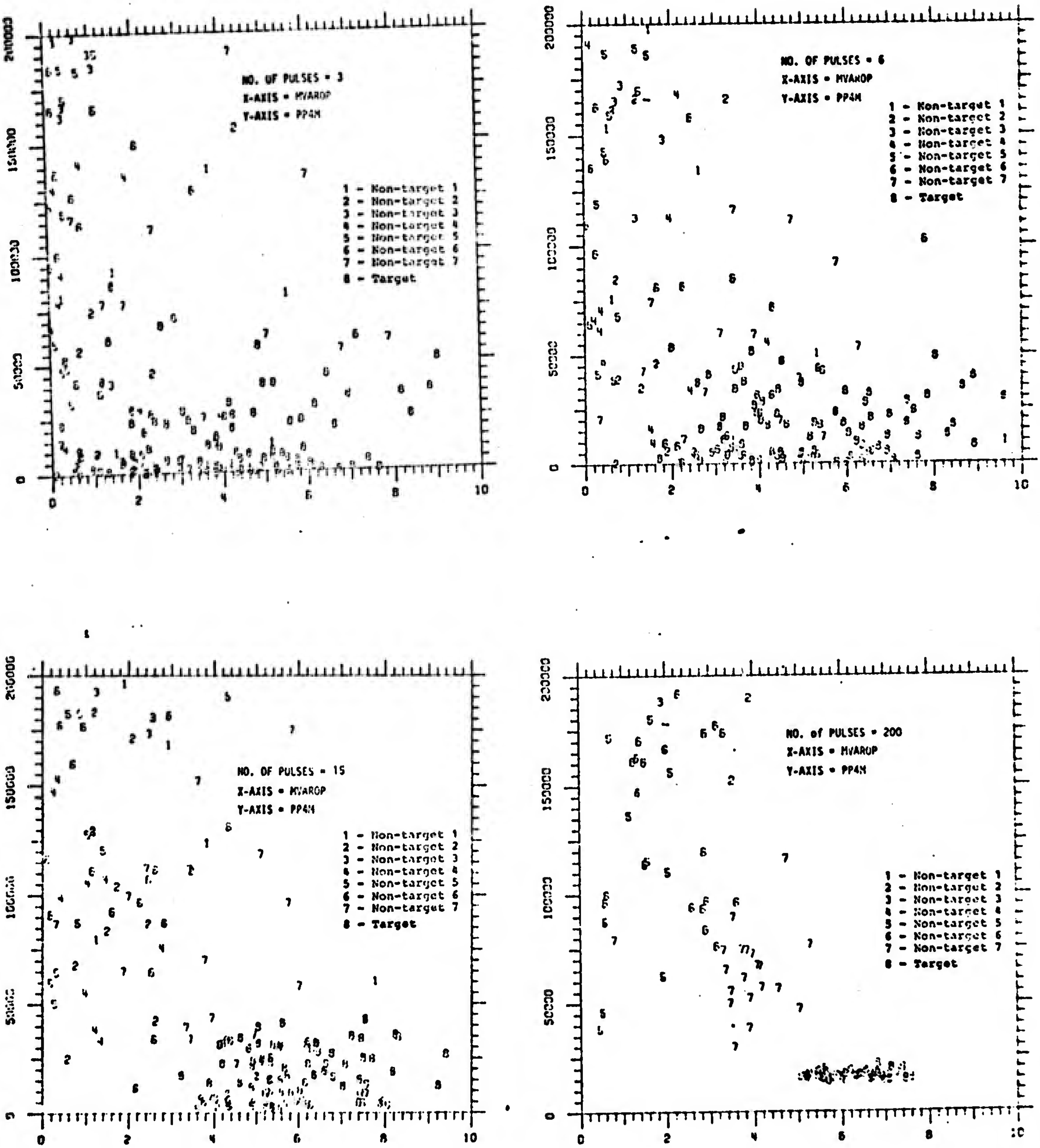
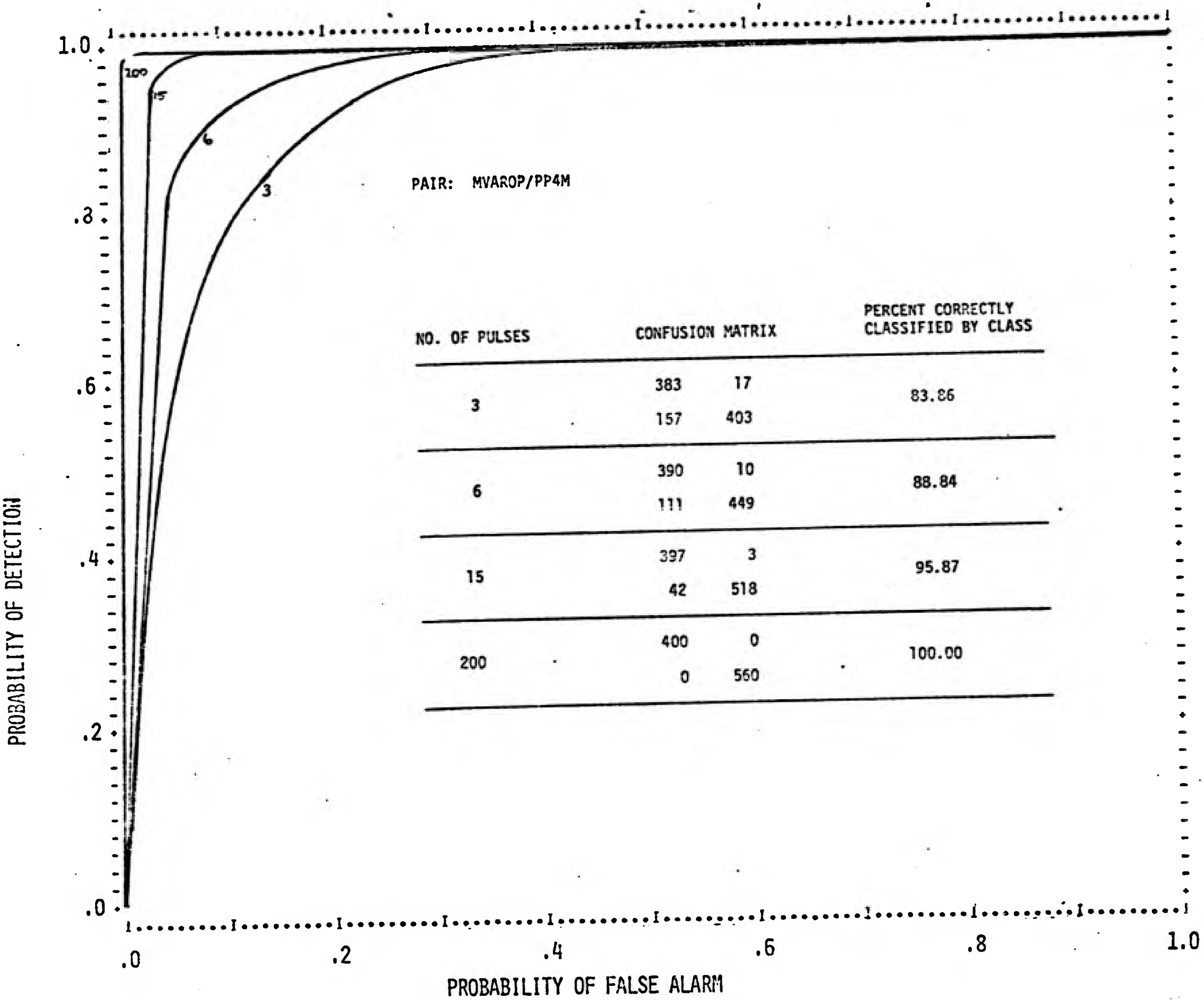

NO. OF PULSES = 3
X-AXIS = MVAROP
Y-AXIS = PP4M
1 - Non-target 1
2 - Non-target 2
3 - Non-target 3
4 - Non-target 4
5 - Non-target 5
6 - Non-target 6
7 - Non-target 7
8 - Target
NO. OF PULSES = 6
X-AXIS = MVAROP
Y-AXIS = PP4M
1 - Non-target 1
2 - Non-target 2
3 - Non-target 3
4 - Non-target 4
5 - Non-target 5
6 - Non-target 6
7 - Non-target 7
8 - Target
NO. OF PULSES = 15
X-AXIS = MVAROP
Y-AXIS = PP4M
1 - Non-target 1
2 - Non-target 2
3 - Non-target 3
4 - Non-target 4
5 - Non-target 5
6 - Non-target 6
7 - Non-target 7
8 - Target
NO. of PULSES = 200
X-AXIS = MVAROP
Y-AXIS = PP4M
1 - Non-target 1
2 - Non-target 2
3 - Non-target 3
4 - Non-target 4
5 - Non-target 5
6 - Non-target 6
7 - Non-target 7
8 - Target

Figure 7-20



NO. OF PULSES	CONFUSION MATRIX		PERCENT CORRECTLY CLASSIFIED BY CLASS
3	383	17	83.86
	157	403	
6	390	10	88.84
	111	449	
15	397	3	95.87
	42	518	
200	400	0	100.00
	0	560	

Figure 7-21

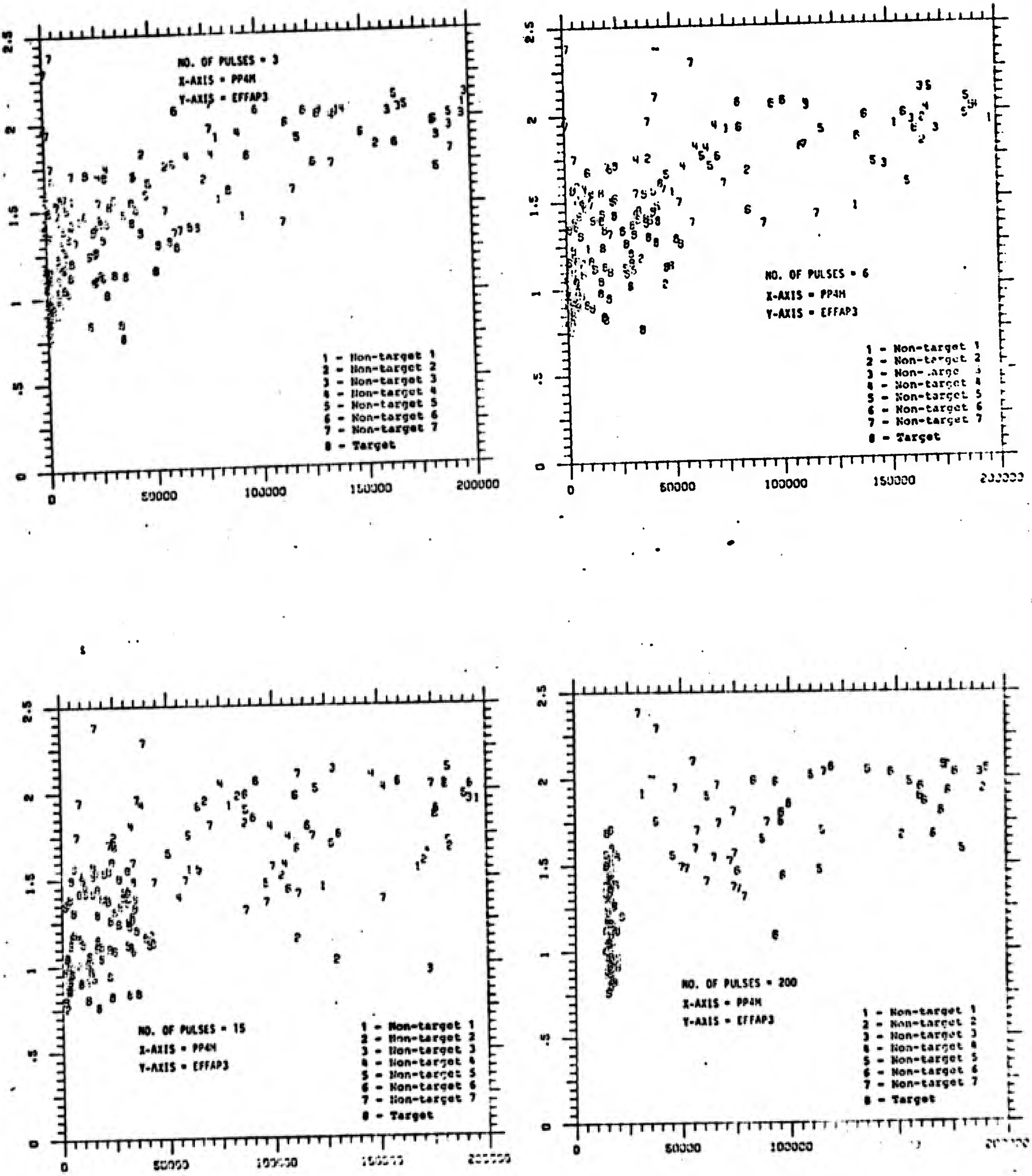
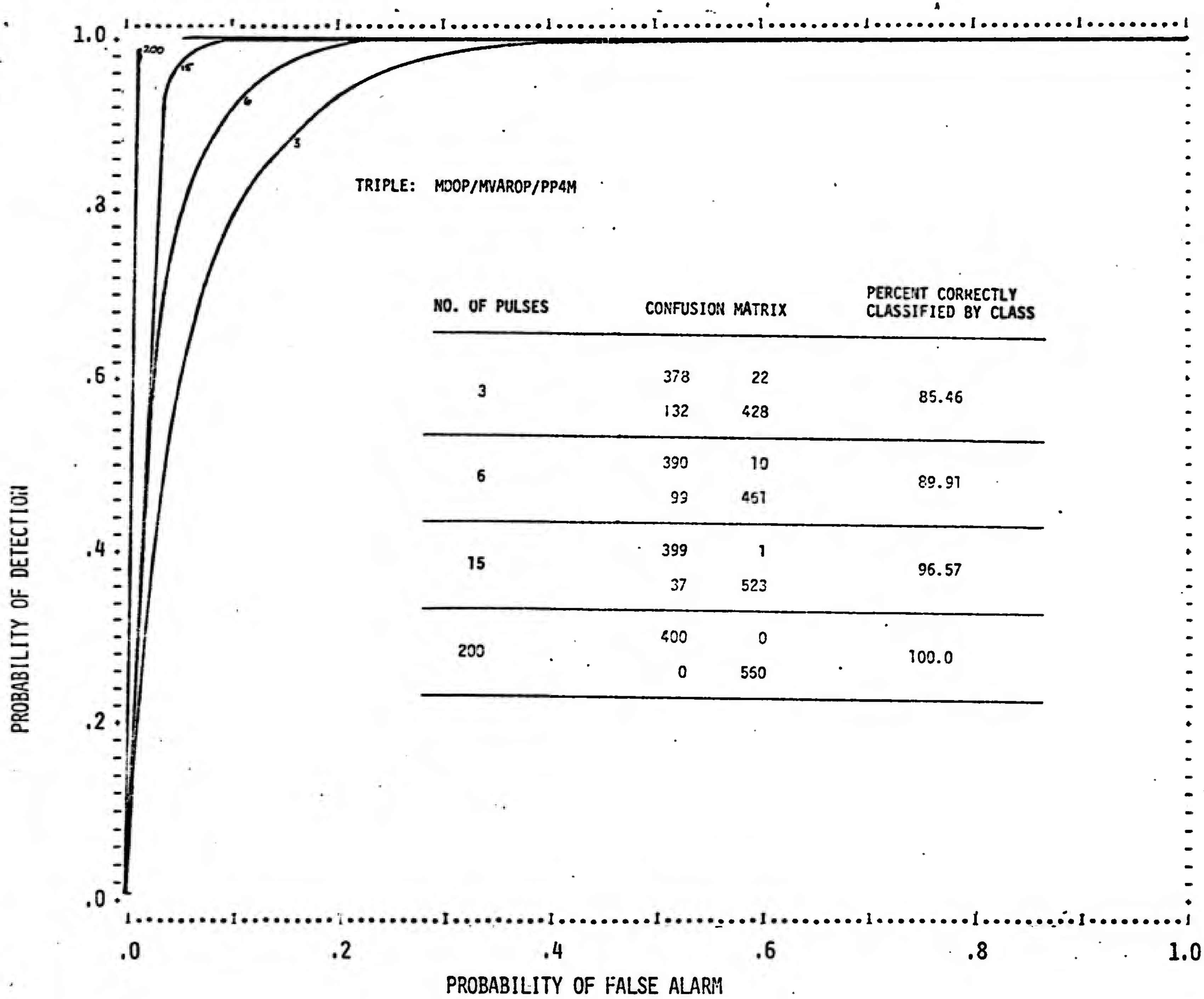
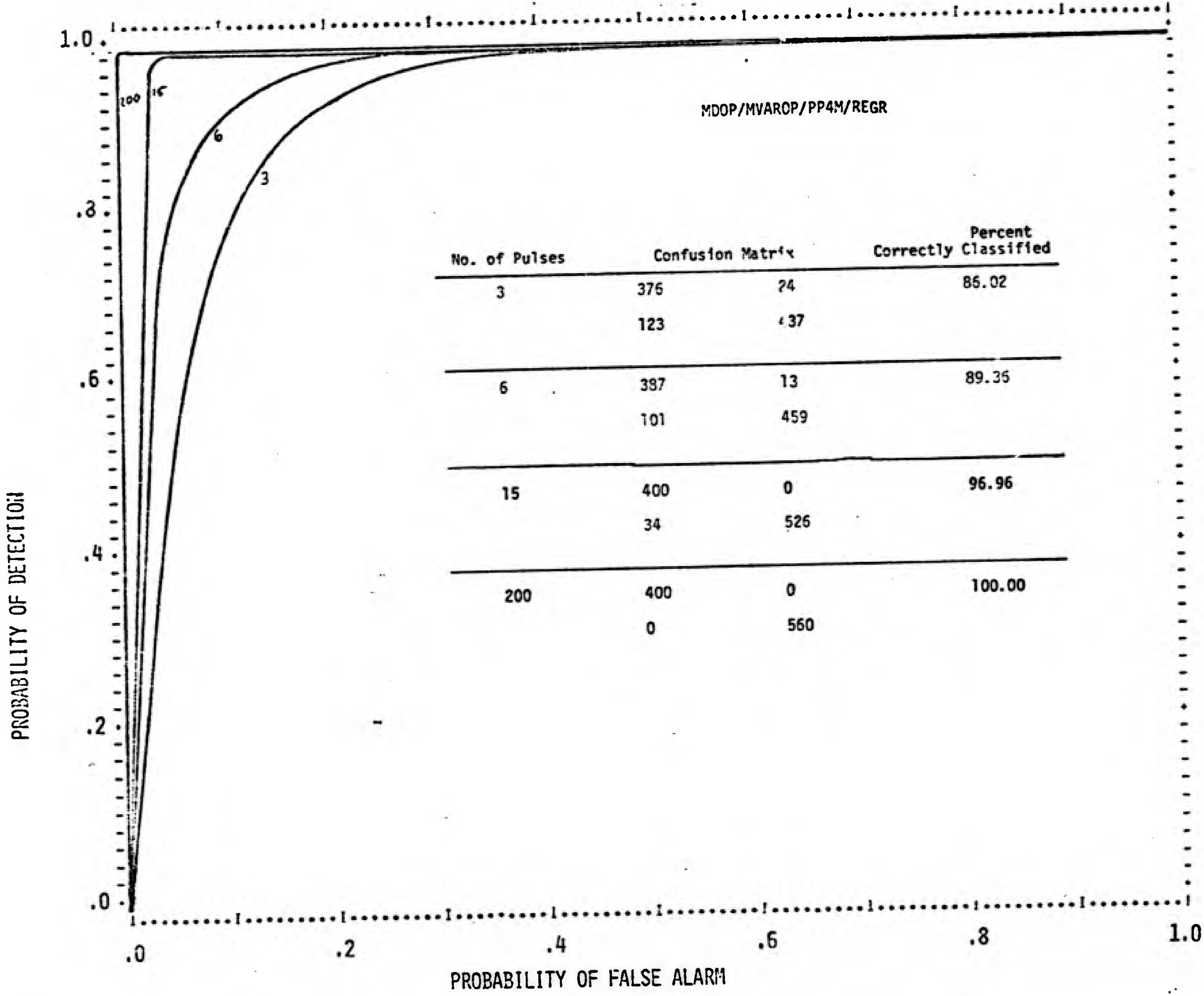

NO. OF PULSES = 3
X-AXIS = PP4M
Y-AXIS = EFFAP3
1 - Non-target 1
2 - Non-target 2
3 - Non-target 3
4 - Non-target 4
5 - Non-target 5
6 - Non-target 6
7 - Non-target 7
8 - Target
0
.5
1
1.5
2
2.5
0
50000
100000
150000
200000
NO. OF PULSES = 6
X-AXIS = PP4M
Y-AXIS = EFFAP3
NO. OF PULSES = 15
X-AXIS = PP4M
Y-AXIS = EFFAP3
NO. OF PULSES = 200
X-AXIS = PP4M
Y-AXIS = EFFAP3

Figure 7-22



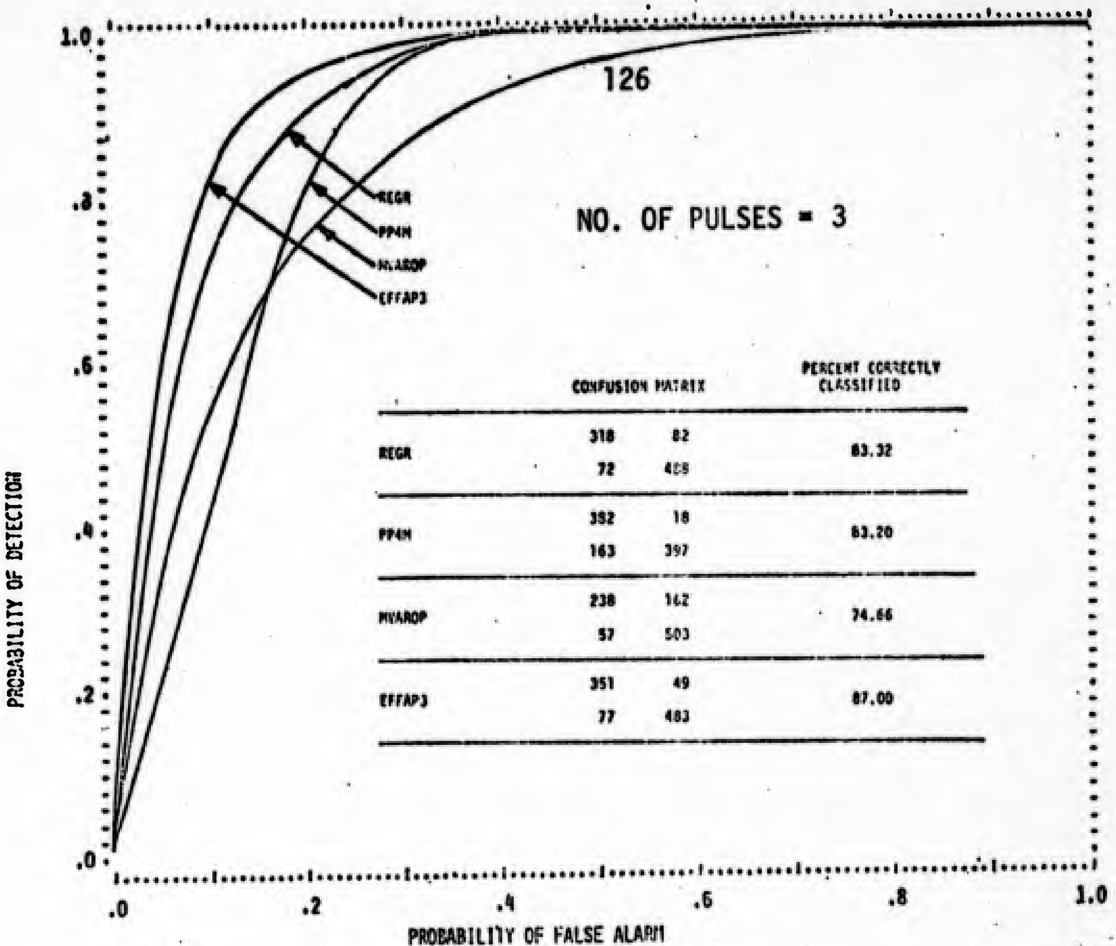
NO. OF PULSES	CONFUSION MATRIX		PERCENT CORRECTLY CLASSIFIED BY CLASS
3	378	22	85.46
	132	428	
6	390	10	89.91
	93	461	
15	399	1	96.57
	37	523	
200	400	0	100.0
	0	550	

Figure 7-23

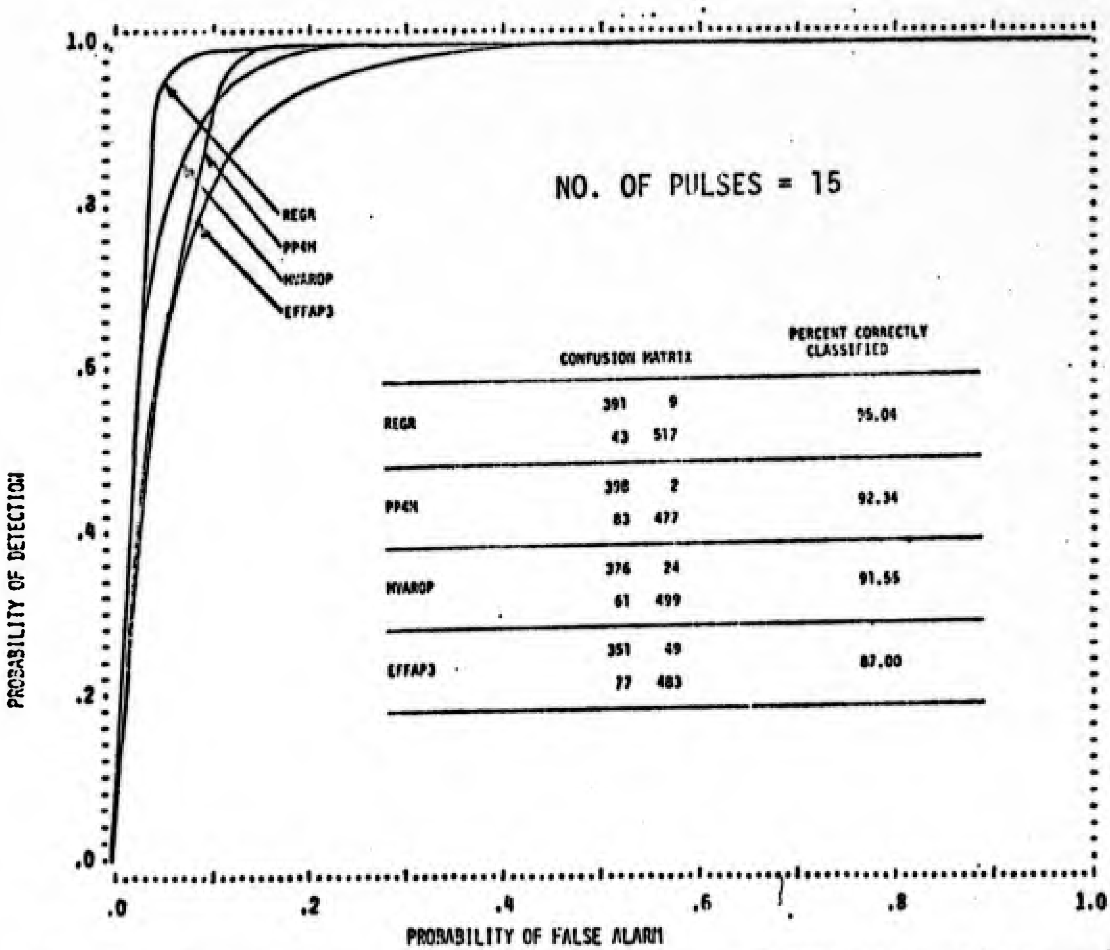


No. of Pulses	Confusion Matrix		Percent Correctly Classified
3	376	24	86.02
	123	437	
6	387	13	89.35
	101	459	
15	400	0	96.96
	34	526	
200	400	0	100.00
	0	560	

Figure 7-24

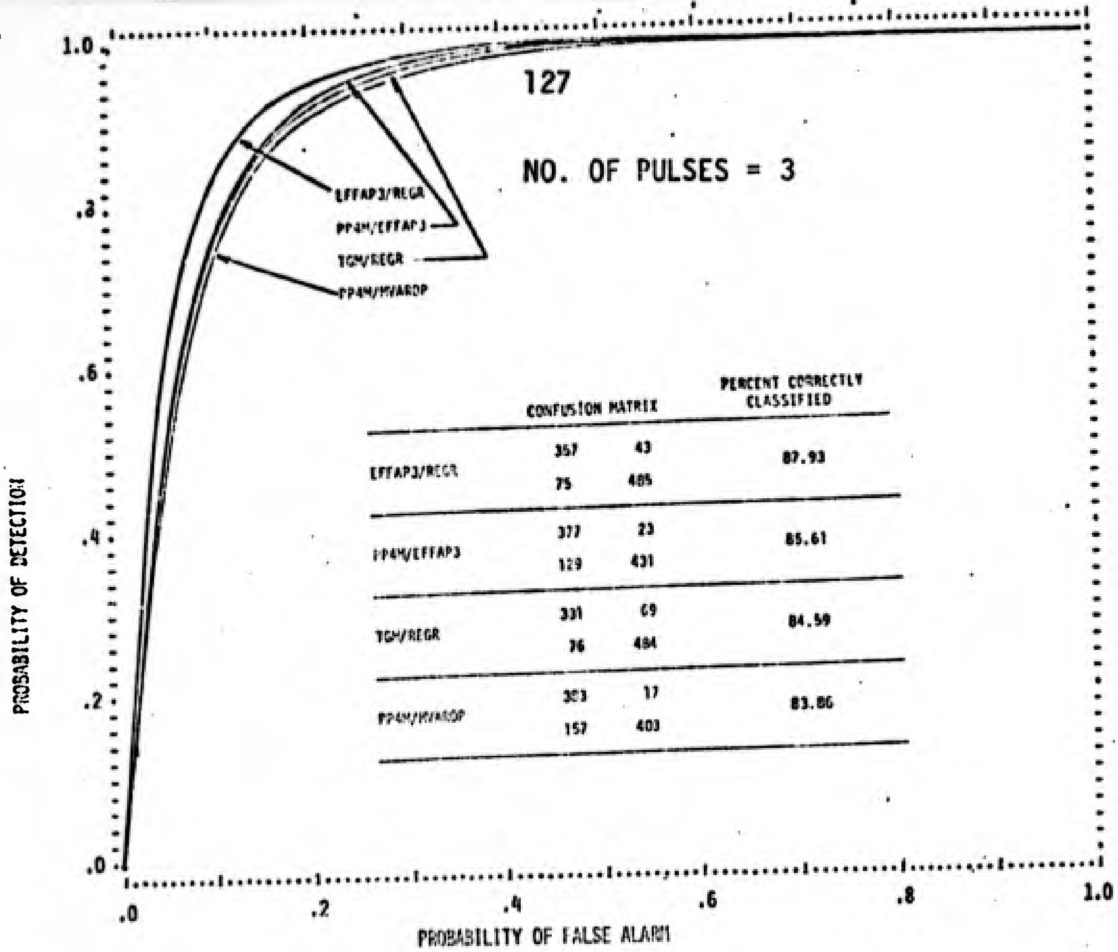


	CONFUSION MATRIX		PERCENT CORRECTLY CLASSIFIED
REGR	318	82	83.32
	72	488	
PP4M	392	18	83.20
	163	397	
MVAROP	238	162	74.66
	57	503	
EFFAP3	351	49	87.00
	77	483	

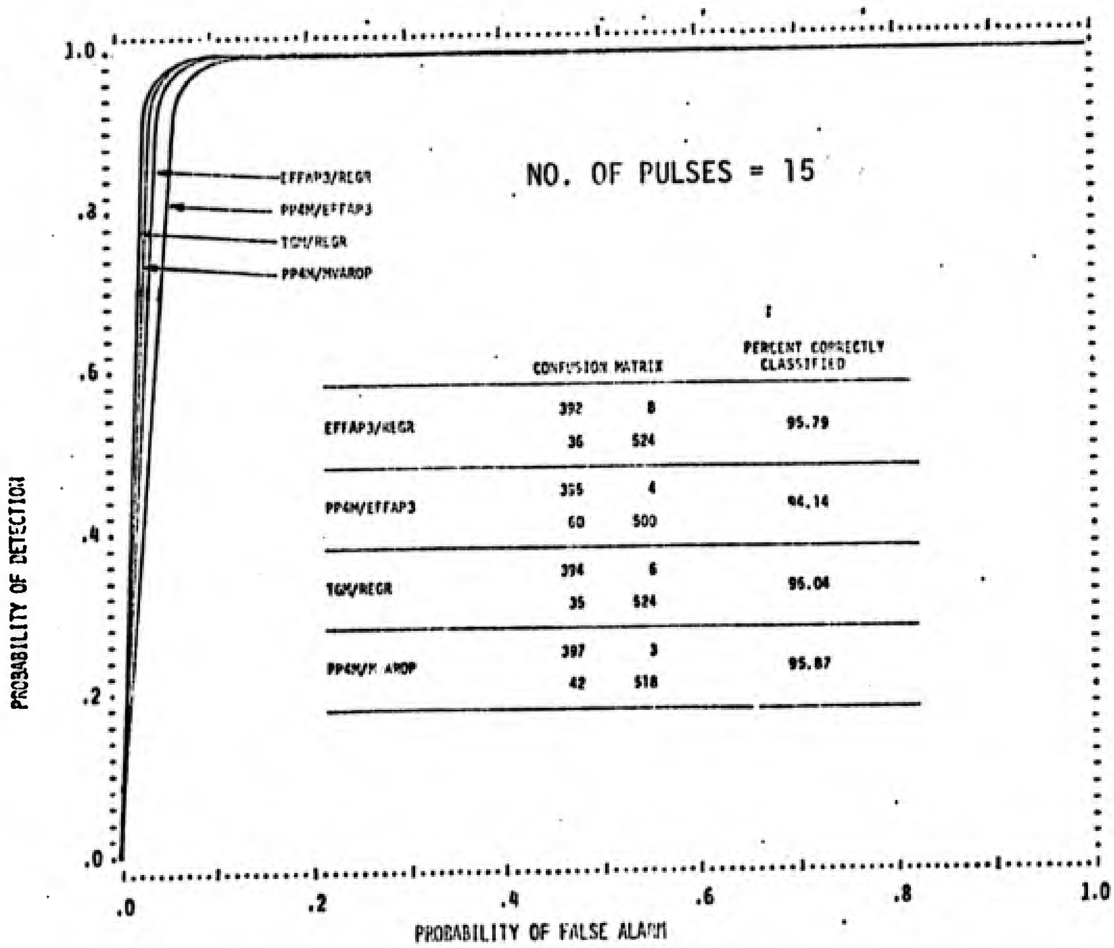


	CONFUSION MATRIX		PERCENT CORRECTLY CLASSIFIED
REGR	391	9	95.04
	43	517	
PP4M	398	2	92.34
	83	477	
MVAROP	376	24	91.55
	61	499	
EFFAP3	351	49	87.00
	77	483	

Figure 7-25: Comparison of single feature performance at N=3 and N=15.

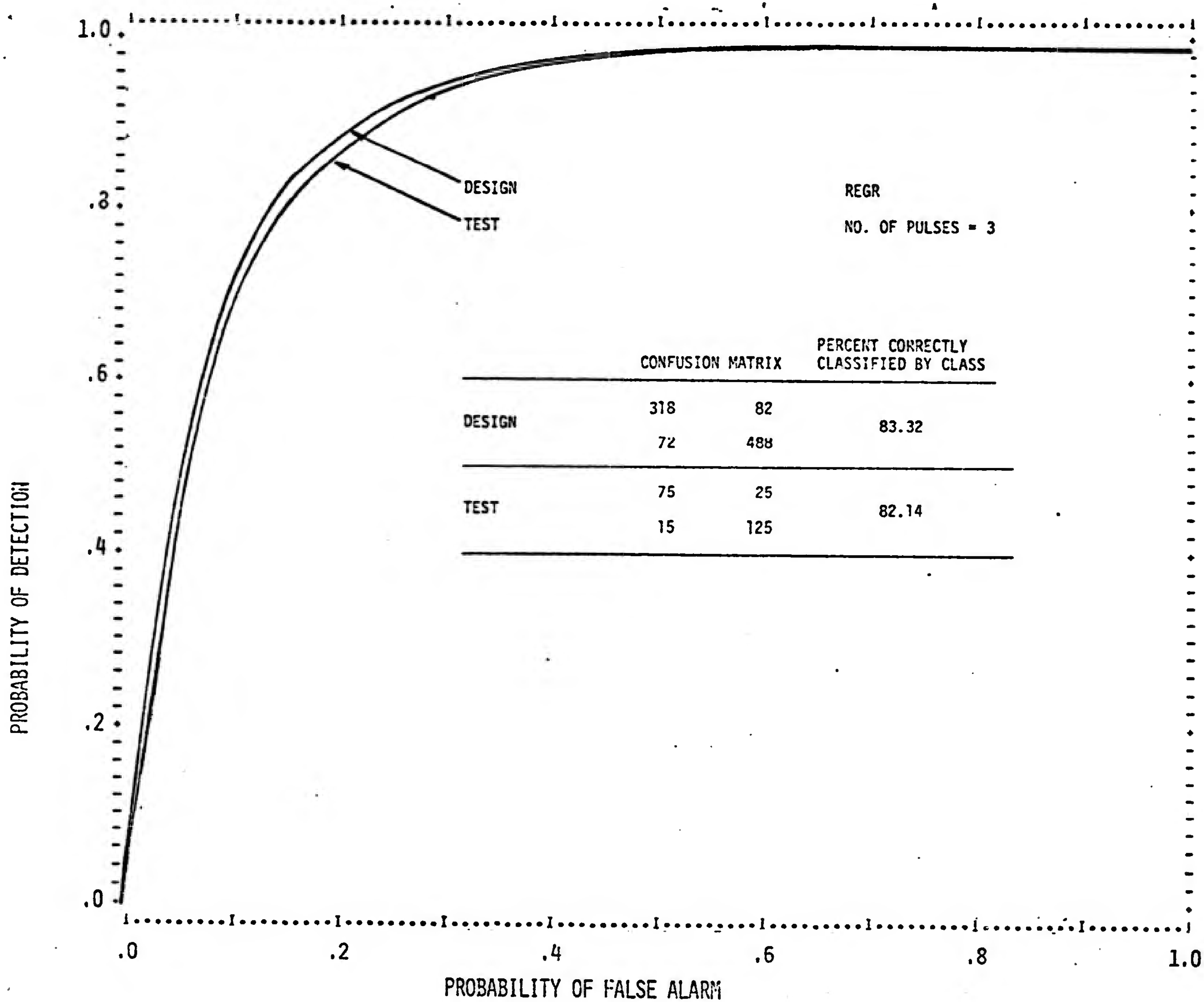


	CONFUSION MATRIX		PERCENT CORRECTLY CLASSIFIED
EFFAP3/REGR	357	43	87.93
	75	485	
PP4M/EFFAP3	377	23	85.61
	129	431	
TGM/REGR	331	69	84.59
	76	484	
PP4M/MVAROP	383	17	83.86
	157	403	



	CONFUSION MATRIX		PERCENT CORRECTLY CLASSIFIED
EFFAP3/REGR	392	8	95.79
	36	524	
PP4M/EFFAP3	396	4	94.14
	60	500	
TGM/REGR	394	6	95.04
	35	524	
PP4M/M AROP	397	3	95.87
	42	518	

Figure 7-26: Comparison of performance of pairs of features at N=3 and N=15.



	CONFUSION MATRIX		PERCENT CORRECTLY CLASSIFIED BY CLASS
DESIGN	318	82	83.32
	72	488	
TEST	75	25	82.14
	15	125	

Figure 7-27

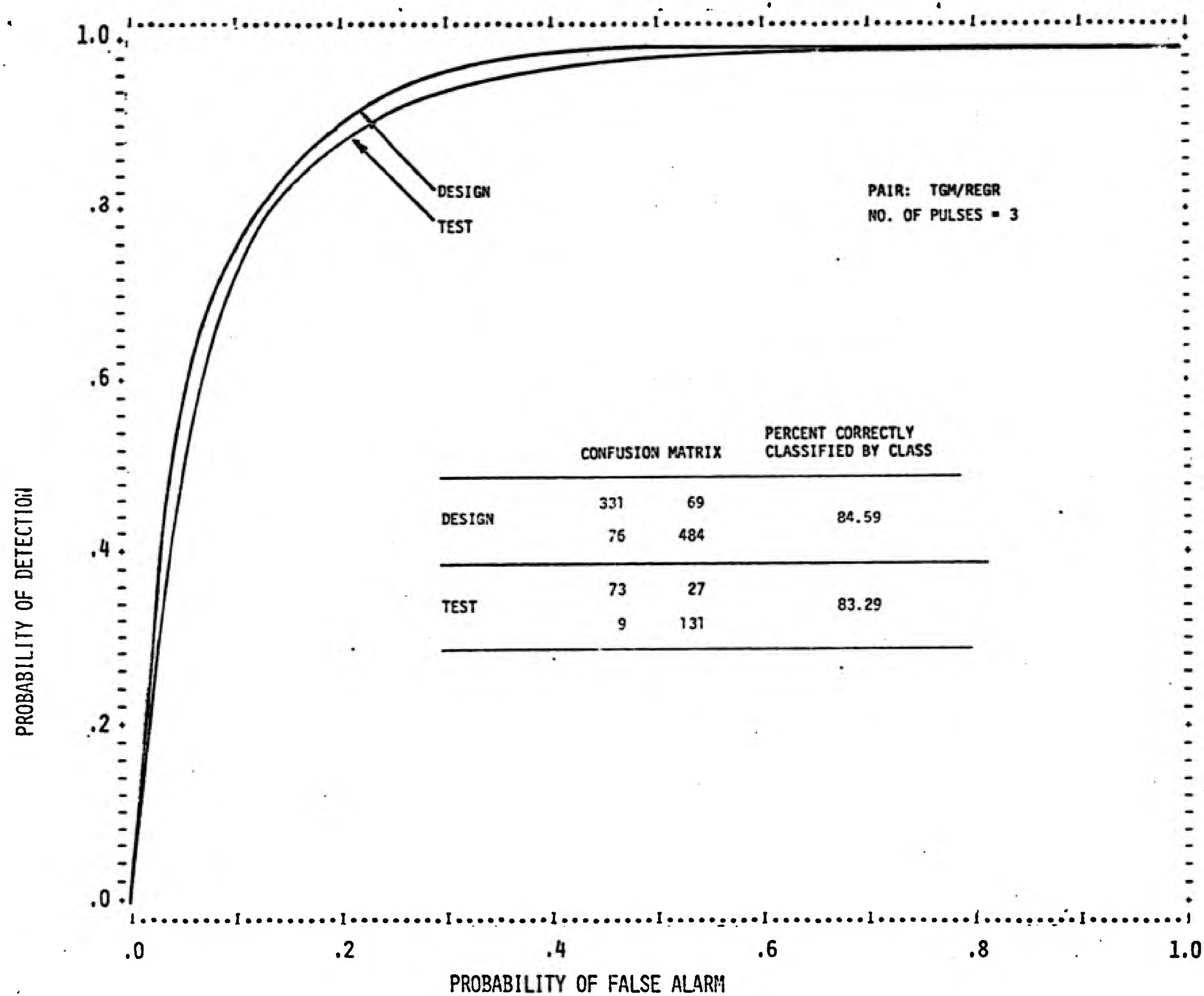

1.0
.8
.6
.4
.2
.0
PROBABILITY OF DETECTION
DESIGN
TEST
PAIR: TGM/REGR
NO. OF PULSES = 3
CONFUSION MATRIX
PERCENT CORRECTLY CLASSIFIED BY CLASS
DESIGN
331 69
76 484
84.59
TEST
73 27
9 131
83.29
.0
.2
.4
.6
.8
1.0
PROBABILITY OF FALSE ALARM

Figure 7-28

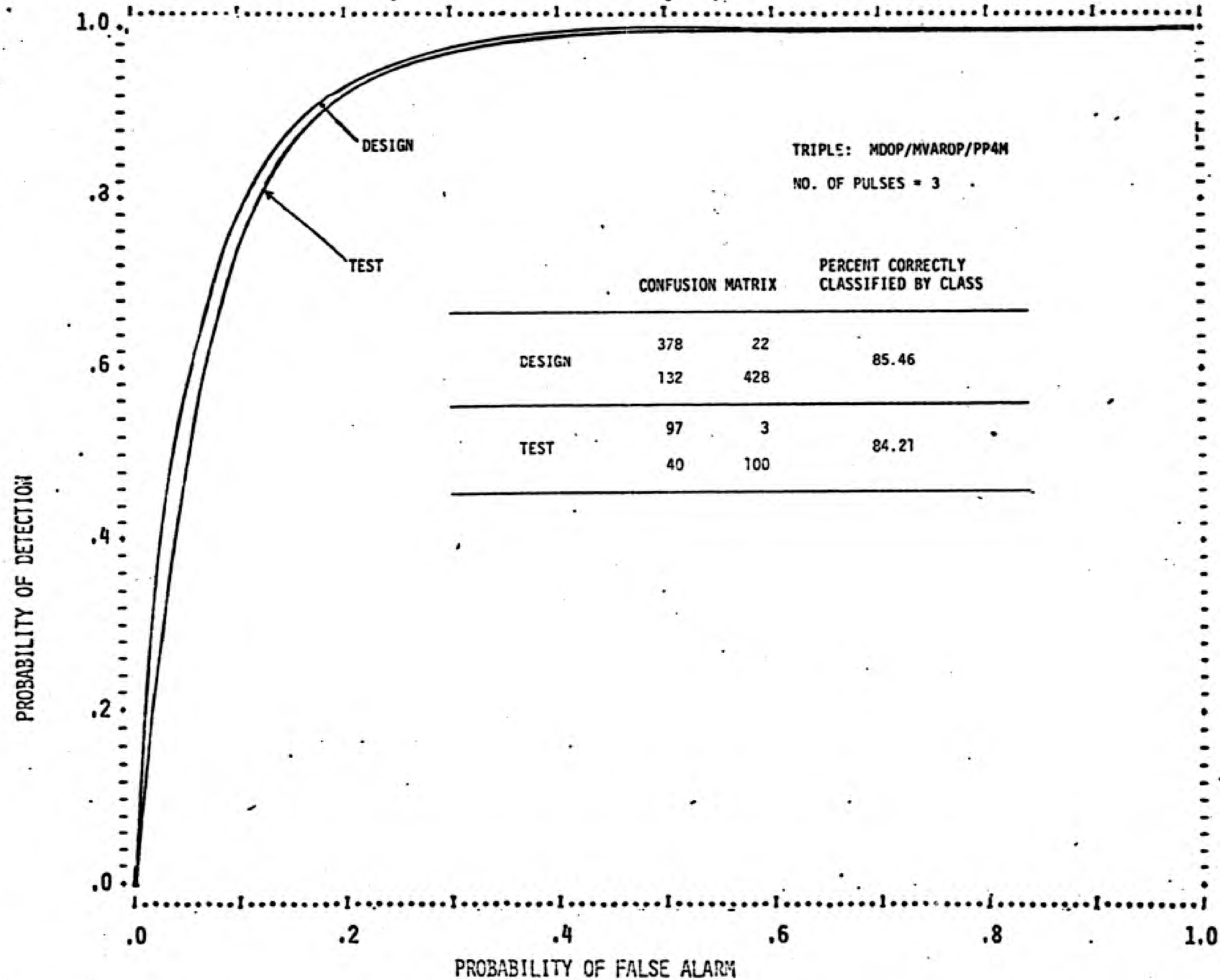

TRIPLE: MDOP/MVAROP/PP4M
NO. OF PULSES = 3
DESIGN
TEST
CONFUSION MATRIX
PERCENT CORRECTLY CLASSIFIED BY CLASS
DESIGN 378 22 132 428 85.46
TEST 97 3 40 100 84.21
1.0
.8
.6
.4
.2
.0
PROBABILITY OF DETECTION
PROBABILITY OF FALSE ALARM

Figure 7.29

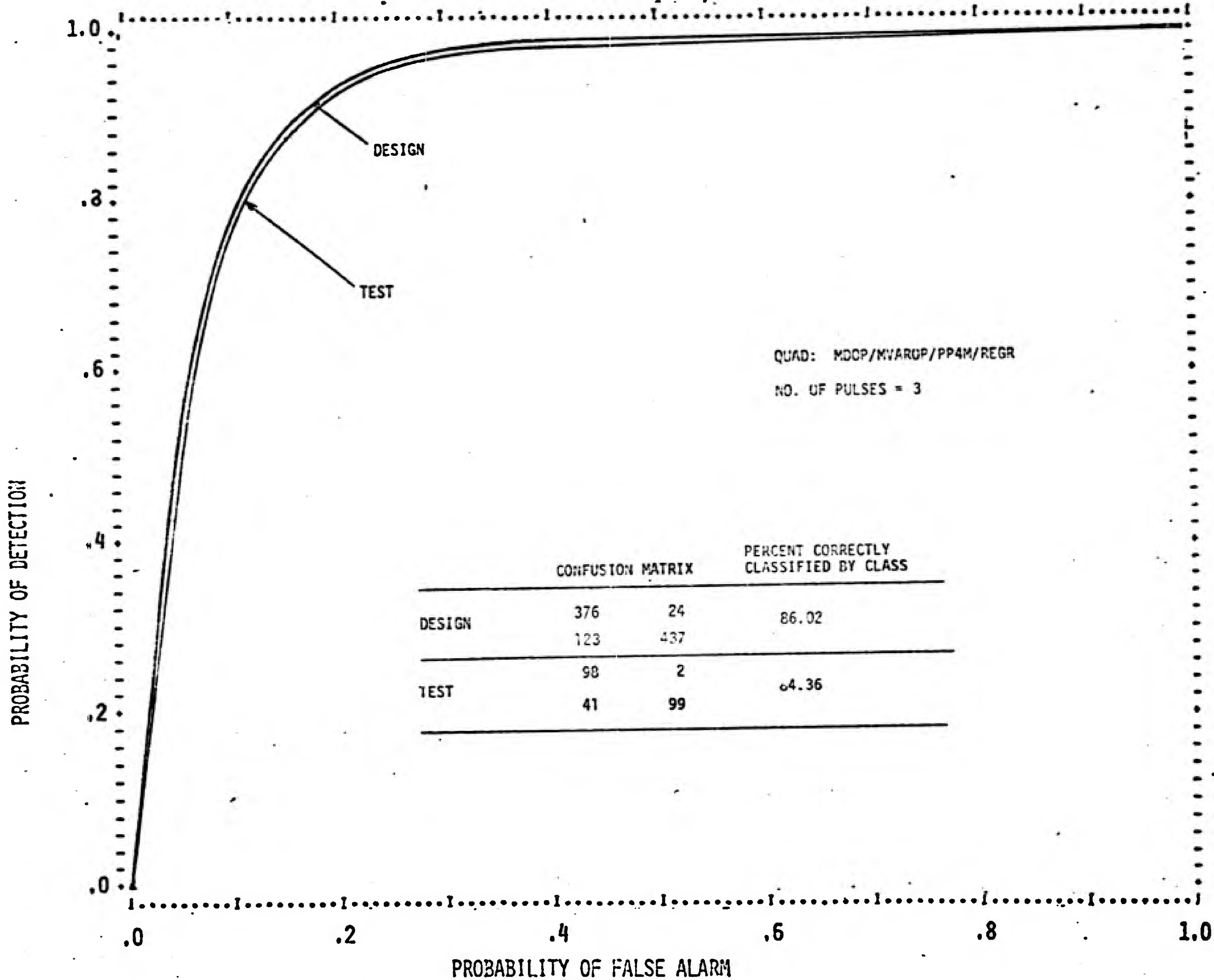

1.0
.8
.6
.4
.2
.0
PROBABILITY OF DETECTION
DESIGN
TEST
QUAD: MDOP/MVARUP/PP4M/REGR
NO. OF PULSES = 3
CONFUSION MATRIX
PERCENT CORRECTLY CLASSIFIED BY CLASS
DESIGN 376 24 86.02
123 437
TEST 98 2 64.36
41 99
.0
.2
.4
.6
.8
1.0
PROBABILITY OF FALSE ALARM

Figure 7.30

8.0 THE INVERSE PROBLEM: WHAT CHARACTERISTICS MAKE DISCRIMINATION DIFFICULT?

Two kinds of results have been reported in previous sections. One type of result is the level of discrimination possible given a collection of non-target classes and a given target class. Another is determination of the type of features and data which make discrimination possible. The inverse problem is to determine which non-target classes cause the most discrimination problems, and which particular signatures cause the most problems in the problem classes.

There are three aspects of the inverse problem we will discuss. The first aspect is examining which of the seven non-target subclasses are most difficult to discriminate from the target. The second aspect is uncovering which characteristics of the signatures and what corresponding physical properties (motion, aspect, etc.) cause discrimination problems. The third aspect is the question of to what degree the performance would degrade if all non-targets came from the most difficult subclass. The scope of this project will not allow examination of the questions in depth; we will note some of the more obvious conclusions and illustrate the approach to further examination in detail.

Figures 8-1 through 8-11 indicate for several single features (for N=15 and N=3) the overlap of the class-conditional probability densities (estimated by Parzen estimators from the samples) of the target class and the seven non-target subclasses. Non-target subclass 7 is a major offender for

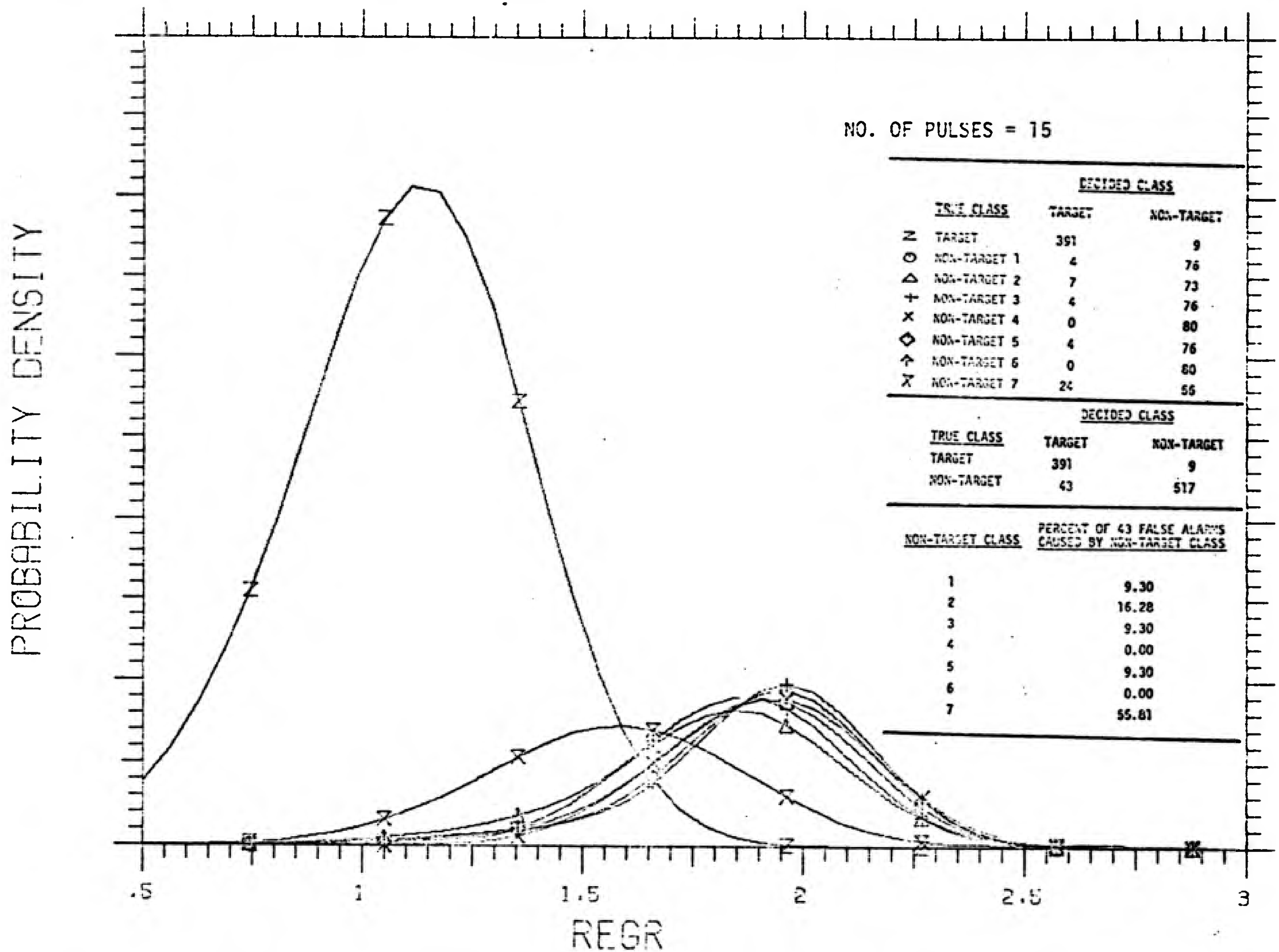


Figure 8-1: Overlap of class conditional probabilities by non-target subclass (indication of percent of false alarms caused by each non-target subclass).

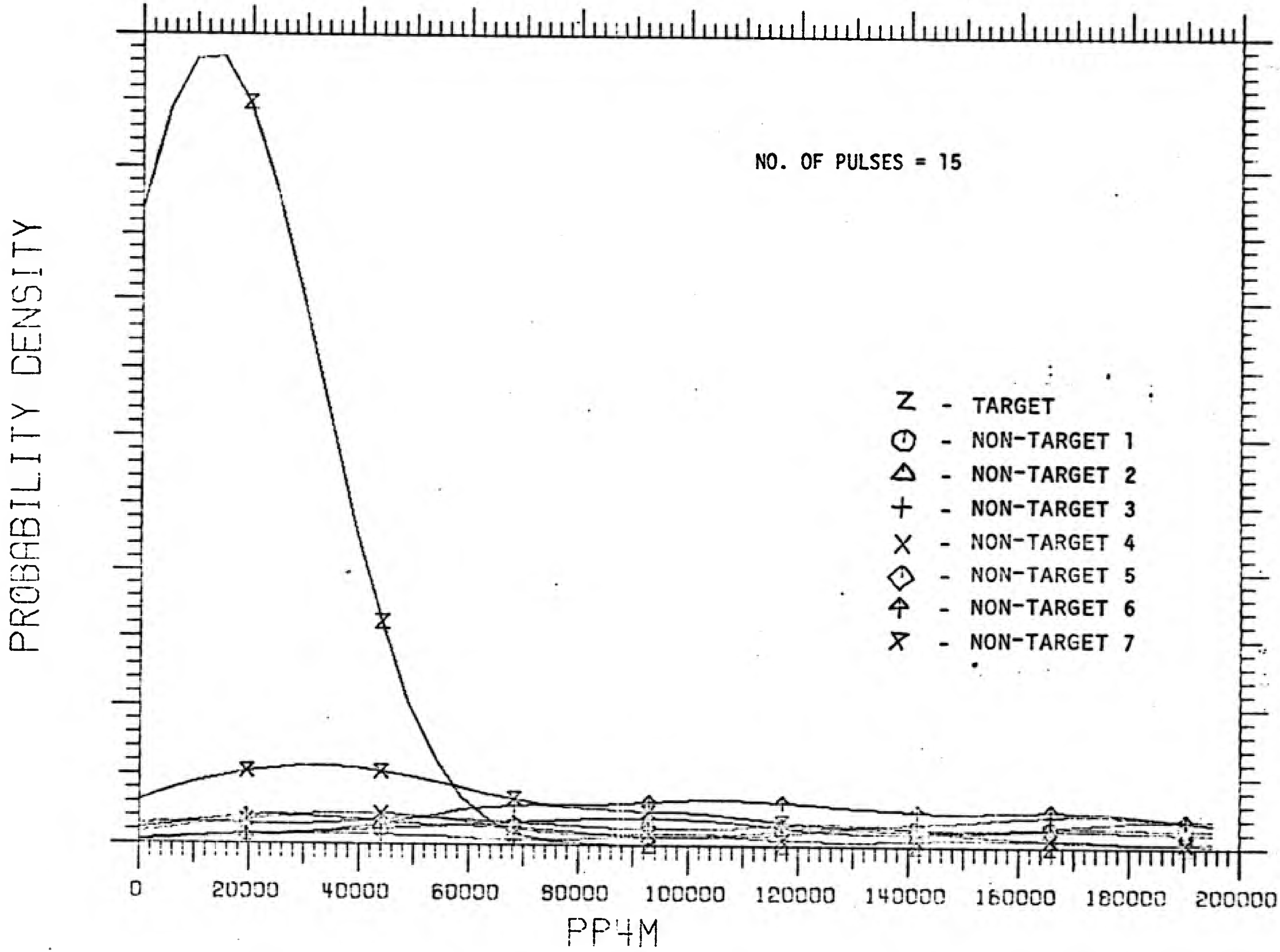

NO. OF PULSES = 15
Z - TARGET
- NON-TARGET 1
- NON-TARGET 2
+ - NON-TARGET 3
× - NON-TARGET 4
- NON-TARGET 5
- NON-TARGET 6
- NON-TARGET 7
0
20000
40000
60000
80000
100000
120000
140000
160000
180000
200000
PP4M
PROBABILITY DENSITY

Figure 8-2

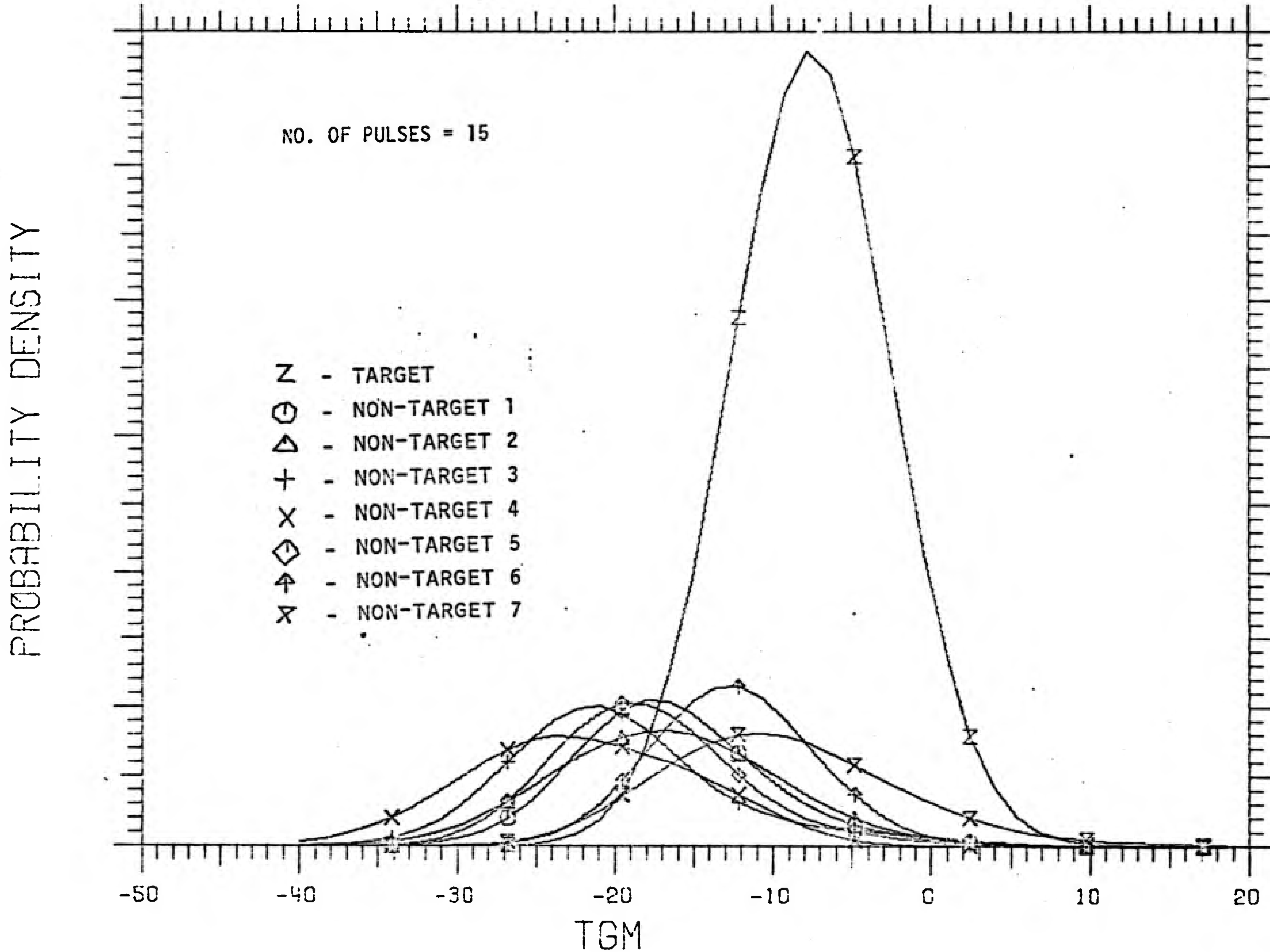

NO. OF PULSES = 15
Z - TARGET
- NON-TARGET 1
- NON-TARGET 2
- NON-TARGET 3
- NON-TARGET 4
- NON-TARGET 5
- NON-TARGET 6
- NON-TARGET 7
PROBABILITY DENSITY
TGM
-50
-40
-30
-20
-10
0
10
20

Figure 8-3

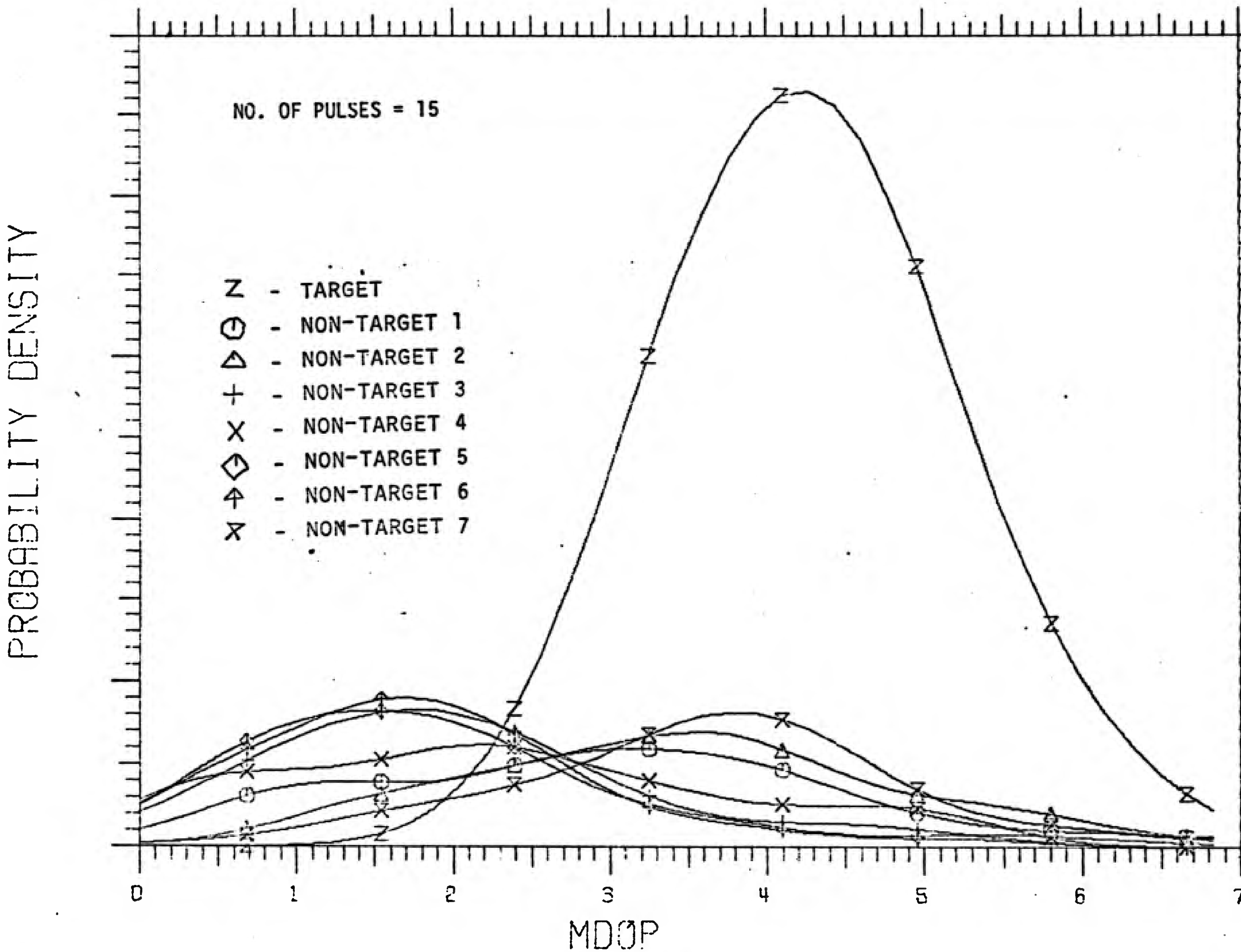

NO. OF PULSES = 15
Z - TARGET
- NON-TARGET 1
- NON-TARGET 2
- NON-TARGET 3
- NON-TARGET 4
- NON-TARGET 5
- NON-TARGET 6
- NON-TARGET 7
PROBABILITY DENSITY
0
1
2
3
4
5
6
7
MDOP

Figure 8-4

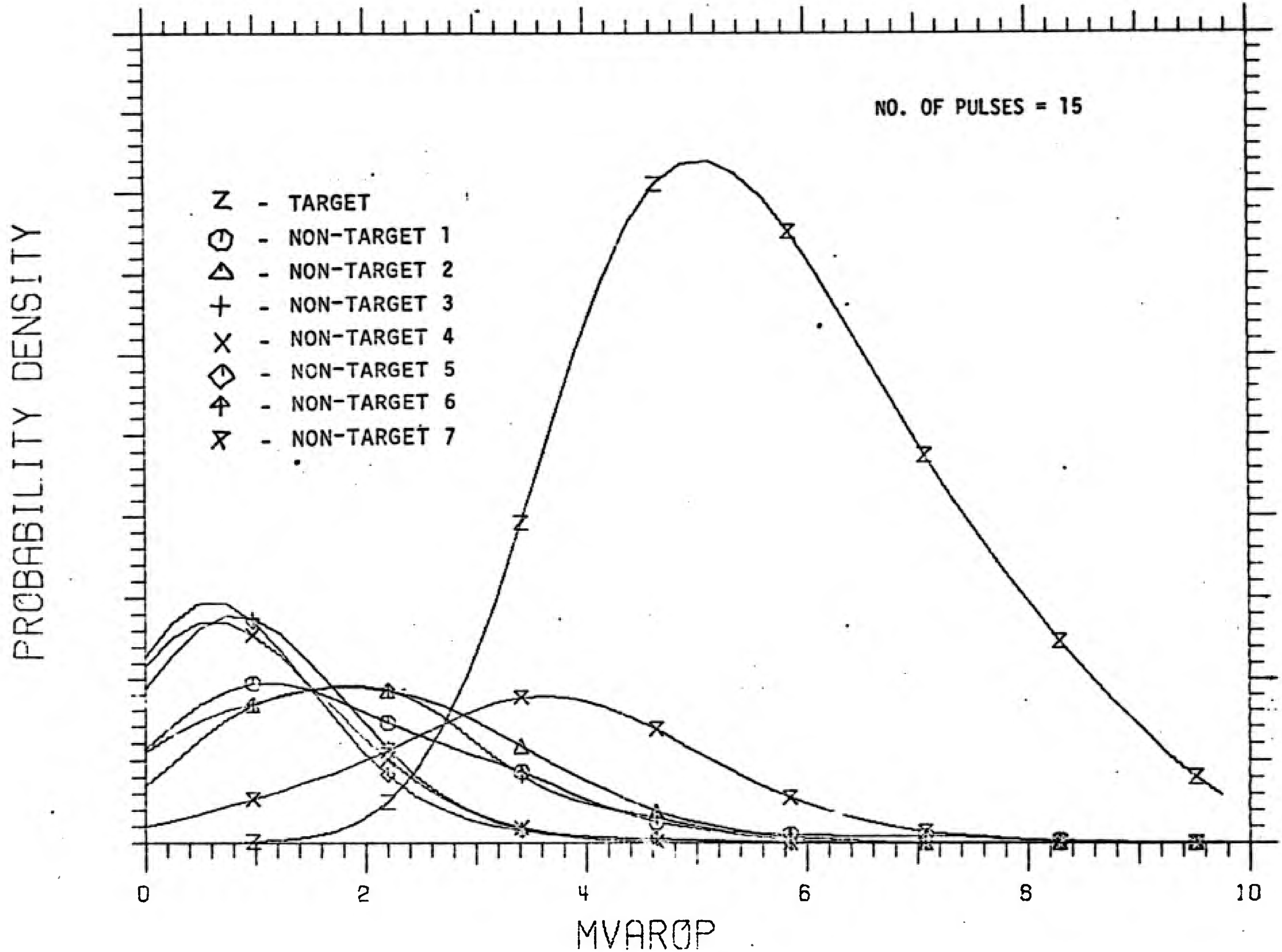

NO. OF PULSES = 15
Z - TARGET
NON-TARGET 1
NON-TARGET 2
NON-TARGET 3
NON-TARGET 4
NON-TARGET 5
NON-TARGET 6
NON-TARGET 7
PROBABILITY DENSITY
0
2
4
6
8
10
MVARØP

Figure 8-5

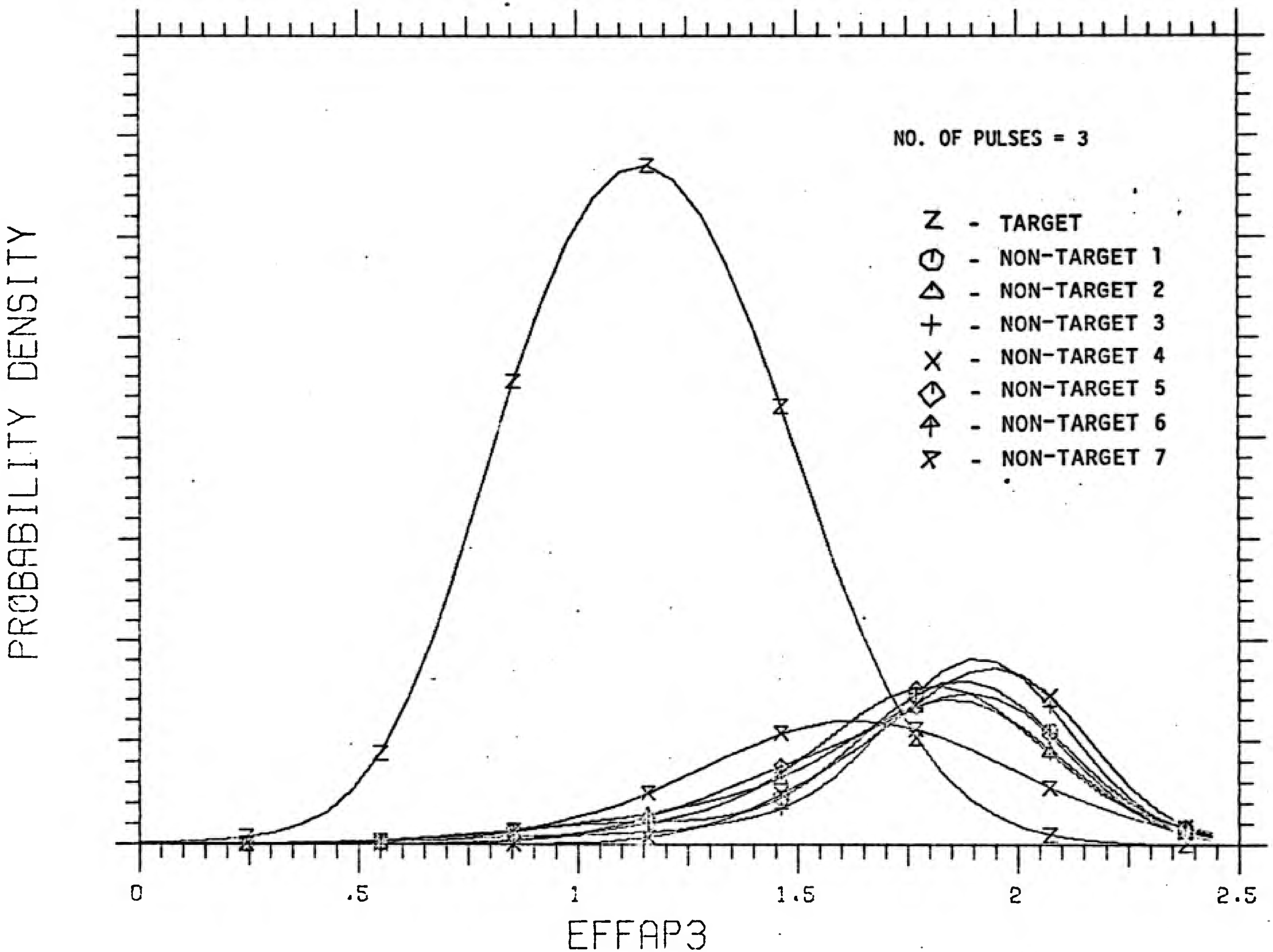

NO. OF PULSES = 3
Z - TARGET
- NON-TARGET 1
- NON-TARGET 2
- NON-TARGET 3
- NON-TARGET 4
- NON-TARGET 5
- NON-TARGET 6
- NON-TARGET 7
PROBABILITY DENSITY
0
.5
1
1.5
2
2.5
EFFAP3

Figure 8-6

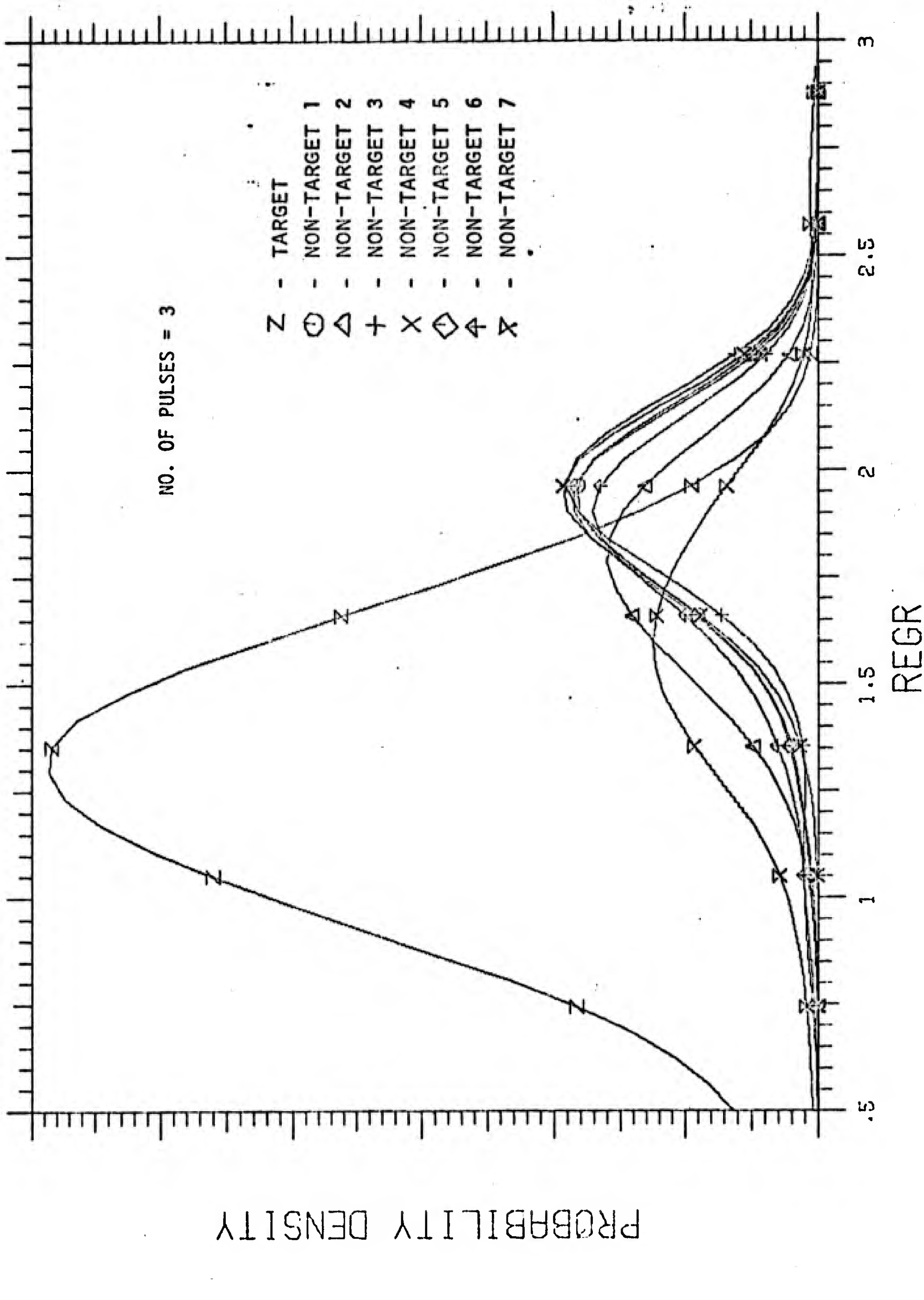

NO. OF PULSES = 3
TARGET
NON-TARGET 1
NON-TARGET 2
NON-TARGET 3
NON-TARGET 4
NON-TARGET 5
NON-TARGET 6
NON-TARGET 7
.5
1
1.5
2
2.5
3
REGR
PROBABILITY DENSITY

Figure 8-7

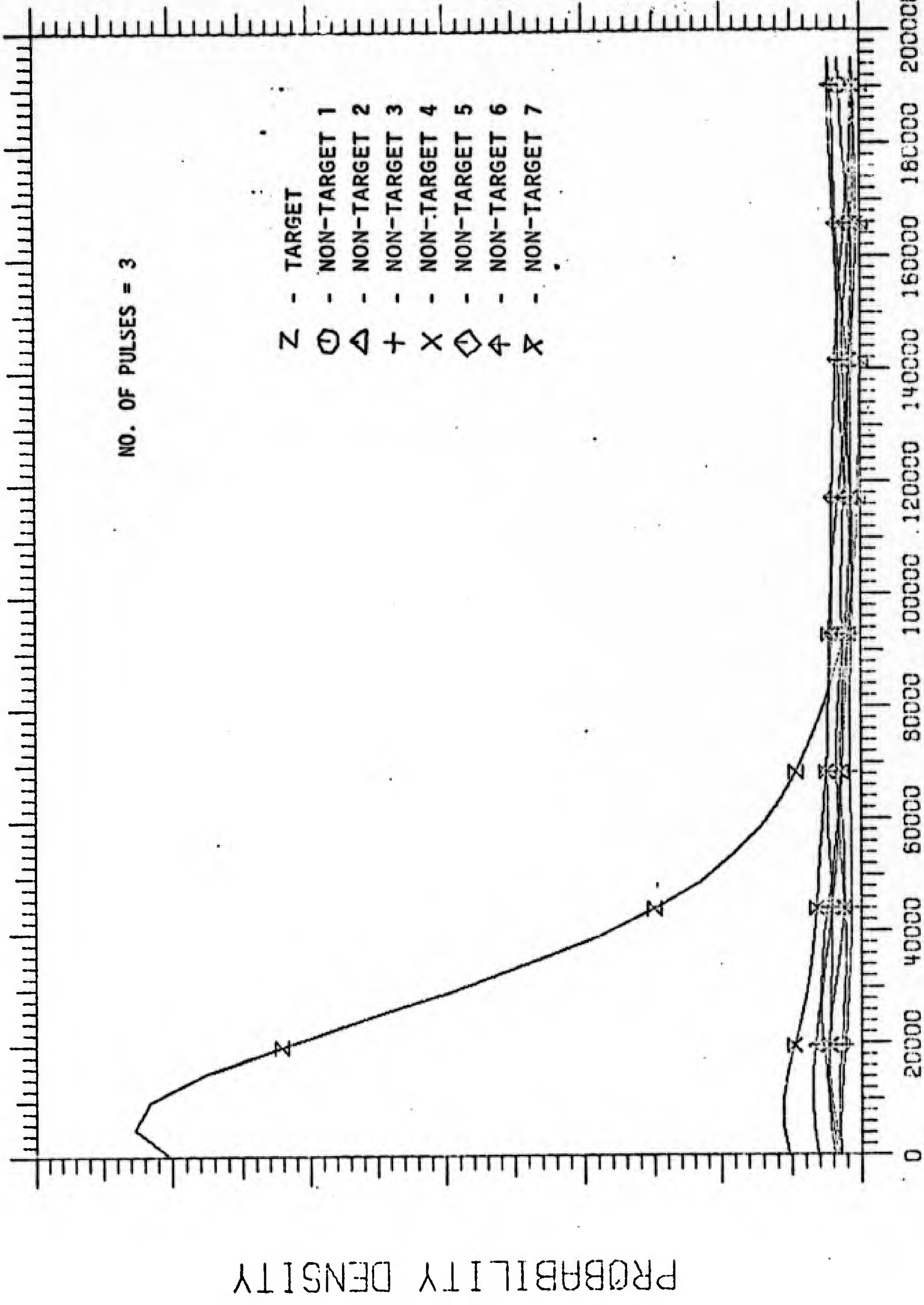

NO. OF PULSES = 3
Z - TARGET
- NON-TARGET 1
- NON-TARGET 2
- NON-TARGET 3
- NON-TARGET 4
- NON-TARGET 5
- NON-TARGET 6
- NON-TARGET 7
0
20000
40000
60000
80000
100000
120000
140000
160000
180000
200000
PP4M
PROBABILITY DENSITY

Figure 8-8

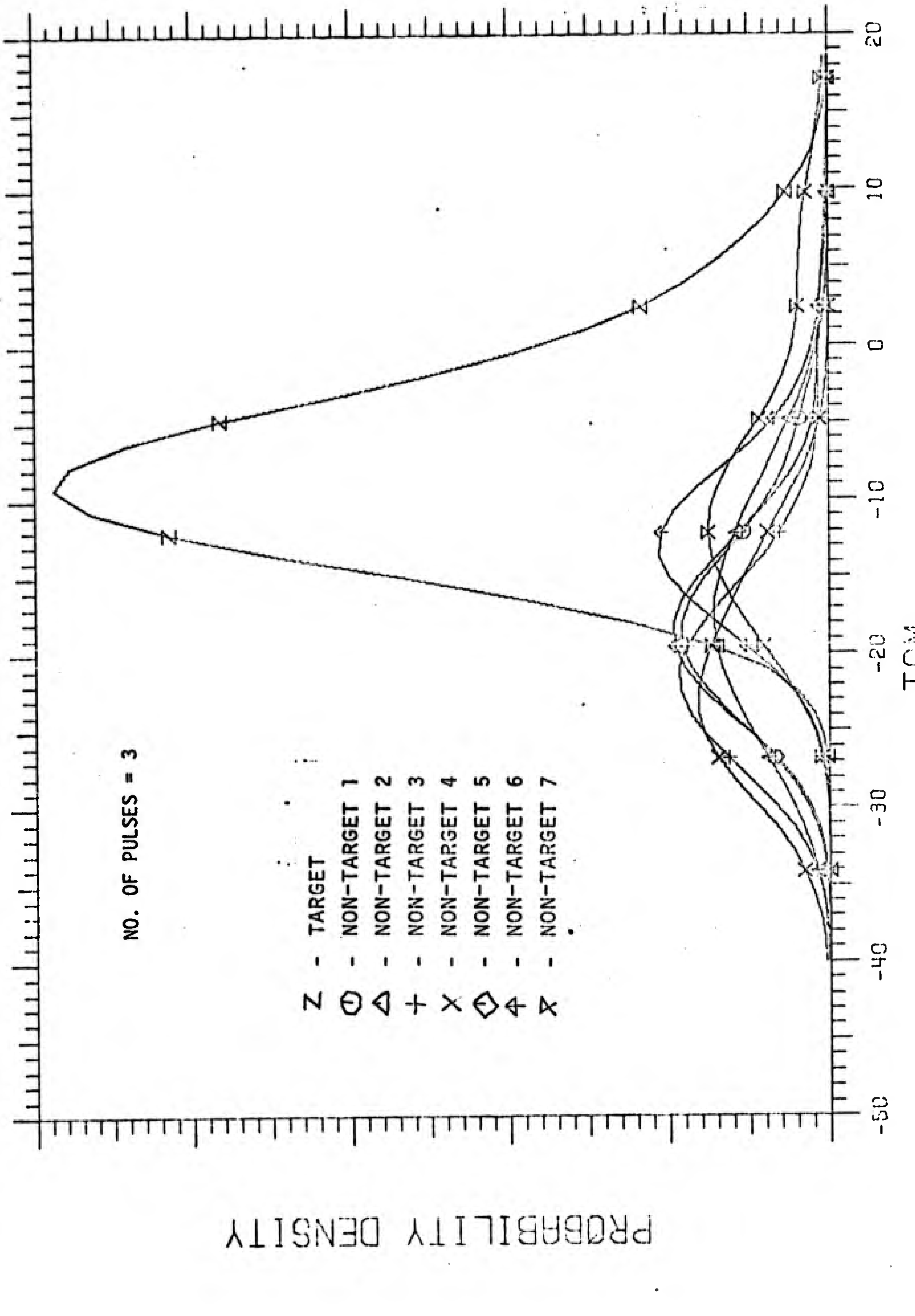
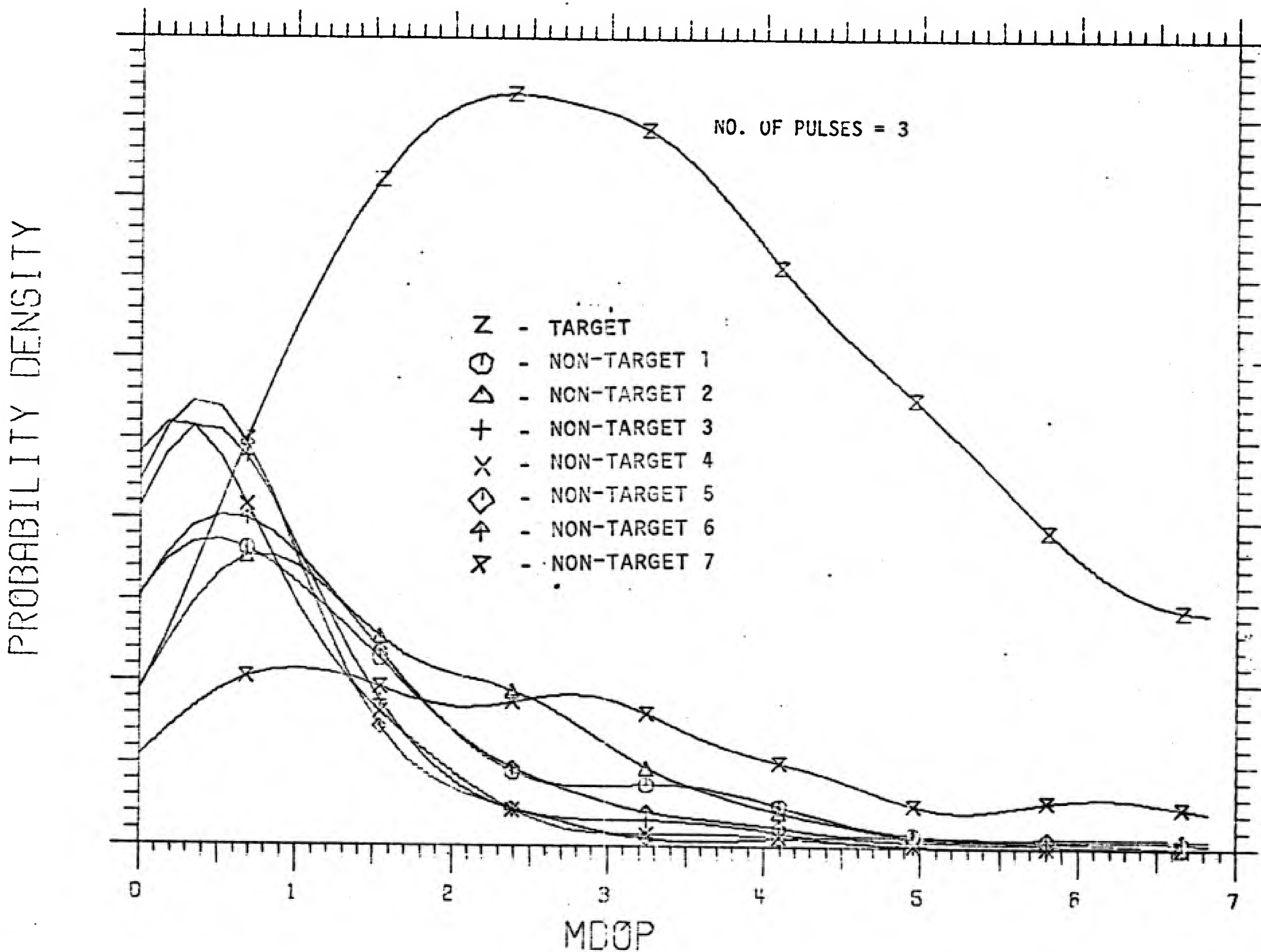

NO. OF PULSES = 3
Z - TARGET
- NON-TARGET 1
- NON-TARGET 2
+ - NON-TARGET 3
X - NON-TARGET 4
- NON-TARGET 5
- NON-TARGET 6
- NON-TARGET 7
-50
-40
-30
-20
-10
0
10
20
TGM
PROBABILITY DENSITY

Figure 8-9



NO. OF PULSES = 3
Z - TARGET
NON-TARGET 1
NON-TARGET 2
NON-TARGET 3
NON-TARGET 4
NON-TARGET 5
NON-TARGET 6
NON-TARGET 7
PROBABILITY DENSITY
MDØP
0
1
2
3
4
5
6
7

Figure 8-10

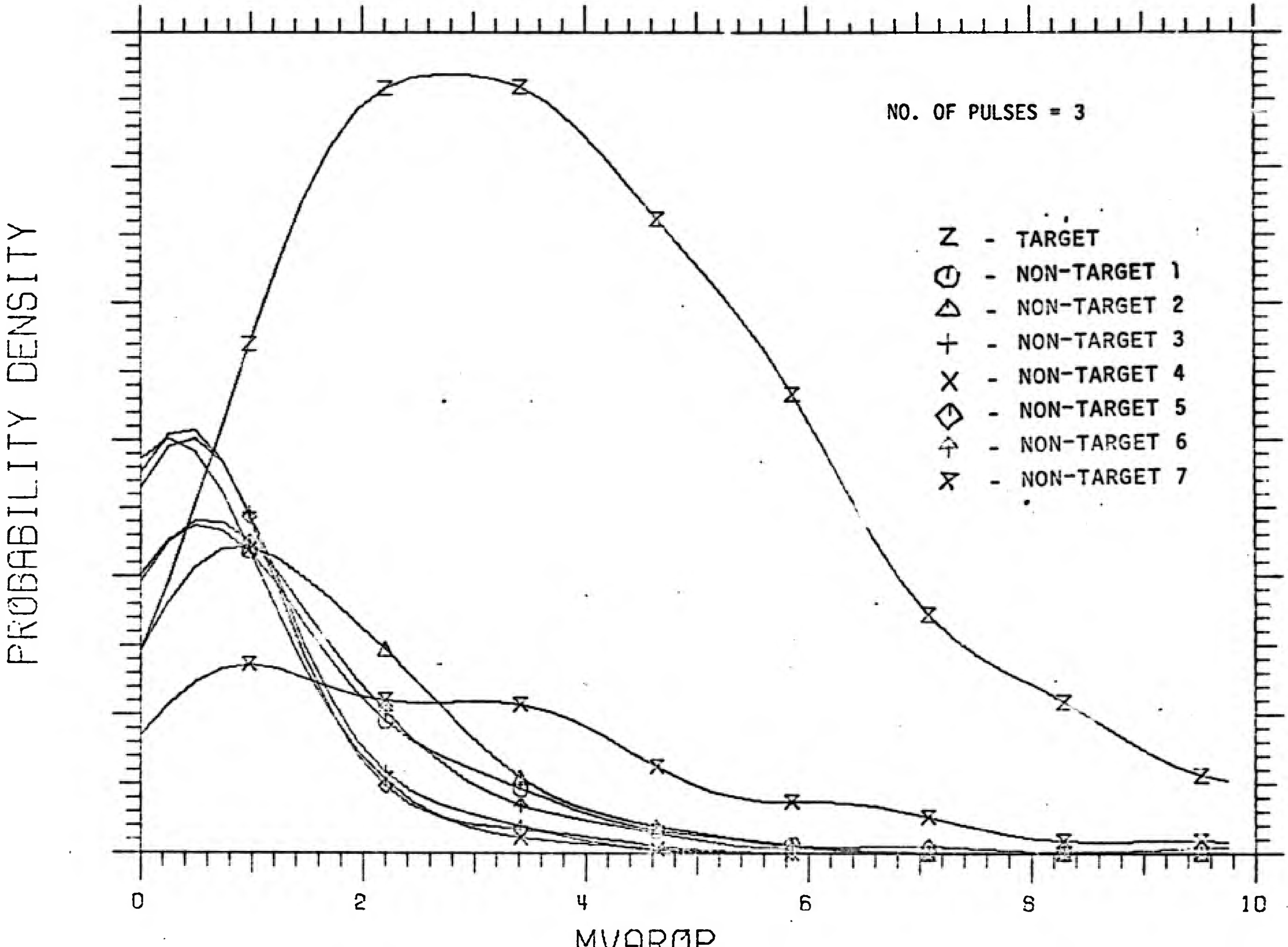

NO. OF PULSES = 3
Z - TARGET
- NON-TARGET 1
- NON-TARGET 2
+ - NON-TARGET 3
X - NON-TARGET 4
- NON-TARGET 5
- NON-TARGET 6
- NON-TARGET 7
PROBABILITY DENSITY
0
2
4
6
8
10
MVARØP

Figure 8-11

all features; several others, such as class 2 and class 6 are problems for specific features. The bivariate scatter plots in Sections 5.0 and 7.0 provide visual aids to the examination of the overlap for the two-feature case. More generally, the specific subclasses contributing to false alarms and to missed detections can be explicitly listed as part of the discrimination analysis (see Table 8-1 and Figure 8-1).

Figures 8-12 and 8-13 indicate the degradation in performance for a single feature (RECR) and a pair of features (TGM/REGR) if the non-target class consisted of only sublcass 7, the most problematical sub-class. Substantial degradation occurs.

Finally, we note that the individual signatures which cause false alarms and missed detections can be identified. Individual signatures are plotted in figures 8-14 through 8-23 with an indication of their correct class and the class into which they were placed by the decision rule. Included are correct classifications, false alarms, and missed detections. Note that <u>all</u> 200 pulses in the data are plotted, but only the first 15 were used for discrimination. Included with each plot is a list of the model initial conditions which generated the signature; systematic examination of all the missed detections or false alarms would indicate the particular characteristics of motion, aspect, etc., which cause discrimination problems.

		DECIDED CLASS	
	TRUE CLASS	TARGET	NON-TARGET
Z	TARGET	391	9
⊙	NON-TARGET 1	4	76
△	NON-TARGET 2	7	73
+	NON-TARGET 3	4	76
×	NON-TARGET 4	0	80
◇	NON-TARGET 5	4	76
↑	NON-TARGET 6	0	80
⧗	NON-TARGET 7	24	56

	DECIDED CLASS	
TRUE CLASS	TARGET	NON-TARGET
TARGET	391	9
NON-TARGET	43	517

NON-TARGET CLASS	PERCENT OF 43 FALSE ALARMS CAUSED BY NON-TARGET CLASS
1	9.30
2	16.28
3	9.30
4	0.00
5	9.30
6	0.00
7	55.81

Table 8-1: Breakdown of false alarms by non-target subclass for feature REGR (refer to figure 8-1).

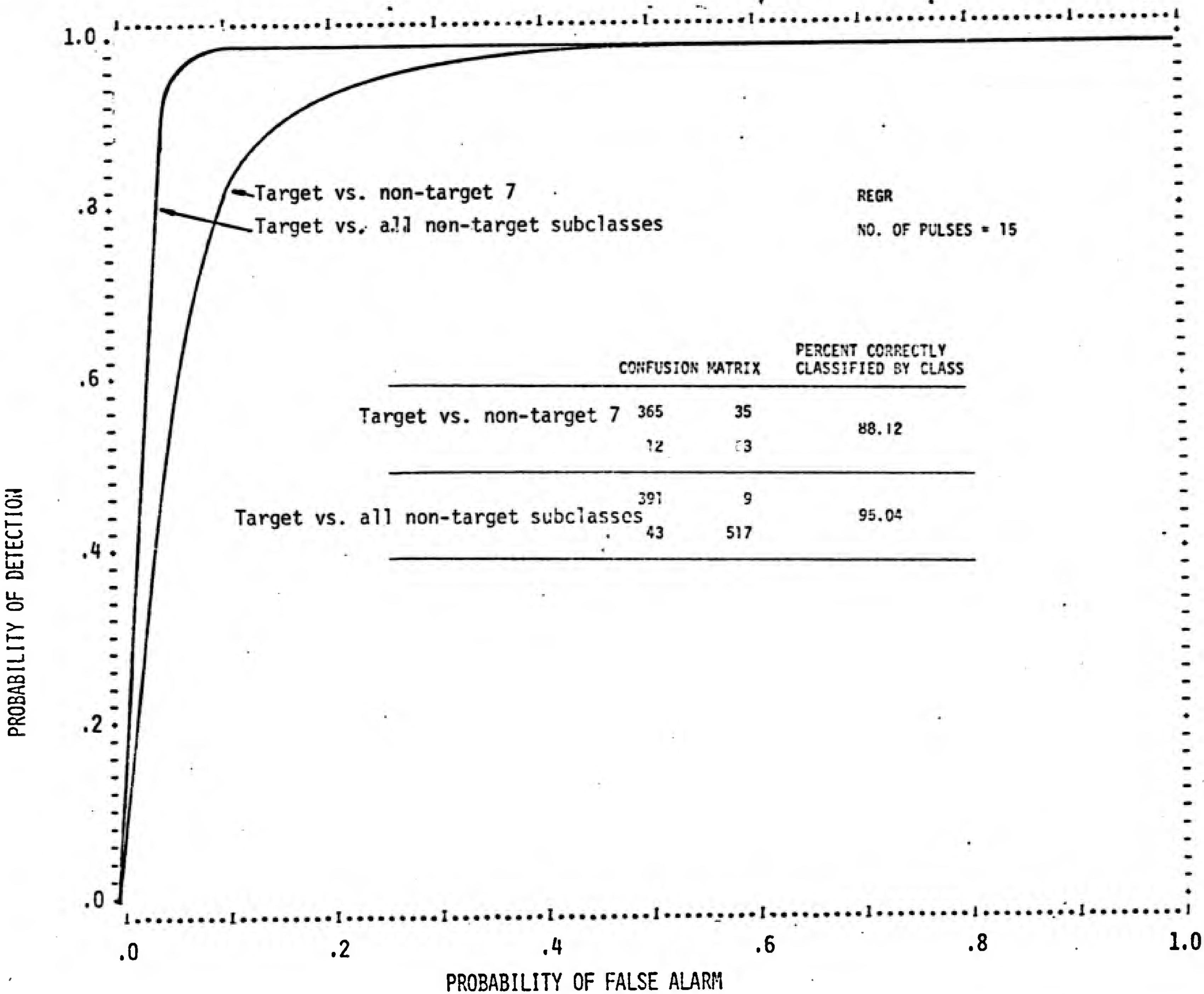
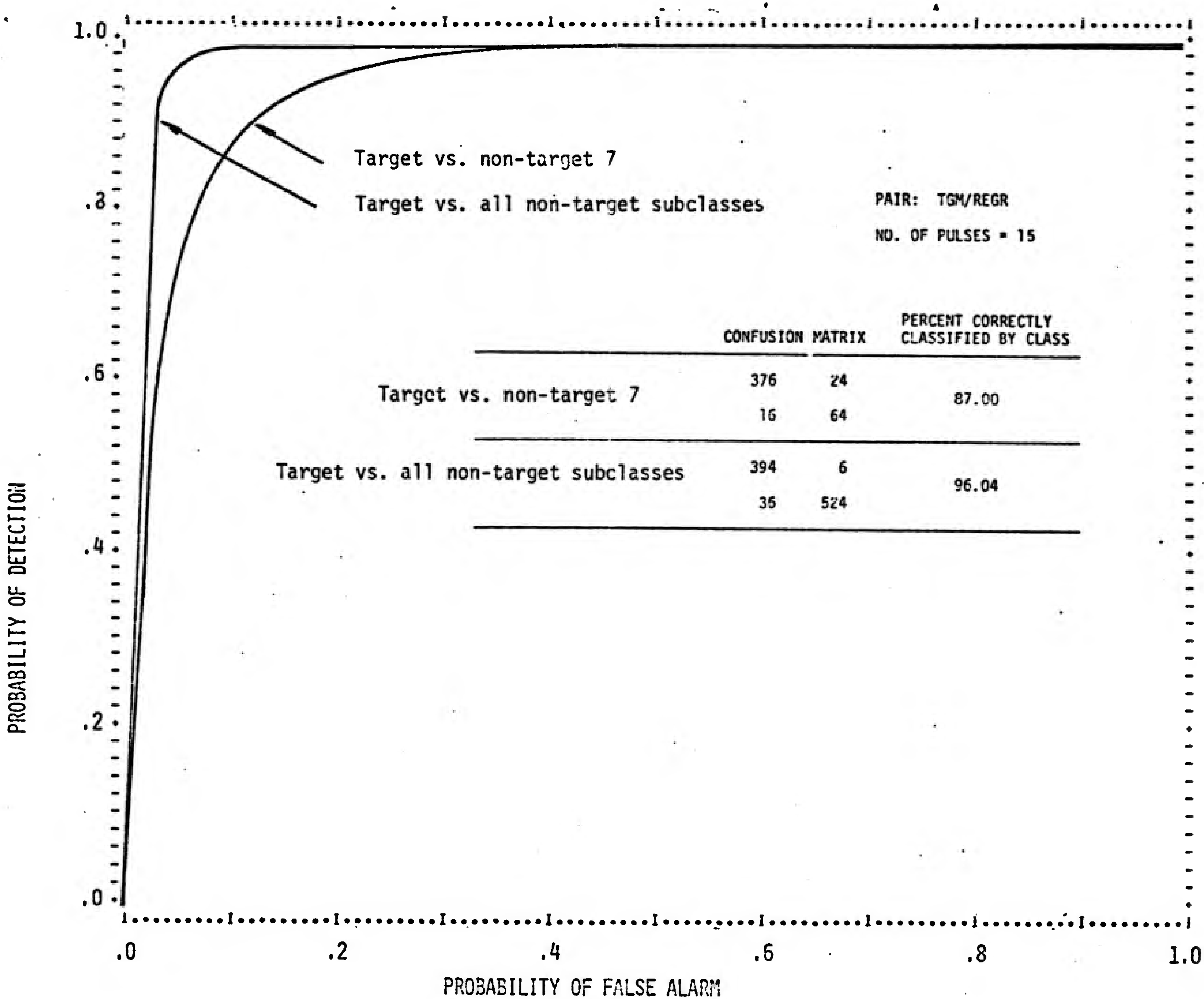
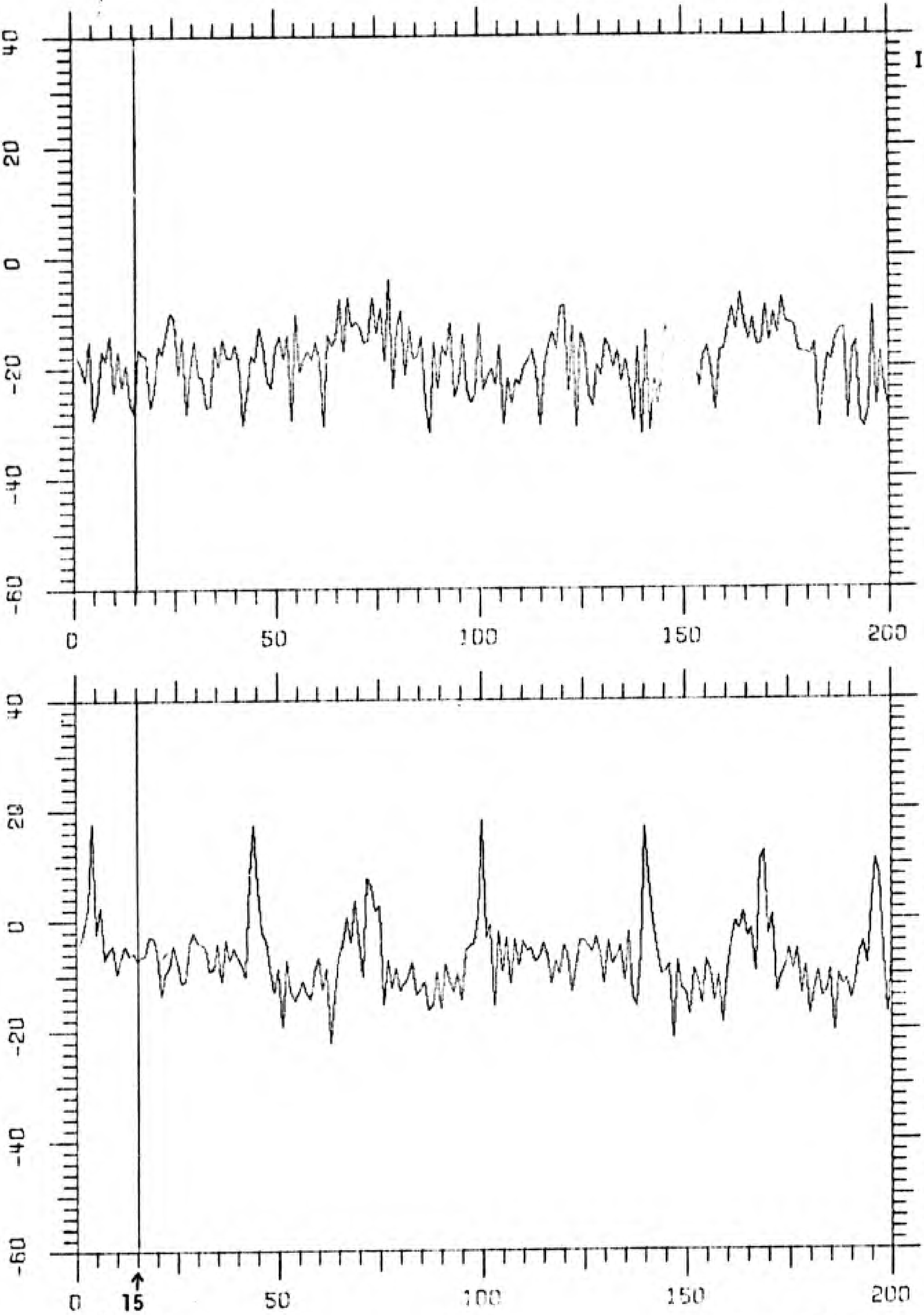

1.0
.8
.6
.4
.2
.0
.0
.2
.4
.6
.8
1.0
PROBABILITY OF DETECTION
PROBABILITY OF FALSE ALARM
Target vs. non-target 7
Target vs. all non-target subclasses
REGR
NO. OF PULSES = 15
CONFUSION MATRIX
PERCENT CORRECTLY CLASSIFIED BY CLASS
Target vs. non-target 7
365 35
12 3
88.12
Target vs. all non-target subclasses
391 9
43 517
95.04

Figure 8-12



	CONFUSION MATRIX		PERCENT CORRECTLY CLASSIFIED BY CLASS
Target vs. non-target 7	376	24	87.00
	16	64	
Target vs. all non-target subclasses	394	6	96.04
	36	524	

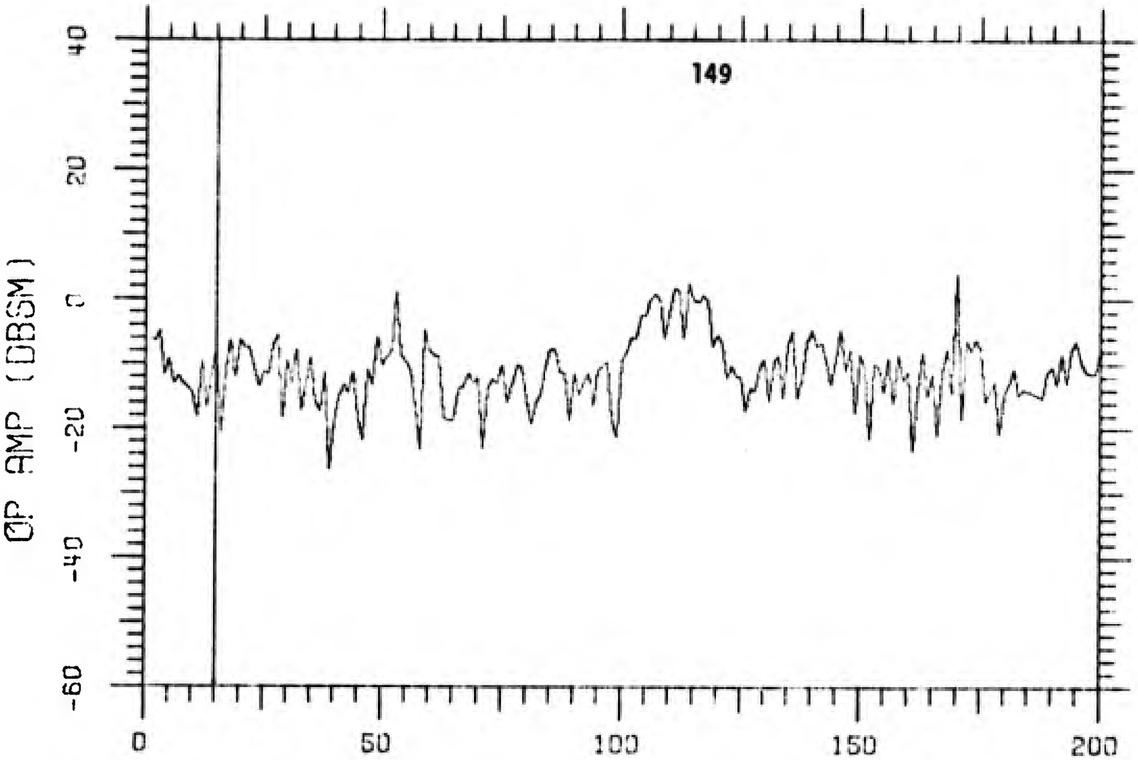
Figure 8-13



I.C. = 273.5625,
7.7b1987,
24.38437,
1.306561,
0.0,
0.0,
9.95,
0.05,
0.0,
0.25,
0.2b

Figure 8-14

Actual: Target Decided: Target



I.C. = 95.56532,
156.8586,
178.8647,
1.011846,
0.368838,
0.0061137,
9.95,
0.05,
0.0,
0.25,
0.25

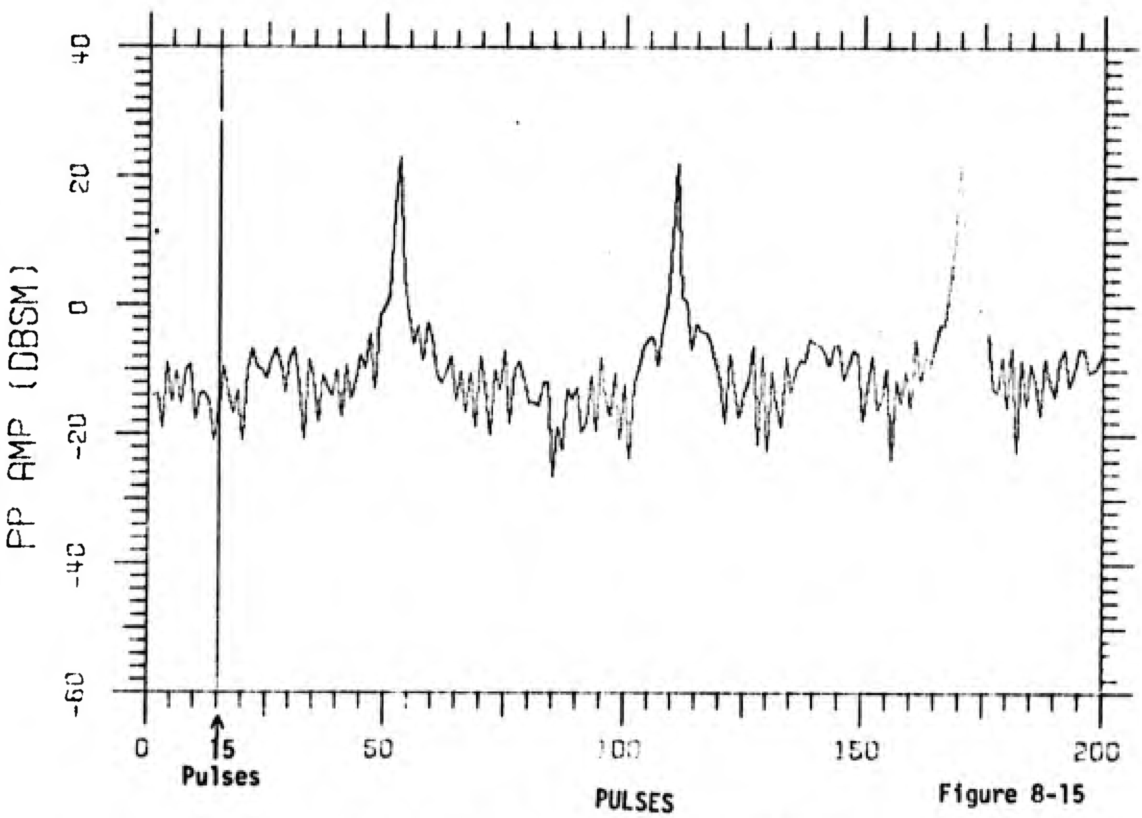


Figure 8-15

Actual: Non-target 7 Decided: Non-target

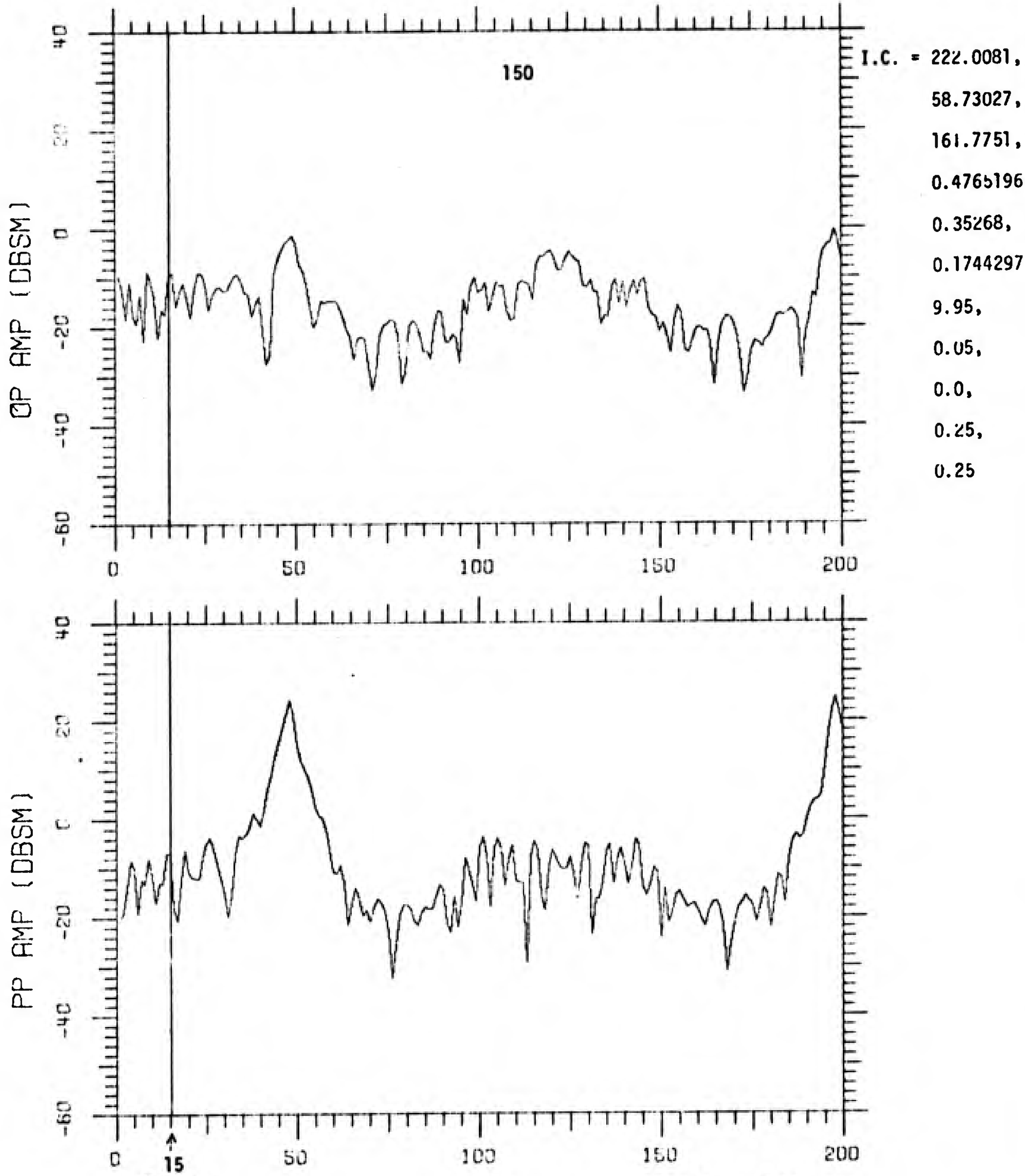


Figure 8-16

Actual: Non-target 7 Decided: Target

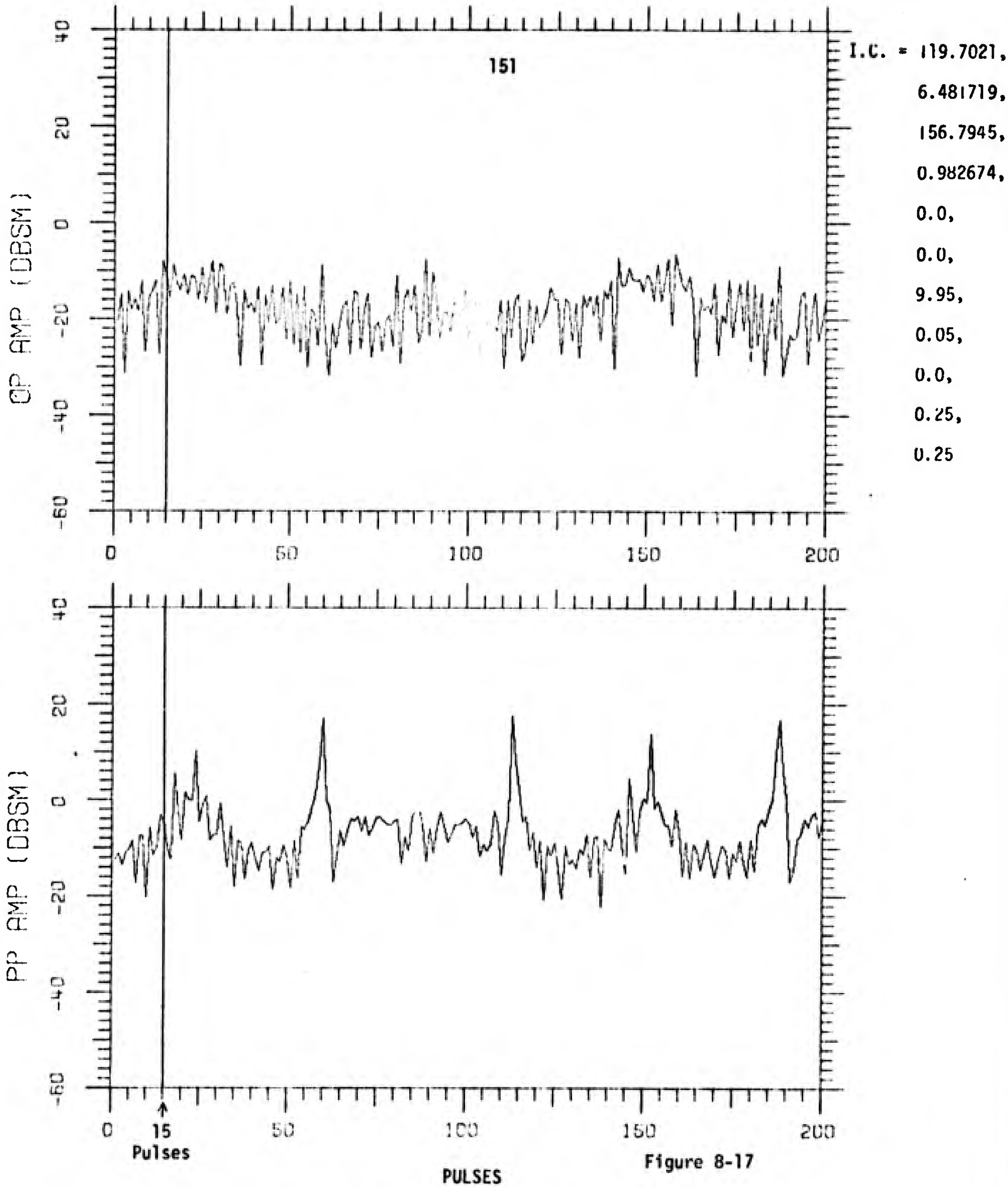


Figure 8-17

Actual: Target Decided: Target

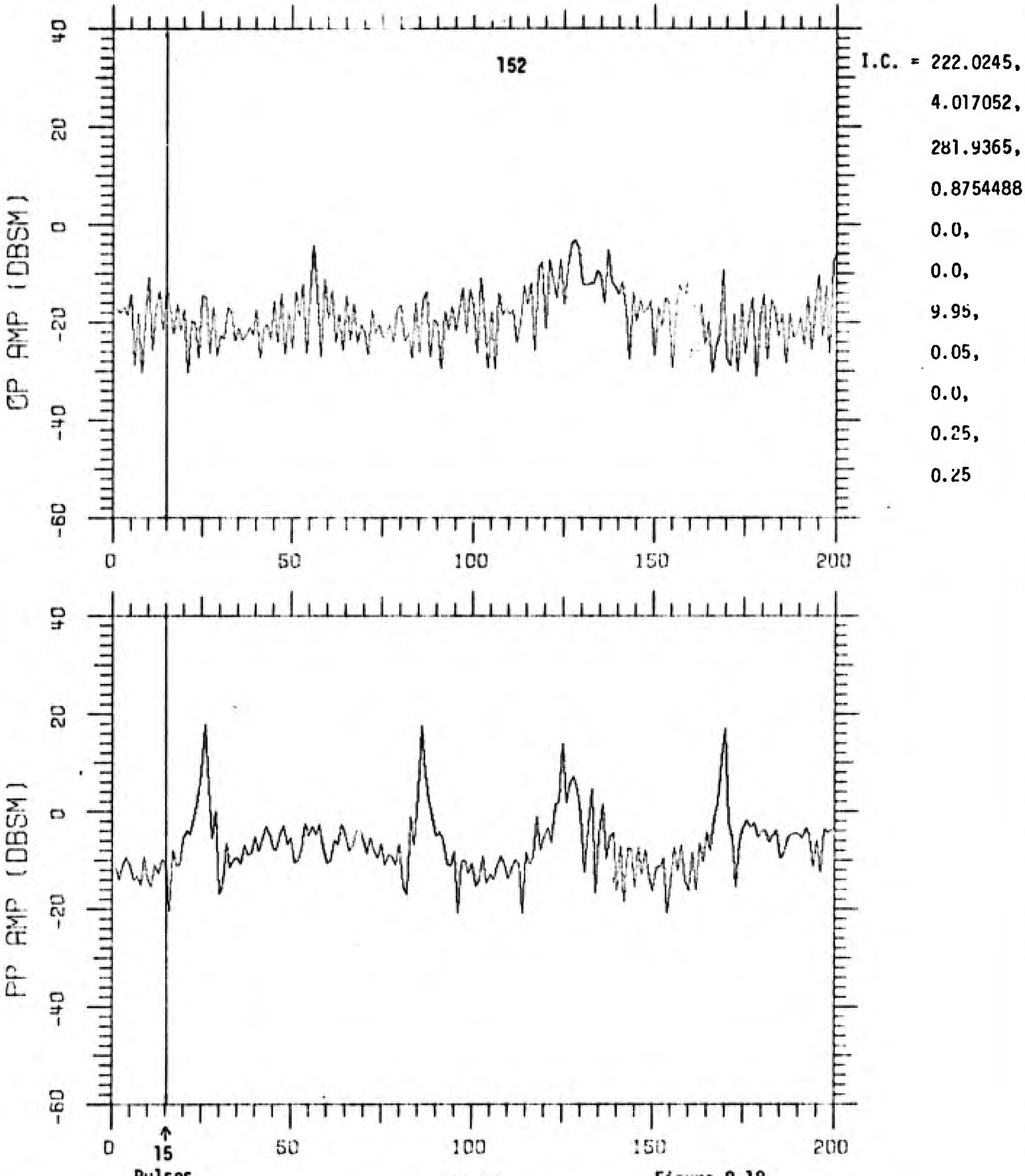


Figure 8-18

Actual: Target   Decided: Target

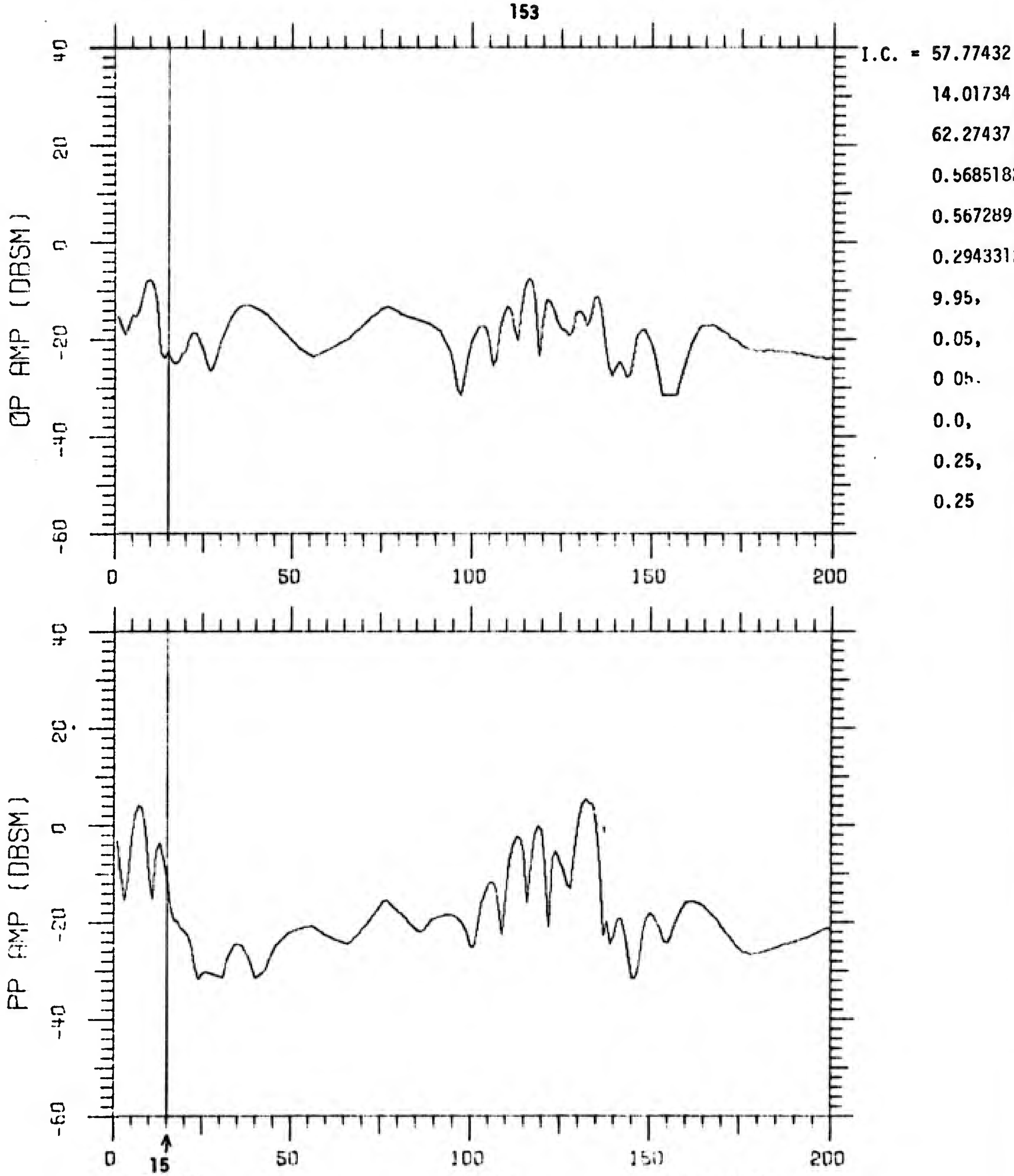


Figure 8-19

Actual: Non-Target 1 Decided: Target

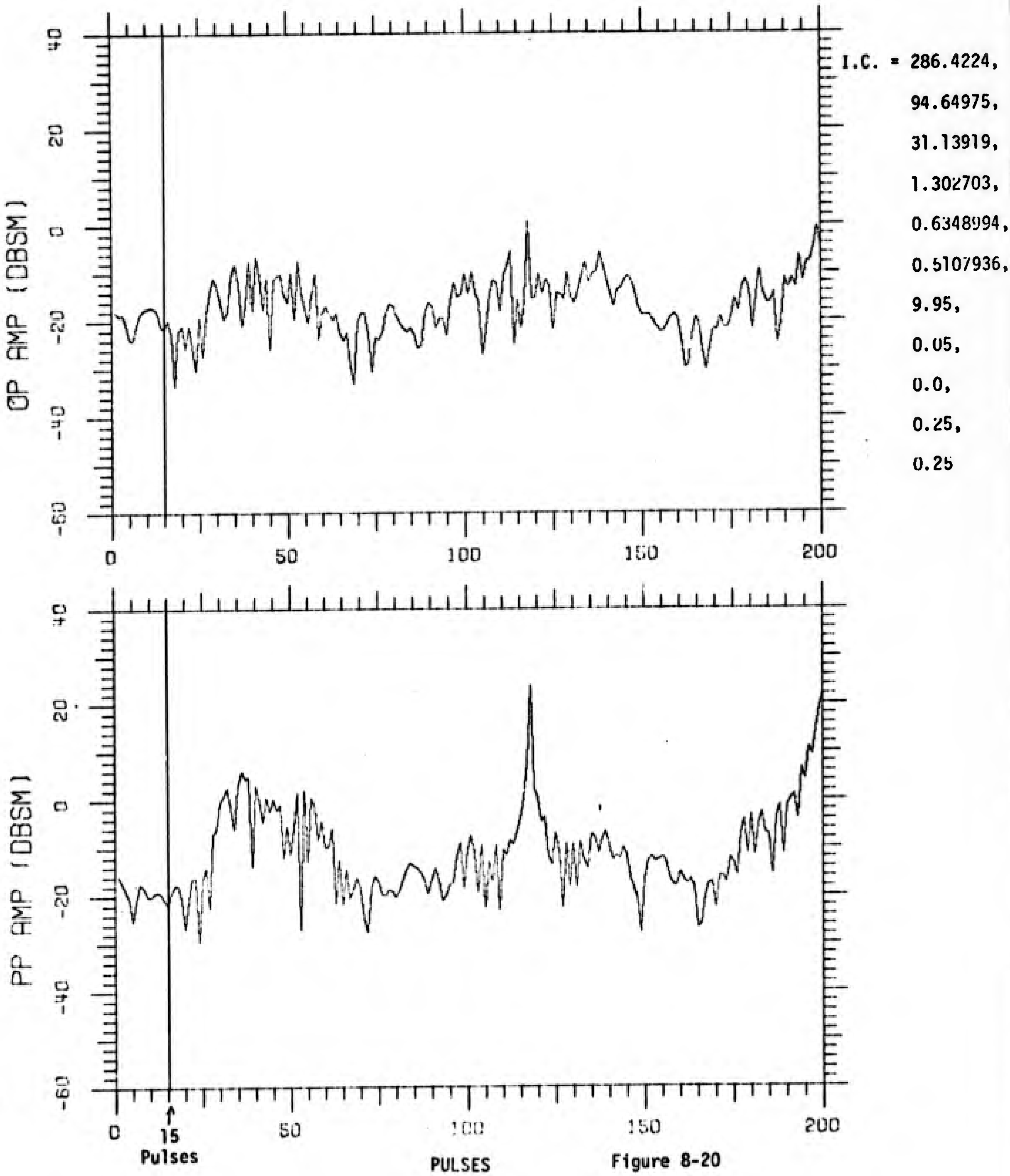


Figure 8-20

Actual: Non-target 7 Decided: Non-target

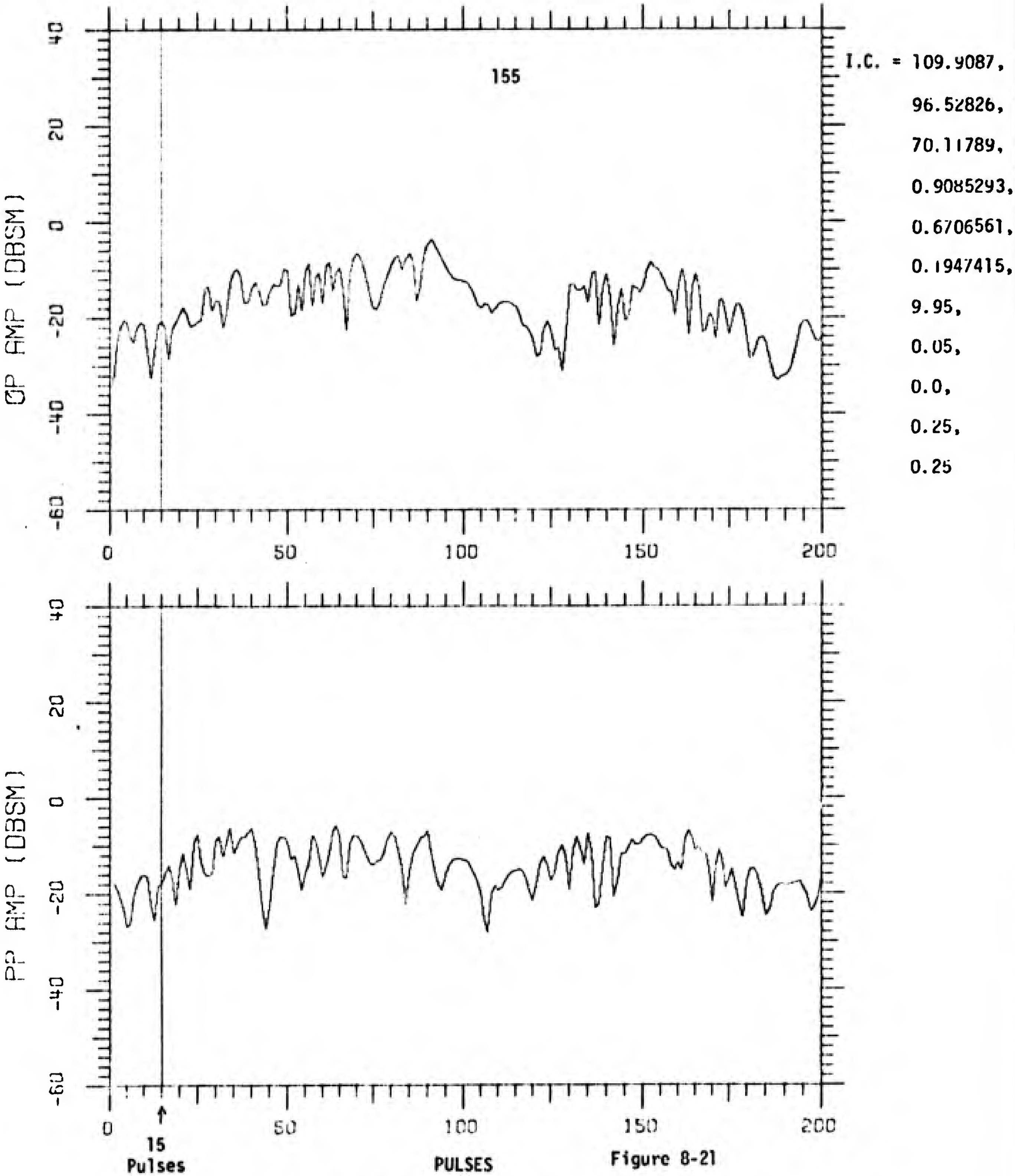


Figure 8-21

Actual: Non-target 7 Decided: Non-target

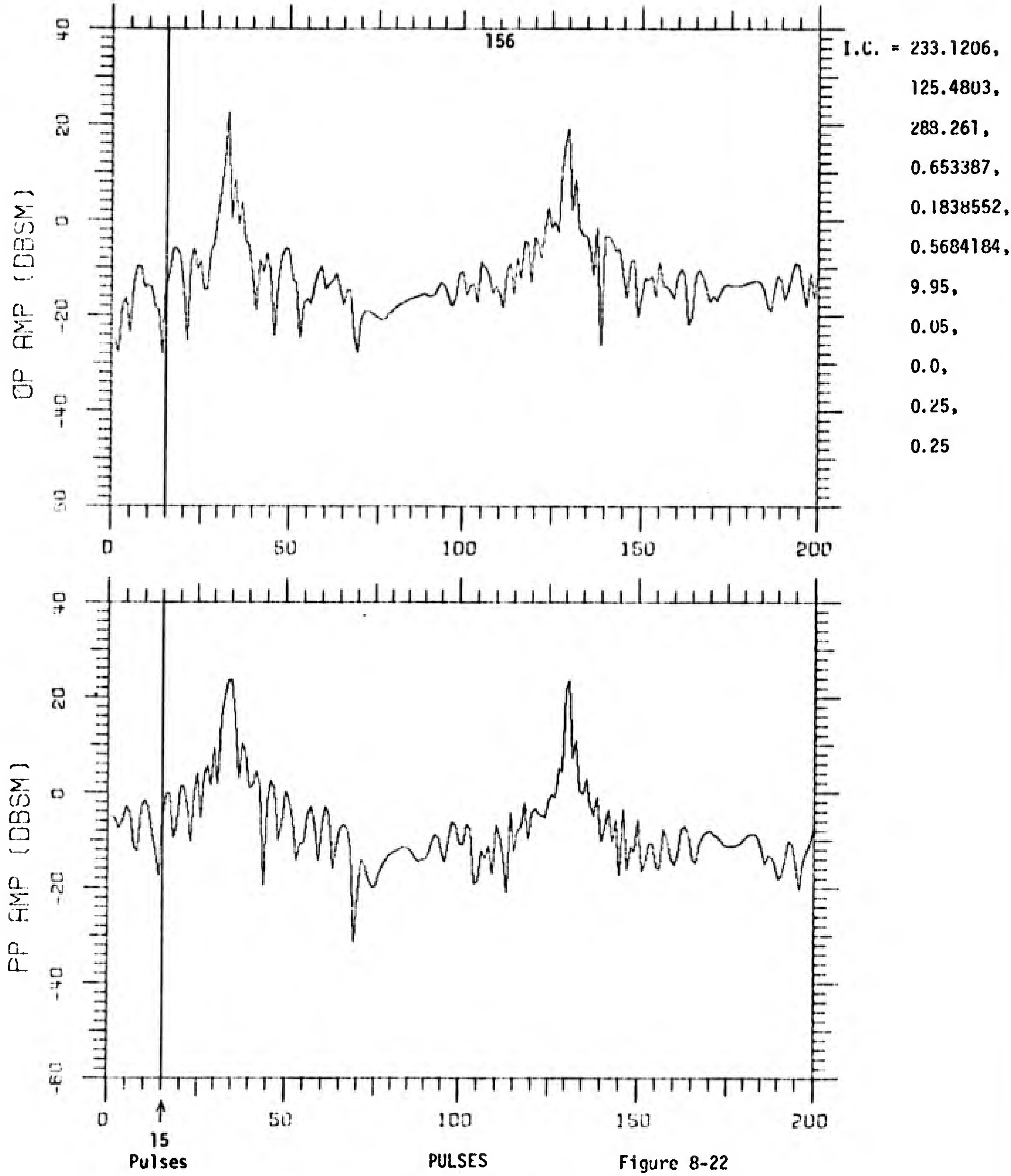
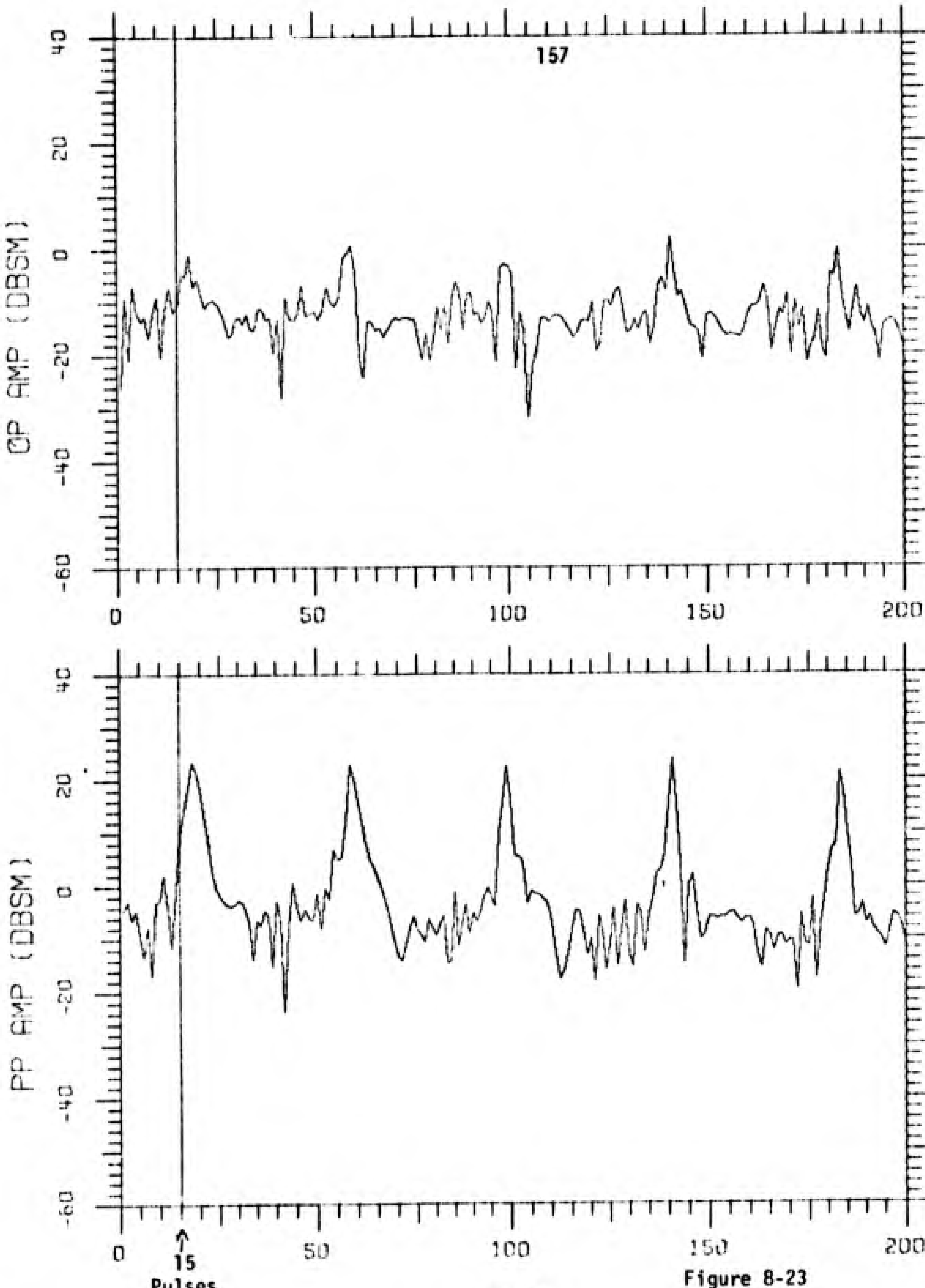


Figure 8-22

Actual: Non-Target 7 Decided: Target



I.C. = 113.1392,
17.12646,
158.5994,
0.2689505,
0.676381,
1.414248,
9.95
0.05,
0.0,
0.25,
0.25

Figure 8-23

Actual: Non-target 7   Decided: Target

9.0 ADDITIONAL TOPICS

As mentioned previously in Section 2.0, the level of effort of this study was limited and thus the objectives of the study were subject to this constraint. However, despite this, some (perforce very limited) analysis effort was applied to determining the impact on the results of attacking the overall problem in further depth. The section discusses some of these analyses and indicates some ways the data analysis approach described in the report could be applied in other contexts to achieve the objectives described in Section 2.0.

9.1 FURTHER DEVELOPMENT OF OBJECTIVE FEATURES

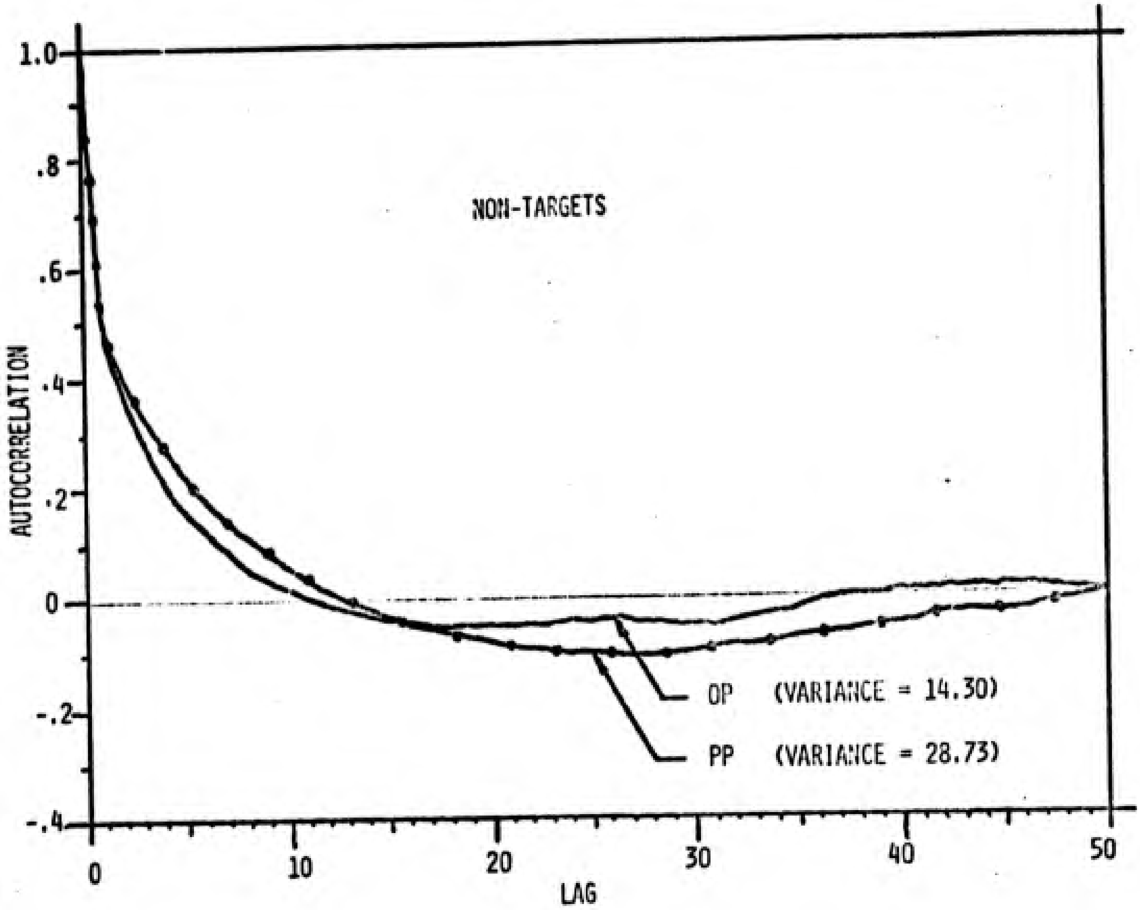
Further features and sets of features should be derived, combined, and ranked in the following contexts:

(1) Range of Numbers of Pulses

The feature ranking described in this report was largely performed at N=15 pulses. Earlier analysis described in Section 7.0 indicates that the ranking of features changes as the number of pulses changes. A "best" set of features for various numbers of pulses (e.g., N=3, N=6, N=10, N=20) might be derived.

(2) Variations in Interpulse Spacing

The pulses selected to form the raw data for the feature creation step were equally spaced in time at an effective PRF (pulse repetition frequency) of 20 cycles per second. The autocorrelations shown in figure 9.1 indicates that the effective PRF should be approximately



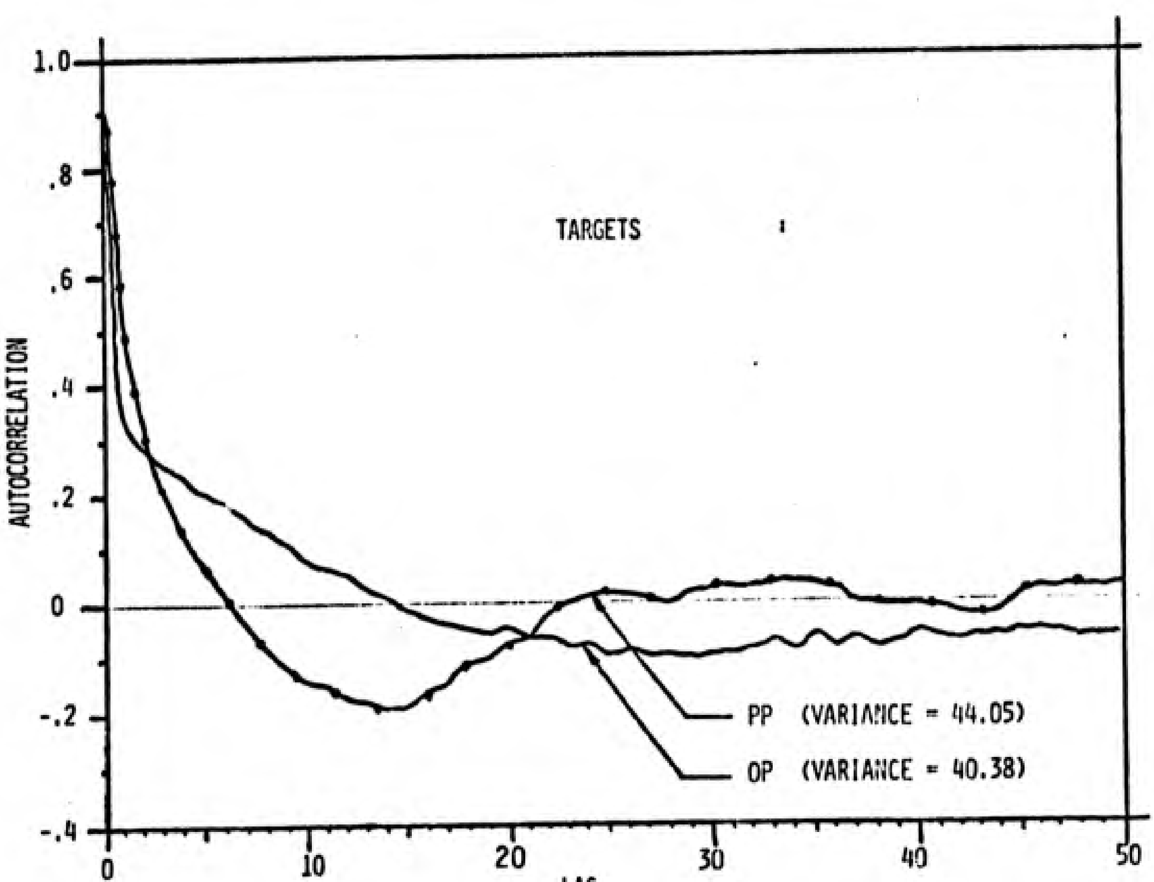


Figure 9.1: Autocorrelation of Target and Non-Target Signatures

2 cycles per second to get fairly uncorrelated data. Additional analysis might indicate the degree to which discrimination performance could be improved by using a larger interpulse spacing and how different interpulse spacings affect the feature ranking results for various numbers of pulses.

(3) Improved Performance on "Difficult" Non-Target Objects

A more detailed analysis of the non-target objects in an object-by-object fashion may lead to features which, though low ranked in the two-class context, are highly ranked for the discrimination of certain "difficult" cases. Additionally, new objects which may have more realistic (and possibly more difficult) characteristics might be modeled and discrimination features developed for them.

(4) Utilization of Phase Data

All of the analysis in this report indicates results for sensor data which contained no phase information. Features might be developed which use this information and which provide enhanced performance.

(5) Derivation (Optimization) of Objective Features in a Multiple Feature Environment

In this report objective features were derived by effectively optimizing a measure of the separation of the scalar or one-dimensional class conditional probability densities. The one-dimensional objective features so derived were then combined with other features to form the best group of dimension two, three, etc.

One can anticipate improved performance if one performs the derivation of an objective feature (to be used for example as one member of a pair of features) in the context of other features. That is, the optimization of the measure of the separation of the class conditional probability densities might be performed in the higher-dimensional space formed by the objective feature and the other features of the group.

(6) Systematic Derivation of Piecewise Linear Transformations of Sets of Heuristic Features

Additional study might be profitably focused on the development of objective features created by deriving piecewise-linear transformations of sets of heuristic features. Techniques for determining the correct selection (and dimension) of the input heuristic features and the correct dimension of the output piecewise linear feature space should be employed.

(7) Real Data

The techniques described in the report are applicable to real (rather than simulated) data. Problems exist in terms of the number of samples available but these can often be overcome by the techniques described in the body of the report which can result in low dimensional decision algorithms.

10.0 CONCLUSION

This report has indicated the versatility of the multivariate empirical approach and of objective feature extraction procedures in (1) determining the possible level of discrimination of target versus non-target classes based on radar cross section signatures, (2) extracting the key characteristics of the signature which allow discrimination, (3) indicating the impact of the number of pulses utilized on the performance of the discrimination algorithm, and (4) isolating the key characteristics which confound discrimination. The many plots and tables provide a wealth of insight into the above aspects of the discrimination problem. The report has demonstrated a systematic approach to the rapid, cost effective investigation of such problems in general and has provided the data for deeper analysis of the given target/non-target classes in particular.

The aspects of the problem analyzed include the following:

(1) the possible level of discrimination of the given classes (PD/PFA curves, error rates, etc.) for the cases of 200, 15, 6, and 3 pulses, with emphasis on the 15- and 3-pulse cases;

(2) the features (objective characteristics of the signals) which best discriminate singly and jointly, their distributions and other attributes;

(3) the subclasses of non-targets which posed the most difficult discrimination problem, and the particular signatures which tended to be misclassified; and

(4) a brief discussion of such additional topics as the effect of using less correlated pulses.

REFERENCES

1. Meisel W. S., Computer-Oriented Approaches to Pattern Recognition, Academic Press, 1972.

2. Meisel, W. S., "The Efficient Representation of Functions for Pattern Clasification and Feature Selection," Proc. 1971 Systems, Man, and Cybernetics Conference, Joint National Conference on Major Systems, October 1971, pp. 51-57.

3. Collins, D. C., and Meisel, W. S., "A General Approach to Objective Feature Creation." Paper in preparation.

4. Collins, D. C., and Meisel, W. S., "The Detailed Discrimination of Radar Signatures through Hybrid Analysis," TSC-097, November 1972.

5. Meisel, W.S., and Collins, D.C., "Repro-modeling: An Approach to Efficient Model Utilization and Interpretation," IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-3, No. 4, July 1973, pp. 349-358.

6. Fukunaga, K., Introduction to Statistical Pattern Recognition, Academic Press, 1972.