

FINAL REPORT

Improved Analysis Algorithms for UXO Filler Identification

SERDP Project MM-1383

FEBRUARY 2009

Robert Holslin
Robert Sullivan
SAIC

This document has been approved for public release.



Strategic Environmental Research and
Development Program

EXECUTIVE SUMMARY

SAIC and Duke University investigated several advanced data analysis algorithms and techniques to apply to Pulsed Elemental Analysis with Neutrons (PELAN) data. The data was collected on shells filled with inert and explosive materials, and chemical simulants. The goal of this investigation was to improve the performance, reliability, and robustness of the PELAN decision-making process and to make it easier to train the PELAN system using target libraries. These studies have provided considerable improvements in performance over the previous methods used to analyze the PELAN spectra and the decision-making process. The focus of this effort concentrated on unexploded ordnance (UXO) items. Both the Least Squares/Generalized Likelihood Ratio Test (LS/GLRT) and the Least Squares/Principal Component Analysis (LS/PCA) combinations showed significant improvement over the LS/decision tree approach. As a result, the LS/GLRT method was implemented into the portable PELAN unit. We are continuing investigation into the PCA spectral analysis method, which shows even more improvement over the LS/GLRT approach. The PCA algorithm was shown to be effective at using the entire spectrum to extract characteristics of the target for improved identification.

During this project, several key results were established and are summarized below:

- GLRT was established as an effective decision making algorithm.
 - GLRT can be used in conjunction with LS, PCA, and other spectral analysis techniques (e.g., Multiple Signal Classification, MUSIC).
- The tertiary declaration was added for GLRT decision making (“unknown”) for explosives/inert-filled shells.
- PCA can perform better than Least Squares on shell targets.
- Background measurements may not be necessary for effective PCA analysis.
 - The need for empty shell in background is eliminated.
 - Less user input is required for recording the environment.
 - Further testing is required for verification.
- The GLRT, tertiary declaration, and entropy-based confidence level were implemented into PELAN Non-invasive Filler Identifier (NFI) systems.
- Data Collection at Indian Head was conducted Dec 6-22, 2004.
 - Strategic Environmental Research and Development Program data was made available to SAIC and Duke University in February 2005.

SAIC, in collaboration with Duke University and Environmental Chemical Corporation (ECC), have been selected by ESTCP to build, test, demonstrate and validate a mobile, multi-detector-based PELAN unit for the classification of UXO filler at cleanup sites. With improved classification algorithms, we can improve the reliability of the target analysis, improve the performance, and, thus, provide a cheaper and safer means for environmental remediation needs and other explosive ordnance disposal-related efforts. The improved spectral analysis and decision-making algorithms developed in this project will be implemented directly into the current PELAN systems. Along these lines, we have already implemented and started testing the tertiary GLRT and entropy-based confidence algorithms in the PELAN IV system.

TABLE OF CONTENTS

EXECUTIVE SUMMARY	ii
TABLE OF FIGURES	iv
TABLE OF TABLES	v
ACRONYMS	vii
1. INTRODUCTION	9
1.1 Project Background	9
1.2 Objective	9
1.3 Technical Approach	11
1.3.1 Technical Description	11
1.3.2 The PELAN System	11
1.3.3 Earlier Spectral Analysis and Decision-making Methods	12
1.3.4 Evaluation of New Analysis Techniques	13
1.4 Summary of Results	14
2. PROJECT ACCOMPLISHMENTS	14
2.1 Background	14
2.1.1 PCA for Spectral Analysis	15
2.1.2 GLRT for Decision Making	16
2.1.3 Preliminary Results	17
2.2 Data Collection	19
2.3 Normalization Studies	20
2.3.1 Data Transformation	20
2.3.2 GLRT Target Grouping and Setup	21
2.3.3 Results	37
2.3.4 Conclusions	38
2.4 Confidence Metrics	38
2.4.1 Entropy-based Confidence Metric	38
2.4.2 Experimental Data Results	40
2.5 Spectral Analysis With PCA	44
2.5.1 Effects of Background Subtraction	44
2.5.2 Variables Affecting Cluster Formation	47
2.6 Spectral Estimation	71
2.7. Processing of PELAN IV Data	73
2.7.1. Introduction	73
2.7.2. Analysis Algorithms	75
2.7.3. Performance Results	76
2.7.4. Performance Results of December 2004 Data	80
2.7.5. Comparison with Neural Net Analysis	81
2.7.6. Summary	86
2.8. Implementation in PELAN IV	86
3. CONCLUSIONS	87
4. TRANSITION PLAN AND RECOMMENDATIONS	88
APPENDIX A: Test Plan	90

TABLE OF FIGURES

Figure 1.3-1 The PELAN III (left) and PELAN IV systems. The PELAN IV system was used for tests at Indian Head in December 2004.....	12
Figure 2.2-1. PELAN IV shown inspecting ordnance in soil test box at Indian Head in December 2004.....	20
Figure 2.4-1: Identification results for ML estimates of fill material when the correlation between the element counts is included in the algorithm.....	41
Figure 2.4-2: Identification results for ML estimate of fill material when the element counts are assumed to be uncorrelated in the algorithm.	42
Figure 2.4-3: Identification results for ML estimate of fill material when the correlation between the element counts is included in the algorithm.....	43
Figure 2.4-4: Identification results for ML estimate of fill material when the element counts are assumed to be uncorrelated in the algorithm.	43
Figure 2.5.1-1. ROC of PCA trained with data taken on a table, then tested on a table. The left ROC plot is with no background subtraction, and the right ROC plot is with background subtracted.	46
Figure 2.5.1-2. ROC of PCA trained with data taken on a table then tested on sand, asphalt, and soil. The left ROC plot is with no background subtraction, and the right ROC plot is with background subtracted.	47
Figure 2.5.2 – 1: Stem plot of the first three principal components for large shells on any background. There is no preprocessing of the data.....	52
Figure 2.5.2 – 2: Stem plot of the first three principal components for large shells on any background. Background signal is subtracted from the data.	53
Figure 2.5.2 – 3: Stem plot of the first three principal components for large shells on any background. The data is mean-centered.....	54
Figure 2.5.2 – 4: Stem plot of the first three principal components for large shells on any background. The data is autoscaled.	55
Figure 2.5.2 – 5: Stem plot of the first three principal components for large shells on any background. The data is mean-centered with background subtraction.....	56
Figure 2.5.2 – 6: Stem plot of the first three principal components for large shells on any background. The data is autoscaled with background subtraction.	57
Figure 2.5.2 – 7: Graphical display of the first three principal components for empty shells on soil background. The data is mean-centered.....	59
Figure 2.5.2 – 8: Graphical display of the first three principal components for empty shells on table background. The data is mean-centered.....	60
Figure 2.5.2 – 9: Graphical display of the first three principal components for empty shells on table or soil background. The data is mean-centered. Color-coding is according to shell size..	61
Figure 2.5.2 – 14: Stem plot of the first three principal components for large shells on any background. The data is mean-centered.....	66
Figure 2.5.2 – 15: Stem plot of the first three principal components for large and medium shells on any background. The data is mean-centered.....	67
Figure 2.5.2 – 16: Stem plot of the first three principal components for large, medium, and small shells on any background. The data is mean-centered.....	68

Figure 2.5.2 – 17: Graphical display of the first three principal components for large and medium shells on any background. The data is mean-centered.....	69
Figure 2.5.2 – 18: Graphical display of second, third and fourth principal components for large and medium shells on any background. The data is mean-centered.....	69
Figure 2.7-1. All shell sizes, with correlation, 0% Don't Know (left) and 15% Don't Know (right), with an empty shell in the background run.....	77
Figure 2.7-2. Small and medium size shells, with correlation, 0% Don't Know (left) and 15% Don't Know (right), with an empty shell in the background run.....	78
Figure 2.7-3. Small (left) and medium (right) shells, with correlation, each 15% Don't Know, with an empty shell in the background run.	79
Figure 2.7-4. Small and medium shells together, with correlation, 0% Don't Know (left), 15% Don't Know (right), with NO empty shell in the background run.....	79
Figure 2.7-5. Small (left) and medium (right) shells analyzed separately, with correlation, 15% Don't Know, with NO empty shell in the background run.....	80
Figure 2.7-6. All December 2004 data (left) and December 2004 data with an empty shell in the background (right), with correlation, 15% Don't Know.....	81
Figure 2.7.5-1. Neural net results of all PELAN IV data with shell size as input (left plot) and with background as input (right plot). The black curve is the baseline curve as described in the text.....	83
Figure 2.7.5-2. Neural net results of all PELAN IV data with baseline curve (no inputs) and with shells size and background as inputs.	83
Figure 2.7.5-3. GLRT results of all PELAN IV data with only SEC input (left plot) and with size as input (right plot). The black curve is for the equivalent analysis using the neural net.....	84
Figure 2.7.5-4. GLRT results of all PELAN IV data with background as input (left plot) and with size and background as input (right plot). The black curve is for the equivalent analysis using the neural net.	84
Figure 2.7.5-5. Neural network results of all PELAN IV data for each shell size group after training on each separately.....	85
Figure 2.7.5-6. GLRT results of all PELAN IV data for small and medium sized shell size group after training on each separately. The black curves are the equivalent analysis using neural networks.....	85

TABLE OF TABLES

Improved Analysis Algorithms for UXO Filler Identification	i
Table 1.3-1 Elemental densities and ratios of three classes of substances.	11
Table 2.3-1 GLRT Test Results of Shell Data Using Empty Shell Background.....	24
Table 2.3-2 GLRT Test Results of Shell Data Using Empty Shell Background.....	25
Table 2.3-3 GLRT Test Results of Shell Data Using Empty Shell Background.....	26
Table 2.3-4 GLRT Test Results of Shell Data Using Empty Shell Background.....	26
Table 2.3-5 GLRT Test Results of Shell Data Using Empty Shell Background.....	27
Table 2.3-6 GLRT Test Results of Shell Data Using Empty Shell Background.....	27
Table 2.3-7 GLRT Test Results of Shell Data Using Empty Shell Background.....	28
Table 2.3-8 GLRT Test Results of Shell Data Using Empty Shell Background.....	28
Table 2.3-9 GLRT Test Results of Shell Data Using Empty Shell Background.....	29
Table 2.3-10 GLRT Test Results of Shell Data Using Empty Shell Background.....	29
Table 2.3-11 GLRT Test Results of Shell Data Using Empty Shell Background.....	30

Table 2.3-12	GLRT Test Results of Shell Data Using Empty Shell Background.....	30
Table 2.3-13	GLRT Test Results of Shell Data Using Empty Shell Background.....	31
Table 2.3-14	GLRT Test Results of Shell Data Using Empty Shell Background.....	31
Table 2.3-15	GLRT Test Results of Shell Data Using Empty Shell Background.....	32
Table 2.3-16	GLRT Test Results of Shell Data Using Empty Shell Background.....	32
Table 2.3-17	GLRT Test Results of Shell Data Using Empty Shell Background.....	33
Table 2.3-18	GLRT Test Results of Shell Data Using Empty Shell Background.....	33
Table 2.3-19	GLRT Test Results of Shell Data Using Empty Shell Background.....	34
Table 2.3-20	GLRT Test Results of Shell Data Using Empty Shell Background.....	34
Table 2.3-21	GLRT Test Results of Shell Data Using Empty Shell Background.....	35
Table 2.3-22	GLRT Test Results of Shell Data Using Empty Shell Background.....	35
Table 2.3-23	GLRT Test Results of Shell Data Using Empty Shell Background.....	36
Table 2.3-24	GLRT Test Results of Shell Data Using Empty Shell Background.....	36
Table 2.7-1.	Data distribution (546 total) for PELAN IV data (excluding December 2004 at Indian Head) taken with an empty shell in the background.....	74
Table 2.7-2.	Data distribution (494 total) for 2004 data taken without an empty shell in the background.....	74

ACRONYMS

ACF	Autocorrelation function
ACM	Autocorrelation matrix
ANFO	Ammonium Nitrate-Fuel Oil mixture
AR	Auto-Regressive
ARMA	Auto-Regressive moving average
ASTM	American Standards of Testing and Materials
AT	Anti-tank
BGO	Bismuth germinate $\text{Bi}_3\text{Ge}_4\text{O}_{12}$
C-4	Composition 4
COMP B	A mixture of TNT and RDX
cps	counts per second
CW	Chemical Warfare agent
DoD	Department of Defense
DOE	Department of Energy
DP	Detection Probability
ESTCP	Environmental Security Technology Certification Program
FA	False Alarm
GLRT	Generalized Likelihood Ratio Test
HE	High explosive
HEAT	High Explosive Anti-Tank
HMX	High Melting eXplosive (one of multiple definitions)
ID	Identification
IEDs	Improvised explosive devices
LS	Least Squares spectral analysis
MCNP	Monte Carlo N-Particle code
ML	Maximum likelihood
NAVEODTECHDIV	Navy Explosive Ordnance Disposal Technology Division, Indian Head
NEODTD	Equivalent to NAVEODTECHDIV
NFI	Non-invasive Filler Identifier (program sponsored by NEODTD)
NG	Neutron generator
NN	Neural Network
NRC	Nuclear Regulatory Commission
pdf	Probability density function
PC	Principal Components
PCA	Principal Component Analysis
PELAN	Pulsed ELemental Analysis with Neutrons
PFTNA	Pulsed Fast/Thermal Neutron Analysis
PoP	Plaster of Paris
PSD	Power spectral density
RDX	A military explosive (1,3,5-Trinitro-1,3,5-triazacyclohexane)
ROC	Receiver operator characteristic
SAIC	Science Applications International Corporation
SEC	Spider Elemental Counts
SERDP	Strategic Environmental Research and Development Program

SIMCA	Soft Independent Modeling of Class Analogies
SNR	Signal to noise ratio
SPIDER	SPectrum Interpolation and DEconvolution Routine
TNA	Thermal neutron analysis
TNT	Tri-nitro toluene
TSWG	Technical Support Working Group
μ Ci	micro Curies (37,000 decays/second)
UXO	Unexploded Ordnance
WKU	Western Kentucky University
WP	White Phosphorous
WSS	Wide-sense stationary
XML	Extensible Markup Language

1. INTRODUCTION

1.1 Project Background

Prior to the selection of a disposal method for unexploded ordnance (UXO), a determination must be made of the filler material. The materials can range from standard military explosives to chemical agents to inert simulants. Currently, trained UXO experts perform this determination using external markings and visual examination. Many times, the UXO has weathered or corroded and the markings and external visual cues are deteriorated or absent. If a conservative approach is used and all questionable UXO is treated as explosive or chemical filled, the cost of clearance operations is greatly increased. If a less conservative approach is used, accidents occur, such as those at the Naval Surface Warfare Center, Indian Head Division, and the San Clemente Test Range, that lead to injury or loss of life. There is the need for a means of rapidly and correctly determining the fill of UXO to permit the rapid disposition of inert rounds and proper handling of explosive or chemical-filled UXO.

The Naval Explosive Ordnance Technology Division (NAVEODTECHDIV) has been investigating the use of the Pulsed ELEMENTAL Analysis with Neutrons (PELAN) developed by the University of Western Kentucky (WKU) and SAIC as part of an Office of Naval Research Applied Research effort and an Environmental Security Technology Certification Program (ESTCP) effort. These efforts have demonstrated the utility of using PELAN to gather data from explosive-, chemical- and inert-filled UXO, but have highlighted the need for more advanced signal processing to increase the probability of detection and reduce the false alarm rate.

This project addressed the Statement of Need Number UXSON-04-02 for Innovative Technology for Identification of Filler Material in Recovered UXO. The PELAN system is being further developed and tested with support from ESTCP and NAVEODTECHDIV. Our tasks were to investigate, test, and demonstrate advanced data analysis and decision making algorithms for the PELAN system for non-intrusively identifying the fillers of UXO in-situ for cost-effective and safer environmental remediation. The goals of these improvements are to increase the filler detection efficiency and accuracy, reduce false alarm rates, and allow the system to be capable of learning the signatures of new targets.

1.2 Objective

The objectives in this project were to investigate, test, and demonstrate advanced data analysis and decision-making algorithms for the PELAN system. The goals of these algorithmic improvements are to increase the filler detection efficiency and accuracy, reduce false alarm rates, and improve the system's ability to learn the signatures of new targets.

We originally proposed a 12-month, three-phase effort for evaluating, testing, and demonstrating advanced analysis algorithms to improve the performance of PELAN. Below is the Statement of Work we suggested for the project.

Phase 1: Concept Study (three months)

- Collect available data (spectra and elemental intensities) measured with the PELAN system
- Apply matching pursuits and/or other algorithms to analyze spectral data for elemental features
- Investigate generalized likelihood ratio decision algorithms for the decision-making process
- Determine detection and false alarm rates (or receiver operating characteristics, ROC's) as part of the analysis
- Compare results to current decision tree approach

Phase 2: Optimization (six months)

- Evaluate additional spectral algorithms, including the Multiple Signal Classification (MUSIC) method
- Evaluate additional decision algorithms, including Support Vector Machines and Relevance Vector Machines
- Evaluate combinations of spectral and decision-making analyses for optimal performance
- Extend model to include potentially nonlinear factors
- Conduct targeted experiments to support optimization and model evaluation

Phase 3: Implementation (3 months)

- Convert MATLAB®-based (MathWorks, Inc.) algorithms to software code
- Incorporate algorithms into PELAN system and test
- Demonstrate PELAN with improved analysis algorithms

SAIC's focus in this project were the following tasks:

- Provided project management
- Provided subcontract to Duke
- Developed test plans for tests at NAVEODTECHDIV
- Evaluated and provided data sets to Duke and NAVEODTECHDIV for analysis
- Implemented Generalized Likelihood Ratio Test (GLRT) and entropy-based confidence level into PELAN
- Examined normalization methods to reduce effects from moisture changes, shell size, and target position
- Examined variables and analysis approaches affecting the principal component analysis (PCA) results
- Provided guidance to Duke University during their investigations

Duke University was a major contributor and provided the following tasks:

- Developed confidence metrics for fill material classification and identification
- Examined the effects of background subtraction on PCA analysis
- Investigated several spectral estimation techniques to improve the spectral features
- Analyzed the performance of GLRT trained on the results of several spectral analyses using ROC
- Compared neural net and GLRT decision-making results using same analysis approaches
- Provided guidance for the interpretation of the analysis and for implementation

These tasks and results are described in this final report. The results of the studies in this project are key to the continuing development of PELAN for improved discrimination of UXO fills.

1.3 Technical Approach

1.3.1 Technical Description

High explosives (TNT, RDX, C-4, etc.) are composed primarily of the chemical elements hydrogen, carbon, nitrogen, and oxygen. Many innocuous materials are also primarily composed of these same elements. However, these elements are found in each material with very different elemental ratios and concentrations. It is thus possible to identify and differentiate, for example, TNT from paraffin. Table 1.3-1 shows the atomic density of elements for various materials, along with the atomic ratios. For narcotics, the C/O ratio is at least a factor of two larger than the innocuous materials. Explosives have been shown to be differentiated by utilization of both the C/O ratio and the C/N ratio. The problem of identifying explosives and other threat materials is thus reduced to the problem of elemental identification.

Nuclear techniques present a number of advantages for non-destructive elemental characterization. These advantages include the ability to examine bulk quantities with speed, high elemental specificity, and no memory effects from the previously measured object. These qualities are important for an effective detection system for explosives and drugs.

Neutrons are highly penetrating particles, so their intensity is not diminished significantly by the thickness of commonly utilized containers. Furthermore, the outgoing gamma rays are also very penetrating, easily exiting the interrogated volume. Thus, the method is non-intrusive (the interrogation can take place from a distance of several centimeters) and non-destructive because of the very small amount of radiation absorbed by the interrogated object.

Density or Ratio	H	C	N	O	Cl	C/O	C/N	Cl/O
<i>Narcotics</i>	High	High	Low	Low	Medium	High, >3	High	Very High
<i>Explosives</i>	Low-Medium	Med	High	Very High	Medium to None	Low, <1	Low, <1	Low to Medium
<i>Plastics</i>	Medium-High	High	High to Low	Medium	Medium to None	Medium	Very High	-

Table 1.3-1 Elemental densities and ratios of three classes of substances.

1.3.2 The PELAN System

Developed by WKU with support from Technical Support Working Group (TSWG) and Navy Explosive Ordnance Disposal, PELAN utilizes a pulsing deuterium-tritium (d-T) neutron generator.

By using fast neutron reactions, capture reactions, and activation analysis, a large number of elements can be identified in a continuous mode without sampling. PELAN is a man-portable device designed for portability and rapid deployment. The PELAN III prototype is shown in Figure 1.3-1. Under an existing program sponsored by TSWG, SAIC has extensively upgraded PELAN III to improve reliability, ease of use, and data handling. The new, upgraded version, PELAN IV, has been fabricated and undergone testing. This system, shown to the right in Figure 1.3-1, consists of two equal weight portions. The upper section is the neutron generator and accompanying digital control system. The lower section contains the embedded computer, detector system, detector shielding, and operator interfaces such as Ethernet communication link to a laptop. The controller provides fully automatic operation of PELAN. With a single touch command, all necessary power supplies are energized, neutrons are produced, and data is collected for a predetermined time. Upon the completion of data acquisition, the data are automatically reduced, analyzed, and the results of the interrogation are displayed on the screen.

SAIC has an exclusive license with NuMaT, Inc. and WKU to build and sell PELAN systems. In May 2002, the Phase III prototype was demonstrated at NAVEODTECHDIV in Indian Head, Md. At Indian Head, the system was used to acquire over 230 measurements on a variety of shells on a number of different soil types.

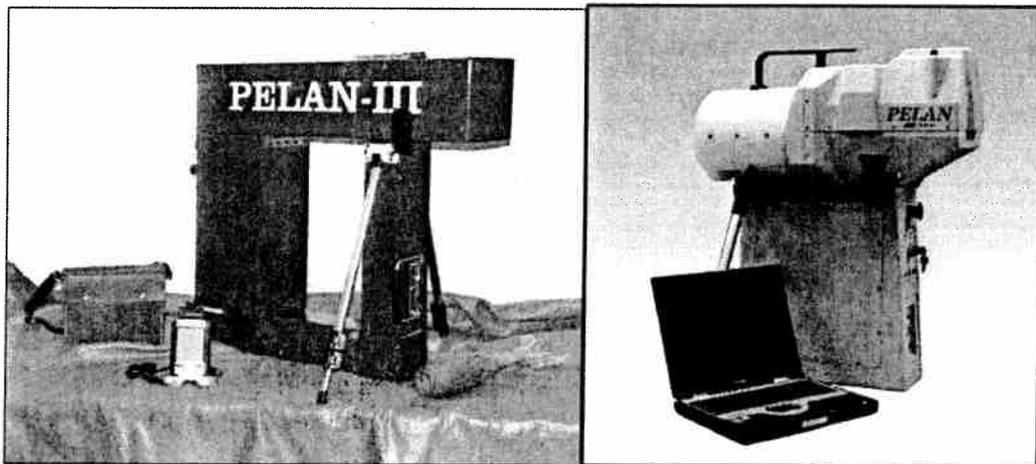


Figure 1.3-1 The PELAN III (left) and PELAN IV systems. The PELAN IV system was used for tests at Indian Head in December 2004.

1.3.3 Earlier Spectral Analysis and Decision-making Methods

In the current PELAN system, data analysis of the resulting gamma-ray spectra is performed with the computer code Spectrum Interpolation and Deconvolution Routine (SPIDER), a spectrum deconvolution code developed for Microsoft Corporation's Windows® 95/98/Windows NT® platforms. In the absence of any sample placed in front of the detector, the detector records gamma rays emanating from the materials surrounding the detector, as well as from the materials inside and around the neutron generator. This spectrum is called the background spectrum. When a sample is placed in front of the detector and a gamma-ray spectrum is acquired, the spectrum of a sample can

be represented as a linear combination of the background spectrum and the response spectra of the various elements utilized to fit the spectrum. SPIDER employs a Least-Squares (LS) algorithm to fit the equation.

The elemental intensities resulting from the LS fit and the computed elemental ratios are used in a decision tree to determine the composition of the filler material. In order to automatically identify an object through its elemental composition, a library of the substances of interest must reside in the PELAN's computer.

The successful characterization of materials using elemental analysis on PELAN data depends upon pattern recognition and experimental investigation using known substances. Explosives, chemical warfare agents, and contraband drugs are all substances with distinct elemental signatures. After interrogating the substances listed above, as well as a variety of harmless substances, characteristic differences become evident and can be exploited in order to determine an unknown substance. From the experimental results, elemental ratios as well as the presence or absence of specific chemical elements can be used in making a decision on whether an unknown substance belongs to a group of "dangerous" substances. The PELAN responses can be dependent on the material of the container and its physical properties as well as any other objects in the container. Attention should be given to the background and the container, particularly if the amount of substance to be analyzed is small.

1.3.4 Evaluation of New Analysis Techniques

In order to improve the detection sensitivity, reliability, reduce false negative and positive alarms, and to allow the system to more easily learn signatures of new targets and backgrounds, new analytical and decision-making analysis techniques were investigated. Several approaches were evaluated to develop a robust automated filler detection scheme for PELAN.

Currently, the count data for the elements of interest (namely, C, N, O and H) is estimated using a linear signal model and a measure of the signal and background spectrum. A Least Squares fit is utilized to obtain estimates of the counts of the various elements. Our objectives were to (1) investigate alternative signal models that may more accurately reflect the underlying physics associated with the sensing modality, (2) investigate alternative spectral estimation procedures (e.g., Auto-Regressive moving average (ARMA)) methods, matching pursuits, and MUSIC methods), and (3) investigate statistical algorithms to more effectively process the count data. These lines of investigation were separated into three stages. In the first, or proof-of-concept phase, we applied pre-existing matching pursuits algorithms and generalized likelihood ratio decision algorithms to existing data sets to demonstrate the feasibility of the proposed approaches and to assess performance improvements over the techniques currently in use. In the second phase, we pursued the remaining spectral estimation and statistical decision techniques to determine which combination provides the most robust and optimal performance. In the final phase, our intention was to transition the MATLAB-based algorithms developed at Duke to the PELAN system and to perform a field demonstration using these algorithms.

1.4 Summary of Results

Both the LS/GLRT and LS/PCA combinations showed significant improvement over the LS/decision tree approach, and the LS/GLRT method was implemented in the portable PELAN unit. Work is continuing on the PCA spectral analysis method, which shows even more improvement over the LS/GLRT approach. The PCA algorithm is much more effective at using the entire spectrum to extract characteristics of the target for improved identification.

During this project, several key results were established, and are summarized below:

- GLRT was established as an effective decision-making algorithm.
 - GLRT can be used in conjunction with LS, PCA, and other spectral analysis techniques (e.g., MUSIC).
- The tertiary declaration was added for GLRT decision making (“unknown”) for explosives/inert-filled shells.
- PCA can perform better than Least Squares on shell targets.
- Background measurements may not be necessary for effective PCA analysis.
 - The need for empty shell in background is eliminated.
 - Less user input is required for recording the environment.
 - Further testing is required for verification.
- The GLRT, tertiary declaration and entropy-based confidence level were implemented into PELAN Non-invasive Filler Identifier (NFI) systems.
- Data collection at Indian Head was conducted December 6-22, 2004.
 - Strategic Environmental Research and Development Program (SERDP) data was made available to SAIC and Duke University in February 2005.

2. PROJECT ACCOMPLISHMENTS

2.1 Background

Prior to the start of the SERDP project, SAIC and Duke University conducted some preliminary studies into various spectral analysis and decision-making methods. A good summary of these studies is found in a paper presented at the SPIE Orlando conference in April 2004¹. In this section, we provide a summary of the methods that were investigated and the results using data collected with PELAN III in 2002 and 2003.

The PELAN data is currently processed using a Least Squares analysis of the spectrum, which is assumed to be a linear combination of the measured background and a number of elemental response functions that correspond to elements of interest. For each measured spectrum and associated background, the output of the Least Squares analysis provides a number of elemental

¹ Giancarlo Borgonovi, Daniel Holslin, Leslie Collins, Stacy Tantum, "Data analysis for classification of UXO filler using pulsed neutron techniques," in Detection and Remediation Technologies for Mines and Mine-like Targets IX, edited by Russell S. Harmon, J. Thomas Broach, John H. Holloway, Jr., Proceedings of SPIE Vol. 5415 (SPIE, Bellingham, WA, 2004).

intensities. For small amounts of explosive material or, in general, when the sample spectrum is not very different from the background, the elemental intensities are not directly proportional to the fraction of elements in the target. They are, however, useful indicators, which are representative of the spectra and display correlation with some properties (such as explosive or non-explosive) of the substances that produced the spectra.

The correlation of the elemental intensities with the material properties useful for decision making has historically been demonstrated and exploited by using a decision tree approach. The decision tree is a set of rules or inequalities on the elemental intensities that is developed through inspection of the data plus trial and error. The development of a tree is a laborious process, since the data must be visualized in multi-dimensions.

Two well-established techniques of statistical data analysis have been investigated for the analysis of PELAN data. The first one is the GLRT, which offers a simple and automated way of correlating the indicators to the properties. The GLRT method can be an alternative to the decision tree approach and can be used on the results of the Least Squares analysis as well as on the results from other spectral analysis techniques. An advantage of the GLRT approach is that it provides a natural way of assessing performance, through the ROC curve.

The second technique is the PCA, which uses an eigenvector approach to derive sets of numbers (indicators) directly from the spectra, without the need for a model and elemental response functions. These indicators are also representative of the spectra and display correlation to the properties. Thus, the decomposition of the spectra into principal components is an alternative to the characterization of the spectra by elemental intensities.

2.1.1 PCA for Spectral Analysis

The method of principal component analysis is based on a particular expansion in terms of orthonormal functions. Any vector, including a spectrum, can be decomposed into a sum of vectors as follows:

$$|S\rangle = c_1|P_1\rangle + c_2|P_2\rangle + \dots + c_n|P_n\rangle,$$

where the $c_1 \dots c_n$ coefficients are numbers, and the vectors $|P_1\rangle \dots |P_n\rangle$ form an orthonormal basis. The above equation is true for any orthonormal set. The question is, is there a preferred way to choose the $|P_i\rangle$?

When we have many vectors $|S\rangle$, we can arrange them to form a matrix X . This matrix is not square. However, the matrix $X^T X$ that is proportional to the covariance of the matrix X (provided that the data has been mean-centered) is square. The eigenvectors of this square covariance matrix, by definition, form an orthonormal set. They can be ordered in descending order according to the magnitude of the corresponding eigenvalues. The advantage of this procedure is that often one does not need all of the eigenvectors to expand the spectra. Instead, a relatively small number of components may be sufficient for the spectral expansion, and most of the variance in the data is

captured by the first few principal components. Once the principal components have been determined, a particular spectrum is represented by a small number of indicators, also known as scores, which are obtained by projecting the vector onto the principal components, as follows:

$$c_i = \langle P_i | S \rangle$$

In summary, the PCA method works as follows:

- Form a matrix of all the data, for example spectra (taking into account the background).
- Calculate the covariance of this matrix.
- Compute the set of eigenvectors of the covariance matrix, and order them in descending order by eigenvalue.
- Select a smaller subset of eigenvectors from the top of the ordered list. These are the principal components to be used.
- Project each data vector (spectrum) on the above components.

The method just described extracts a group of indicators, or scores, for each data vector. These numbers take the place of the elemental intensities obtained from the Least Squares approach, and can be used to make declarations, either via a decision tree⁵ or via a GLRT method.

2.1.2 GLRT for Decision Making

The GLRT is an excellent choice for carrying out the secondary data analysis. It can be applied to any set of indicators, either elemental intensities, or scores on principal components. For application of the GLRT method, we need a training set consisting of samples. For each sample, we know the indicators and whether the sample is explosive or inert. For example, if the primary data analysis is obtained using the Least Squares method, a sample may consist of a vector W with four components (C, H, N, and O intensities, obtained from the primary data analysis). The samples could include all the measurements or a subset, for example, all the measurements made on a particular kind of environment, such as on a concrete surface. We compute the mean and the covariance matrix for the explosives (m_1, Cov_1) and for the inert items (m_0, Cov_0), respectively. For a generic vector W compute

$$\lambda = (W - \mu_1)^T (Cov_1)^{-1} (W - \mu_1) - (W - \mu_0)^T (Cov_0)^{-1} (W - \mu_0),$$

where T indicates the transpose. The quantity λ can be used to make a declaration by comparing its value to a threshold level. The procedure for the declaration is as follows:

If $\lambda < \text{Threshold}$ the declaration is explosive

If $\lambda > \text{Threshold}$ the declaration is inert

By comparing the result of the declarations to the known state (explosive or inert) of the sample, we can calculate the detection probability (DP) and the probability of false alarms (FA) for that

particular threshold. If we let the threshold vary over the entire range of 0 to 1, we obtain a ROC curve, which is a plot of detection probability versus probability of false alarms.

The ROC curve is a good global way of assessing the performance of both primary and secondary data analysis methods. The GLRT method can then be used for making decisions on new data. We select a threshold value that corresponds to an acceptable level on the ROC curve (for example 10% FA, 80% DP) and use the same procedure as above for making the decision. The result of the decision may be explosive or inert. The GLRT parameter has also been used to calculate a confidence value to associate with the declaration. If enough information is available for a range of different substances, the GLRT parameter can also be used for substance identification. The substance is identified as belonging to the class (either explosive or inert) corresponding to the highest confidence.

The general conclusion from this analysis is that the GLRT method is easier to implement and, on average, produces better results (in terms of detection and false alarm probabilities) than the decision tree approach.

The advantages of the GLRT method are as follows:

- It can be applied to any number of dimensions.
- The parameters determining the decision can be obtained with a completely automated analysis.
- The method generates well-behaved boundaries and does not result in "over training."
- Once a threshold has been selected, a confidence level can be associated with the declaration.

The GLRT is fundamentally a binary decision, in this case, explosive or inert. In some cases, where the distributions of explosives and inerts overlap, the result is not clear because the probabilities of each choice are nearly equal. In these situations, it may be desirable to add a third decision ("unknown" or "don't know"). For the NFI 6.2 Program, the requirement allowed for up to 25% "don't know" results. The GLRT was modified to include this choice, making it a tertiary decision. For this case, the ROC curve is relabeled with probability of detection on the vertical axis (as in the standard ROC curve) and with probability of correct rejection, P_{CR} , (that is, detecting an inert) on the horizontal axis. The probability of false alarm, P_f , is then given by $P_f = 1 - P_{CR}$. Throughout this report, we use the tertiary GLRT for decision making.

2.1.3 Preliminary Results

Data were collected with the PELAN IV system at the NAVEODTECHDIV, Stump Neck, Indian Head, Md., during December 2003. The data were from a number of shells filled with both explosive and inert material. The analysis of the data was carried out using the GLRT applied to the following indicators:

- Elemental intensities (C, H, N, O) as determined with the Least Squares method.
- Principal components determined directly from the sample and background spectra.

The PELAN system collects gamma spectra, a sample spectrum and an associated background spectrum, which can be analyzed with one of two methods. The first method is based on a Least Squares fit and uses elemental response functions. The second method is based on the PCA approach, uses principal components that are derived from measurements conducted on an ensemble of test items spanning the items of interest and is currently implemented through specialized MATLAB programs. Both methods result in indicators that, for the purpose of the subsequent analysis, replace the run and background spectra. The set of indicators for a certain ensemble of items is then processed through the GLRT algorithm, which provides estimates of parameters to be used for future decisions on new items. Since the first application of the GLRT technique is a training procedure, one must know the characteristic (explosive or inert) of each item used in the training set. The GLRT also provides a measure of performance (the ROC curve) of the system on the training set.

We note that there is not a unique way of applying the PCA method. For example, different regions of the spectra can be utilized, or different ways of combining fast and thermal spectra, as well as different ways of combining run and background signals. Accordingly, the PCA method was exercised, both at SAIC and at Duke University, in different forms, showing performance (as measured by the ROC curves) generally superior to the LS method.

Usually, the ROC curves produced from training show good performance, which means acceptably high detection probability and acceptably low probability of false alarms. The real test comes when a new set of data is analyzed using parameters derived from a previous test, which makes the new test a blind test. We have made a comparison using two data sets taken on explosive and inert shells, one set measured in April 2003 and one set measured in December 2003. The April set is more comprehensive since it consisted of a much larger number of data points. For a comparison of LS/GLRT and PCA/GLRT results, see Figures 3 and 4 in the SPIE publication.

We have found that the GLRT method has several advantages over the decision tree approach. The main advantage is that the method does not require the cumbersome manual analysis involved in developing sets of inequalities and translating them into conditional "if-then" statements. Instead, one uses the indicators obtained from measuring a known set to obtain mean and covariance for explosives and inert items. The mean and covariance are then used to generate a ROC curve, which is a powerful way to assess the performance of the system.

The GLRT method has been tested on several sets of elemental intensities derived from data collected during 2002 and 2003. In all cases, the GLRT method has been found to be superior to the decision tree approach and, therefore, has been implemented on the PELAN IV system.

The current LS approach, which uses response functions, generates elemental intensities for selected elements. It appears that these numbers cannot be used in an absolute, that is to say stoichiometric sense, but only with reference to a calibration on known sets of samples. Therefore, the data analysis approach must be based on some kind of training (such as the decision tree) and the problem becomes one of pattern recognition. Under these circumstances, it is a legitimate question to ask whether PCA can perform the same function. With PCA, the spectra themselves become the items to be recognized, and the indicators are the coefficients of a small number of (orthogonal) principal components, which are sufficient to characterize the spectra.

In addition to providing better performance, as illustrated by the ROC curves, we have found that the PCA method also results in better stability when the parameters obtained from one test (training) are applied to a second, totally independent test. Thus, the heuristic pattern recognition approach inherent in the PCA method appears to be a valid alternative to the more classic LS approach for the current PELAN system design and targets of interest.

2.2 Data Collection

The purpose of the data collection was to support the algorithm development in the SERDP project. A test plan, shown in Appendix A, was developed and provided to NAVEODTECHDIV and SERDP. This plan for SERDP data collection was incorporated into one written by NAVEODTECHDIV, which can provide SERDP a copy if needed. The data was collected during the period December 6 through 22, 2004.

These were the issues we wanted to address in this testing:

1. Study the effects of variations in target-to-detector distance and filler size and to evaluate methods to correct these effects. Reducing this affect would also allow the training on smaller, more readily available shells, to be used for identification of larger shell fills.
2. Study and correct for changes on H signal (or any other thermal capture gamma ray) due to variations in the moisture content of the soil (or changes in neutron thermalization caused by the presence of a nearby wall).
3. Evaluate the effect on PCA results due to variation in background environment (especially dry versus wet soil).
4. Investigate methods to eliminate the need for using an empty shell in the background measurement.
5. Improve the tertiary explosives identifier by separating the particular types of inert and explosive fills. Develop GLRT parameters separately for the separate fill clusters.
6. Acquire additional data on new targets for addition to the library.

Under the guidance of Denice (Forsht) Lee, NAVEODTECHDIV personnel operated the PELAN IV units and collected the data. A typical setup is shown in Figure 2.2-1. Because testing for the NFI program was also being conducted during this time, SAIC personnel were not allowed to be present at the test site. Rachel Kinney from NAVEODTECHDIV provided the data to SAIC on January 31, 2005. The data were analyzed using the SPIDER algorithm with an empty shell in the background. This data was also provided to Duke University and NAVEODTECHDIV for supporting their algorithm studies. A total of 70 runs were made.

The data variables included the following:

- Explosive fills: TNT, RDX, HMX
- Inert fills: Plaster of Paris, wax, sand
- Shell size: 60mm-155mm
- Environment: Soil, wet soil sand, wet sand, metal table
- Moisture variation of sand/soil: 3%, 17% and 30%



Figure 2.2-1. PELAN IV shown inspecting ordnance in soil test box at Indian Head in December 2004.

2.3 Normalization Studies

The current NFI Target Identification Algorithm applies either GLRT or Tertiary test to the SPIDER results, directly. Due to the variations in both shell sizes and environment, the signal intensity ranges are quite large, even for some of the same targets. Applying ratios among different SPIDER element results could reduce that variation theoretically. However, direct application of the ratios could cause some problem, as the denominators are close to zero or even negative. This report summarizes the study of utilizing data transformation to the SPIDER data before applying GLRT tests. The proposed data transformation could eliminate the situations of denominator close to zero or being negative and still preserve the proper “signal intensity” information.

2.3.1 Data Transformation

To explore the performance of using ratios among the Least-Squares estimated element amounts, i.e., SPIDER results, a data transformation is applied to eliminate the cases of dividing by negative or close to zero numbers. For a given set of data, X , with some data being close to zero or negative, the following linear transformation is applied to calculate the transformed data, X' :

$$x_i' = \frac{(x_i - X_{\min})}{(X_{\max} - X_{\min})} \times (1 - X_0) + X_0,$$

where x_i and x_i' represent the i -th data point before and after transformation, respectively; X_{\min} and X_{\max} represent the minimum and maximum of the data, respectively; and X_0 is a threshold (a small positive constant), for example 0.1.

The above transformation is applied to the SEC (SPIDER Element Counts) of the interested elements, such as H, C, N, and O. Note that the X_{\min} and X_{\max} are different for each element, but X_0 is the same.

The GLRT parameters, including the means and inverse of the covariance matrix, are derived from the transformed data.

Before making the GLRT test, the SPIDER results are transformed first using the same X_{\min} , X_{\max} , and X_0 for each element used before, respectively. Both “transformed data” and “untransformed data” are tested with GLRT.

A special way of transforming data is by applying offset only. For a given set of data, X , with some data being close to zero or negative, the following linear transformation is applied to calculate the transformed data, X' :

$$x_i' = x_i + |X_{\min}| + X_0,$$

where x_i and x_i' represent the i -th data point before and after transformation, respectively; $|X_{\min}|$ represents the absolute value of the minimum of the data; and X_0 is a threshold (a small positive constant), for example 0.1.

The “Offset Data” and the “Transformed Data” are treated similarly.

2.3.2 GLRT Target Grouping and Setup

Tests are performed using the SEC, either “transformed” or “untransformed,” directly (C, H, N, and O), or using the ratios among elements (H/C, N/C, and C/O).

Two groupings on targets are tested, respectively.

a. Dual targets:

The targets are categorized into two groups, “Explosives” and “Inerts.”

b. Multi-targets:

The targets are categorized into different substances. There are six groups for explosives, i.e., TNT, RDX, HMX, CompB, HEAT, and unknown explosives. There are five groups for inerts, i.e., cement, sand, plaster of Paris, wax, and empty.

The shell sizes are grouped into three groups.

a. Small shells: for shell size smaller than 90mm

b. Medium shells: for shell size between and including 90mm and 105mm

c. Large shells: for shell size larger than 105mm

The environments are grouped into three groups.

a. Concrete: including concrete, gravel, sand, and asphalt

b. Metal table: including both metal and metal table

- c. **Soil:** including soil, plastic table, wood table, and all the other environments

With different combinations of GLRT parameters and grouping of parameters, as described above, the following 24 GLRT setups are tested.

1. **Data:** *Untransformed*
GLRT parameters: *C, H, N, and O*
Target grouping: *Dual targets*
Training grouping: *Target type only*
2. **Data:** *Untransformed*
GLRT parameters: *C, H, N, and O*
Target grouping: *Dual targets*
Training grouping: *Target type and Shell size*
3. **Data:** *Untransformed*
GLRT parameters: *C, H, N, and O*
Target grouping: *Dual targets*
Training grouping: *Target type, Shell size, and Environment*
4. **Data:** *Untransformed*
GLRT parameters: *C, H, N, and O*
Target grouping: *Multi-targets*
Training grouping: *Target type only*
5. **Data:** *Untransformed*
GLRT parameters: *C, H, N, and O*
Target grouping: *Multi-targets*
Training grouping: *Target type and Shell size*
6. **Data:** *Untransformed*
GLRT parameters: *C, H, N, and O*
Target grouping: *Multi-targets*
Training grouping: *Target type, Shell size, and Environment*
7. **Data:** *Untransformed*
GLRT parameters: *C, H, N, and O*
Target grouping: *Dual targets*
Training grouping: *Target type only {Using only Small and Medium Shell data}*
8. **Data:** *Untransformed*
GLRT parameters: *C, H, N, and O*
Target grouping: *Dual targets*
Training grouping: *Target type only {Using only Medium Shell data}*
9. **Data:** *Transformed*
GLRT parameters: *H/C, N/C, and O/C*
Target grouping: *Dual targets*
Training grouping: *Target type only*
10. **Data:** *Transformed*
GLRT parameters: *H/C, N/C, and O/C*

- Target grouping:** Dual targets
Training grouping: Target type and Shell size
11. **Data:** Transformed
GLRT parameters: H/C, N/C, and O/C
Target grouping: Dual targets
Training grouping: Target type, Shell size, and Environment
12. **Data:** Transformed
GLRT parameters: H/C, N/C, and O/C
Target grouping: Multi-targets
Training grouping: Target type only
13. **Data:** Transformed
GLRT parameters: H/C, N/C, and O/C
Target grouping: Multi-targets
Training grouping: Target type and Shell size
14. **Data:** Transformed
GLRT parameters: H/C, N/C, and O/C
Target grouping: Multi-targets
Training grouping: Target type, Shell size, and Environment
15. **Data:** Transformed
GLRT parameters: H/C, N/C, and O/C
Target grouping: Dual targets
Training grouping: Target type only {Using only Small and Medium Shell data}
16. **Data:** Transformed
GLRT parameters: H/C, N/C, and O/C
Target grouping: Dual targets
Training grouping: Target type only {Using only Medium Shell data}
17. **Data:** Offset adjusted
GLRT parameters: H/C, N/C, and O/C
Target grouping: Dual targets
Training grouping: Target type only
18. **Data:** Offset adjusted
GLRT parameters: H/C, N/C, and O/C
Target grouping: Dual targets
Training grouping: Target type and Shell size
19. **Data:** Offset adjusted
GLRT parameters: H/C, N/C, and O/C
Target grouping: Dual targets
Training grouping: Target type, Shell size, and Environment
20. **Data:** Offset adjusted
GLRT parameters: H/C, N/C, and O/C
Target grouping: Multi-targets
Training grouping: Target type only

21. *Data: Offset adjusted*
GLRT parameters: H/C, N/C, and O/C
Target grouping: Multi-targets
Training grouping: Target type and Shell size
22. *Data: Offset adjusted*
GLRT parameters: H/C, N/C, and O/C
Target grouping: Multi-targets
Training grouping: Target type, Shell size, and Environment
23. *Data: Offset adjusted*
GLRT parameters: H/C, N/C, and O/C
Target grouping: Dual targets
Training grouping: Target type only {Using only Small and Medium Shell data}
24. *Data: Offset adjusted*
GLRT parameters: H/C, N/C, and O/C
Target grouping: Dual targets
Training grouping: Target type only {Using only Medium Shell data}

The result of each GLRT setup is tabulated in the corresponding table numbers, for example Table 2.3-7 using the GLRT setup No.7, etc. The data used here were shared with Duke University and tabulated in the file *PELAN_Runs_Summary_to_Duke.xls*.

**Table 2.3-1 GLRT Test Results of Shell Data Using Empty Shell Background
 (GLRT Parameters: C, H, N, O of Untransformed Data)
 (Trained on Dual Targets Only)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	18	24	100	97	3	42.9%	97.0%
	Ground	47	20	27	54	53	1	42.6%	98.1%
	Total	89	38	51	154	150	4	42.7%	97.4%
Medium	Table	56	49	7	37	36	1	87.5%	97.3%
	Ground	59	53	6	37	36	1	89.8%	97.3%
	Total	115	102	13	74	72	2	88.7%	97.3%
Large	Table	44	38	6	22	21	1	86.4%	95.5%
	Ground	65	65	0	21	21	0	100.0%	100.0%
	Total	109	103	6	43	42	1	94.5%	97.7%

**Table 2.3-2 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: C, H, N, O of Untransformed Data)
(Trained on Dual Targets and Shell Size)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	35	7	100	91	9	83.3%	91.0%
	Ground	47	39	8	54	38	16	83.0%	70.4%
	Total	89	74	15	154	129	25	83.1%	83.8%
Medium	Table	56	56	0	37	37	0	100.0%	100.0%
	Ground	59	58	1	37	35	2	98.3%	94.6%
	Total	115	114	1	74	72	2	99.1%	97.3%
Large	Table	44	38	6	22	21	1	86.4%	95.5%
	Ground	65	58	7	21	21	0	89.2%	100.0%
	Total	109	96	13	43	42	1	88.1%	97.7%

**Table 2.3-3 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: C, H, N, O of Untransformed Data)
(Trained on Dual Targets, Shell Size, and Environment)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	37	5	100	90	10	88.1%	90.0%
	Ground	47	43	4	54	46	8	91.5%	85.2%
	Total	89	80	9	154	136	18	89.9%	88.3%
Medium	Table	56	56	0	37	37	0	100.0%	100.0%
	Ground	59	59	0	37	36	1	100.0%	97.3%
	Total	115	115	0	74	73	1	100.0%	98.6%
Large	Table	44	38	6	22	22	0	86.4%	100.0%
	Ground	65	65	0	21	21	0	100.0%	100.0%
	Total	109	103	6	43	43	0	94.5%	100.0%

**Table 2.3-4 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: C, H, N, O of Untransformed Data)
(Trained on Multi-targets Only)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	24	18	100	96	4	57.1%	96.0%
	Ground	47	25	22	54	53	1	53.2%	98.1%
	Total	89	49	40	154	149	5	55.1%	96.8%
Medium	Table	56	52	4	37	36	1	92.9%	97.3%
	Ground	59	56	3	37	36	1	94.9%	97.3%
	Total	115	108	7	74	72	2	93.9%	97.3%
Large	Table	44	41	3	22	21	1	93.2%	95.5%
	Ground	65	65	0	21	21	0	100.0%	100.0%
	Total	109	106	3	43	42	1	97.2%	97.7%

**Table 2.3-5 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: C, H, N, O of Untransformed Data)
(Trained on Multi-targets and Shell Size)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	38	4	100	80	20	90.5%	80.0%
	Ground	47	43	4	54	37	17	91.5%	68.5%
	Total	89	81	8	154	117	37	91.0%	76.0%
Medium	Table	56	55	1	37	37	0	98.2%	100.0%
	Ground	59	57	2	37	37	0	96.6%	100.0%
	Total	115	112	3	74	74	0	97.4%	100.0%
Large	Table	44	39	5	22	22	0	88.6%	100.0%
	Ground	65	61	4	21	21	0	93.8%	100.0%
	Total	109	100	9	43	43	0	91.7%	100.0%

**Table 2.3-6 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: C, H, N, O of Untransformed Data)
(Trained on Multi-targets, Shell Size, and Environment)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	37	5	100	85	15	88.1%	85.0%
	Ground	47	43	4	54	46	8	91.5%	85.2%
	Total	89	80	9	154	131	23	89.9%	85.1%
Medium	Table	56	55	1	37	37	0	98.2%	100.0%
	Ground	59	57	2	37	37	0	96.6%	100.0%
	Total	115	112	3	74	74	0	97.4%	100.0%
Large	Table	44	39	5	22	20	2	88.6%	90.9%
	Ground	65	65	0	21	21	0	100.0%	100.0%
	Total	109	104	5	43	41	2	95.4%	95.3%

**Table 2.3-7 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: C, H, N, O of Untransformed Data)
(Trained on Dual Targets {Small & Medium Shells})**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	27	15	100	97	3	64.3%	97.0%
	Ground	47	34	13	54	46	8	72.3%	85.2%
	Total	89	61	28	154	143	11	68.5%	92.9%
Medium	Table	56	53	3	37	35	2	94.6%	94.6%
	Ground	59	58	1	37	36	1	98.3%	97.3%
	Total	115	111	4	74	71	3	96.5%	95.9%
Large	Table	44	41	3	22	21	1	93.2%	95.5%
	Ground	65	65	0	21	20	1	100.0%	95.2%
	Total	109	106	3	43	41	2	97.2%	95.3%

**Table 2.3-8 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: C, H, N, O of Untransformed Data)
(Trained on Dual Targets {Medium Shells})**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	30	12	100	85	15	71.4%	85.0%
	Ground	47	39	8	54	31	23	83.0%	57.4%
	Total	89	69	20	154	116	38	77.5%	75.3%
Medium	Table	56	56	0	37	37	0	100.0%	100.0%
	Ground	59	58	1	37	35	2	98.3%	94.6%
	Total	115	114	1	74	72	2	99.1%	97.3%
Large	Table	44	41	3	22	21	1	93.2%	95.5%
	Ground	65	65	0	21	20	1	100.0%	95.2%
	Total	109	106	3	43	41	2	97.2%	95.3%

**Table 2.3-9 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: H/C, N/C, C/O of Transformed Data, X0=0.1)
(Trained on Dual Targets Only)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	28	14	100	96	4	66.7%	96.0%
	Ground	47	29	18	54	45	9	61.7%	83.3%
	Total	89	57	32	154	141	13	64.0%	91.6%
Medium	Table	56	51	5	37	36	1	91.1%	97.3%
	Ground	59	49	10	37	36	1	83.1%	97.3%
	Total	115	100	15	74	72	2	87.0%	97.3%
Large	Table	44	41	3	22	21	1	93.2%	95.5%
	Ground	65	64	1	21	20	1	98.5%	95.2%
	Total	109	105	4	43	41	2	96.3%	95.3%

**Table 2.3-10 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: H/C, N/C, O/C of Transformed Data, X0=0.1)
(Trained on Dual Targets and Shell Size)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	34	8	100	78	22	81.0%	78.0%
	Ground	47	38	9	54	41	13	80.9%	75.9%
	Total	89	72	17	154	119	35	80.9%	77.3%
Medium	Table	56	54	2	37	37	0	96.4%	100.0%
	Ground	59	50	9	37	36	1	84.7%	97.3%
	Total	115	104	11	74	73	1	90.4%	98.6%
Large	Table	44	37	7	22	20	2	84.1%	90.9%
	Ground	65	61	4	21	21	0	93.8%	100.0%
	Total	109	98	11	43	41	2	89.9%	95.3%

**Table 2.3-11 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: H/C, N/C, O/C of Transformed Data, X0=0.1)
(Trained on Dual Targets, Shell Size, and Environment)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	33	9	100	82	18	78.6%	82.0%
	Ground	47	42	5	54	43	11	89.4%	79.6%
	Total	89	75	14	154	125	29	84.3%	81.2%
Medium	Table	56	55	1	37	37	0	98.2%	100.0%
	Ground	59	51	8	37	35	2	86.4%	94.6%
	Total	115	106	9	74	72	2	92.2%	97.3%
Large	Table	44	38	6	22	22	0	86.4%	100.0%
	Ground	65	64	1	21	21	0	98.5%	100.0%
	Total	109	102	7	43	43	0	93.6%	100.0%

**Table 2.3-12 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: H/C, N/C, O/C of Transformed Data, X0=0.1)
(Trained on Multi-targets Only)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	28	14	100	91	9	66.7%	91.0%
	Ground	47	28	19	54	45	9	59.6%	83.3%
	Total	89	56	33	154	136	18	62.9%	88.3%
Medium	Table	56	53	3	37	34	3	94.6%	91.9%
	Ground	59	53	6	37	35	2	89.8%	94.6%
	Total	115	106	9	74	69	5	92.2%	93.2%
Large	Table	44	40	4	22	20	2	90.9%	90.9%
	Ground	65	65	0	21	18	3	100.0%	85.7%
	Total	109	105	4	43	38	5	96.3%	88.4%

**Table 2.3-13 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: H/C, N/C, O/C of Transformed Data, X0=0.1)
(Trained on Multi-targets and Shell Size)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	35	7	100	75	25	83.3%	75.0%
	Ground	47	42	5	54	39	15	89.4%	72.2%
	Total	89	77	12	154	114	40	86.5%	74.0%
Medium	Table	56	54	2	37	37	0	96.4%	100.0%
	Ground	59	53	6	37	36	1	89.8%	97.3%
	Total	115	107	8	74	73	1	93.0%	98.6%
Large	Table	44	40	4	22	20	2	90.9%	90.9%
	Ground	65	63	2	21	20	1	96.9%	95.2%
	Total	109	103	6	43	40	3	94.5%	93.0%

**Table 2.3-14 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: H/C, N/C, O/C of Transformed Data, X0=0.1)
(Trained on Multi-targets, Shell Size, and Environment)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	33	9	100	81	19	78.6%	81.0%
	Ground	47	39	8	54	40	14	83.0%	74.1%
	Total	89	72	17	154	121	33	80.9%	78.6%
Medium	Table	56	53	3	37	37	0	94.6%	100.0%
	Ground	59	54	5	37	37	0	91.5%	100.0%
	Total	115	107	8	74	74	0	93.0%	100.0%
Large	Table	44	38	6	22	21	1	86.4%	95.5%
	Ground	65	65	0	21	21	0	100.0%	100.0%
	Total	109	103	6	43	42	1	94.5%	97.7%

**Table 2.3-15 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: H/C, N/C, C/O of Transformed Data, X0=0.1)
(Trained on Dual Targets {Small & Medium Shells})**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	29	13	100	93	7	69.0%	93.0%
	Ground	47	35	12	54	43	11	74.5%	79.6%
	Total	89	64	25	154	136	18	71.9%	88.3%
Medium	Table	56	54	2	37	37	0	96.4%	100.0%
	Ground	59	49	10	37	36	1	83.1%	97.3%
	Total	115	103	12	74	73	1	89.6%	98.6%
Large	Table	44	42	2	22	21	1	95.5%	95.5%
	Ground	65	63	2	21	12	9	96.9%	57.1%
	Total	109	105	4	43	33	10	96.3%	76.7%

**Table 2.3-16 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: H/C, N/C, C/O of Transformed Data, X0=0.1)
(Trained on Dual Targets {Medium Shells})**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	29	13	100	87	13	69.0%	87.0%
	Ground	47	36	11	54	39	15	76.6%	72.2%
	Total	89	65	24	154	126	28	73.0%	81.8%
Medium	Table	56	54	2	37	37	0	96.4%	100.0%
	Ground	59	50	9	37	36	1	84.7%	97.3%
	Total	115	104	11	74	73	1	90.4%	98.6%
Large	Table	44	42	2	22	20	2	95.5%	90.9%
	Ground	65	63	2	21	13	8	96.9%	61.9%
	Total	109	105	4	43	33	10	96.3%	76.7%

**Table 2.3-17 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: H/C, N/C, C/O of Offset Data, X0=0.1)
(Trained on Dual Targets Only)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	29	13	100	85	15	69.0%	85.0%
	Ground	47	28	19	54	42	12	59.6%	77.8%
	Total	89	57	32	154	127	27	64.0%	82.5%
Medium	Table	56	53	3	37	37	0	94.6%	100.0%
	Ground	59	46	13	37	36	1	78.0%	97.3%
	Total	115	99	16	74	73	1	86.1%	98.6%
Large	Table	44	42	2	22	21	1	95.5%	95.5%
	Ground	65	60	5	21	15	6	92.3%	71.4%
	Total	109	102	7	43	36	7	93.6%	83.7%

**Table 2.3-18 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: H/C, N/C, O/C of Offset Data, X0=0.1)
(Trained on Dual Targets and Shell Size)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	35	7	100	76	24	83.3%	76.0%
	Ground	47	39	8	54	39	15	83.0%	72.2%
	Total	89	74	15	154	115	39	83.1%	74.7%
Medium	Table	56	55	1	37	37	0	98.2%	100.0%
	Ground	59	48	11	37	34	3	81.4%	91.9%
	Total	115	103	12	74	71	3	89.6%	95.9%
Large	Table	44	37	7	22	20	2	84.1%	90.9%
	Ground	65	62	3	21	21	0	95.4%	100.0%
	Total	109	99	10	43	41	2	90.8%	95.3%

**Table 2.3-19 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: H/C, N/C, O/C of Offset Data, X0=0.1)
(Trained on Dual Targets, Shell Size, and Environment)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	34	8	100	83	17	81.0%	83.0%
	Ground	47	43	4	54	41	13	91.5%	75.9%
	Total	89	77	12	154	124	30	86.5%	80.5%
Medium	Table	56	54	2	37	37	0	96.4%	100.0%
	Ground	59	48	11	37	32	5	81.4%	86.5%
	Total	115	102	13	74	69	5	88.7%	93.2%
Large	Table	44	39	5	22	21	1	88.6%	95.5%
	Ground	65	65	0	21	20	1	100.0%	95.2%
	Total	109	104	5	43	41	2	95.4%	95.3%

**Table 2.3-20 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: H/C, N/C, O/C of Offset Data, X0=0.1)
(Trained on Multi-targets Only)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	33	9	100	82	18	78.6%	82.0%
	Ground	47	37	10	54	38	16	78.7%	70.4%
	Total	89	70	19	154	120	34	78.7%	77.9%
Medium	Table	56	55	1	37	36	1	98.2%	97.3%
	Ground	59	44	15	37	31	6	74.6%	83.8%
	Total	115	99	16	74	67	7	86.1%	90.5%
Large	Table	44	43	1	22	20	2	97.7%	90.9%
	Ground	65	62	3	21	13	8	95.4%	61.9%
	Total	109	105	4	43	33	10	96.3%	76.7%

**Table 2.3-21 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: H/C, N/C, O/C of Offset Data, X0=0.1)
(Trained on Multi-targets and Shell Size)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	35	7	100	72	28	83.3%	72.0%
	Ground	47	41	6	54	35	19	87.2%	64.8%
	Total	89	76	13	154	107	47	85.4%	69.5%
Medium	Table	56	53	3	37	36	1	94.6%	97.3%
	Ground	59	44	15	37	34	3	74.6%	91.9%
	Total	115	97	18	74	70	4	84.3%	94.6%
Large	Table	44	41	3	22	20	2	93.2%	90.9%
	Ground	65	64	1	21	15	6	98.5%	71.4%
	Total	109	105	4	43	35	8	96.3%	81.4%

**Table 2.3-22 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: H/C, N/C, O/C of Offset Data, X0=0.1)
(Trained on Multi-targets, Shell Size, and Environment)**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	33	9	100	79	21	78.6%	79.0%
	Ground	47	41	6	54	42	12	87.2%	77.8%
	Total	89	74	15	154	121	33	83.1%	78.6%
Medium	Table	56	52	4	37	37	0	92.9%	100.0%
	Ground	59	49	10	37	34	3	83.1%	91.9%
	Total	115	101	14	74	71	3	87.8%	95.9%
Large	Table	44	39	5	22	21	1	88.6%	95.5%
	Ground	65	64	1	21	17	4	98.5%	81.0%
	Total	109	103	6	43	38	5	94.5%	88.4%

**Table 2.3-23 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: H/C, N/C, C/O of Offset Data, X0=0.1)
(Trained on Dual Targets {Small & Medium Shells})**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	35	7	100	80	20	83.3%	80.0%
	Ground	47	33	14	54	40	14	70.2%	74.1%
	Total	89	68	21	154	120	34	76.4%	77.9%
Medium	Table	56	54	2	37	36	1	96.4%	97.3%
	Ground	59	45	14	37	33	4	76.3%	89.2%
	Total	115	99	16	74	69	5	86.1%	93.2%
Large	Table	44	43	1	22	20	2	97.7%	90.9%
	Ground	65	60	5	21	9	12	92.3%	42.9%
	Total	109	103	6	43	29	14	94.5%	67.4%

**Table 2.3-24 GLRT Test Results of Shell Data Using Empty Shell Background
(GLRT Parameters: H/C, N/C, C/O of Offset Data, X0=0.1)
(Trained on Dual Targets {Medium Shells})**

Shell Size	Environment	Explosives			Inerts			Correct ID Rate	
		Number of Data A	Correctly Identified B	Miss Identified C	Number of Data E	Correctly Identified F	Miss Identified G	Explosives (B/A)	Inerts (F/E)
Small	Table	42	31	11	100	77	23	73.8%	77.0%
	Ground	47	32	15	54	39	15	68.1%	72.2%
	Total	89	63	26	154	116	38	70.8%	75.3%
Medium	Table	56	55	1	37	37	0	98.2%	100.0%
	Ground	59	48	11	37	34	3	81.4%	91.9%
	Total	115	103	12	74	71	3	89.6%	95.9%
Large	Table	44	42	2	22	20	2	95.5%	90.9%
	Ground	65	63	2	21	10	11	96.9%	47.6%
	Total	109	105	4	43	30	13	96.3%	69.8%

2.3.3 Results

1. Tables 2.3-1 to 2.3-24 tabulate the performance results of the corresponding GLRT setups listed in Section 2.3.2, respectively.
2. Comparing between Tables 2.3-1 and 2.3-4, using multi-targets can help with the correct identification (ID) rate when GLRT is trained on “targets,” or on “targets and shell size,” especially on small shell runs, but it does not help when the GLRT is trained on “targets, shell sizes, and environments.” Similar results can be observed for “transformed data with GLRT trained on SEC ratios,” as shown in Tables 2.3-7 to 2.3-12.
3. For “untransformed data” and using “dual-targets,” the best performance is the GLRT trained on “targets, shell sizes, and environments,” as shown in Table 2.3-3. Similar performance could be expected with GLRT trained on “targets and shell sizes” for “untransformed data” and using “multi-targets,” as shown in Table 2.3-5.
4. For “transformed data” and using “dual-targets,” the best performance is also for GLRT trained on “targets, shell sizes, and environments,” as shown in Table 2.3-9. But its performance is not as good as the “untransformed data,” compared to Table 2.3-3.
5. In general, the “data transformation” does not help with the performance, even though it makes the SEC ratios look better. This is probably because the data for different shell sizes and environments are brought closer through transformation.

2.3.4 Conclusions

1. The “data transformation” does not help with the performance, even though it makes the SEC ratios look better. This is probably because the data for different shell sizes and environments are brought closer through transformation. Another possible reason is that the number of GLRT parameters on the “transformed data” is less than that of the “untransformed data.”
2. Except for the cases of “inert on the ground,” the “transformed data” performs more consistently than the “untransformed data,” as the GLRT is trained with a subset of the data, i.e., “Small and Medium Shells Only,” or “Medium Shells Only,” as shown in Tables 2.3-13 to 2.3-16.
3. Surprisingly, the “untransformed data” results in better performance when small sets of data are used in training the GLRT parameters, as when comparing the “Correct ID Rates of Explosives” in Tables 2.3-1, 2.3-13, and 2.3-15.
4. The best performance on the training data among all of the 24 GLRT setups is using the GLRT parameters trained with “dual targets, shell sizes, and environments” applied to the “untransformed data,” as shown in Table 2.3-3.

2.4 Confidence Metrics

In a fielded system, a classification decision (inert versus explosive) is usually made by comparing the algorithm output to a fixed threshold that may be either pre-determined or set in the field by a calibration procedure. Similarly, an identification decision (fill type) is often made by determining which of the hypotheses is most likely given the measured data. For both classification and identification, it is desirable to report the confidence in the decision in addition to the decision itself.

Two approaches for determining the confidence in a decision are presented. The first is a probability-based confidence measure which relates the probability of the algorithm output under the declared hypothesis to the confidence. The second is an entropy-based confidence measure, which assigns a confidence based on the likelihoods of all the hypotheses.

In this discussion, we focus on the entropy-based confidence measure for identification of the fill type. This approach was used for identification among several target types and can be used to identify the particular fill within a classification (such as TNT, RDX, or HMX). Details of the probability-based confidence for classification and identification and for classification using an entropy-based confidence measure are found in the final report from Duke.

2.4.1 Entropy-based Confidence Metric

There are several interpretations of entropy, one being a measure of the uncertainty associated with a partition of a space, with higher entropy corresponding to greater uncertainty. Taking this point of view, a measure of confidence, or certainty, can be developed from entropy. The idea of

using entropy as a confidence metric was previously proposed for assigning a confidence to the classification of pixels in high-resolution, remotely sensed data.

The entropy-based identification confidence measure can also be extended to multiple hypotheses. For an S-element space, the entropy, H, is defined as

$$H = - \sum_{s=1}^S p_s \log p_s$$

where p_s is the probability of element s . When all the elements in the partition are equally likely, the entropy of the partition is $\log S$. Therefore, the normalized entropy measure, H_n , whose values range from 0 to 1, is

$$H_n = \frac{- \sum_{s=1}^S p_s \log p_s}{\log S}$$

and the corresponding entropy-based identification confidence metric, C_E , is given by

$$C_E = 1 - H_n = 1 + \frac{\sum_{s=1}^S p_s \log p_s}{\log S}$$

Again, the confidence is a function of only two probabilities, not three, since $p(H_2|\text{data}) = 1 - p(H_0|\text{data}) - p(H_1|\text{data})$. This confidence measure is intuitively appealing because it is close to zero when the hypotheses are nearly equally likely, and it tends to 1 when one of the hypotheses is dominant. The probability-scaled entropy-based identification confidence metric, C_{Ep} , whose values range from $1/S$ to 1, is

$$C_{Ep} = \frac{1 + (S-1)C_E}{S} = 1 + \frac{1-S}{S} H_n$$

Unlike the probability-based identification confidence measure, the entropy-based measure can be determined without first estimating M-dimensional probability density functions (pdfs).

The probability-based and probability-scaled entropy-based classification confidence measures provide similar confidence values. As shown in the Duke report, plots of these confidence measures as a function of the classification algorithm output and the probability of H_1 (explosive), given the measured data, demonstrate that the probability-scaled entropy-based confidence metric provides a reasonable approximation to the rigorous probability-based confidence metric.

The probability-scaled entropy-based identification confidence metric also provides a reasonable approximation to the probability-based identification confidence metric. In addition, the probability-scaled entropy-based identification confidence metric allows for the calculation of

confidence for all possible values of Φ^* . The probability-based identification confidence metric can be determined only for those values of Φ^* which were encountered when the pdfs were estimated.

2.4.2 Experimental Data Results

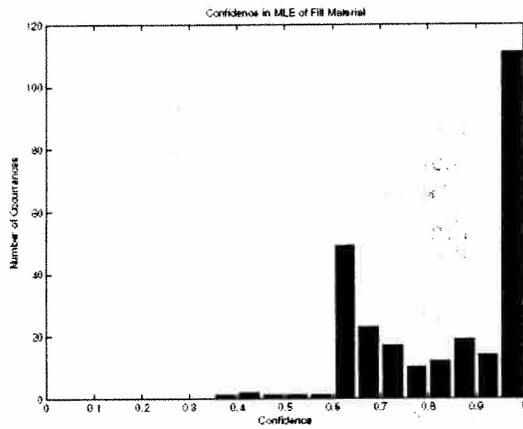
The maximum likelihood (ML) fill material estimation algorithm, previously derived, is applied to SPIDER element counts provided by SAIC (training_mat_full.txt), which were determined for chemical data. The chemicals present in this data set are ANFO, bleach, gasoline, diesel, ammonia, and water.

For the maximum likelihood fill estimation algorithm, the estimated fill material is the fill material, which is most likely given the observed data. The results were determined under two assumptions regarding the covariance structure of the data. The first assumption is that the variables are correlated with the correlation determined by the training data. The second assumption is that the variables are uncorrelated. In both cases, the confidence is determined using the entropy-based identification confidence measure described in the previous section.

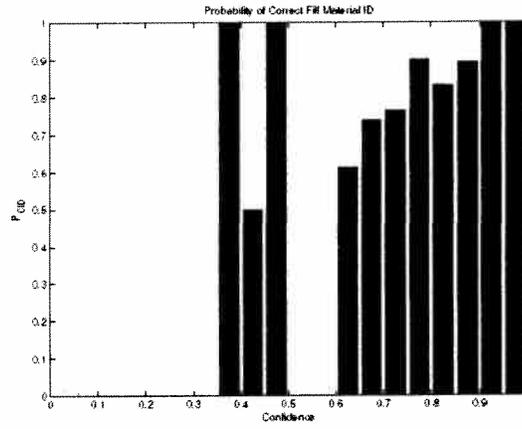
Figures 2.4-1 and 2.4-2 show histograms of the confidence values and the probability of correct identification as a function of the confidence value under the assumption of correlated variables and uncorrelated variables, respectively. Generally, the probability of correct identification increases with the confidence value. It is important to note the number of data points with confidence values within a bin to determine if the associated probability of correct identification is reliable. For instance, if there are only a few cases in which the confidence is within a certain range, then performance of 100% correct identification in that confidence range would be suspect. The figures also show the confidence value versus the probability of the maximum likelihood fill material estimate. The data points (blue asterisks) surrounded by red circles represent the points for which the declared identification is incorrect.

Confusion matrices follow in Figs. 2.4-3 and 2.4-4, again under the assumption of correlated variables and uncorrelated variables, respectively. The average probability of correct identification is 0.855 ($\hat{k} = 0.826$) when correlation between the element counts is considered in the algorithm and 0.738 ($\hat{k} = 0.686$) when the element counts are assumed to be uncorrelated.

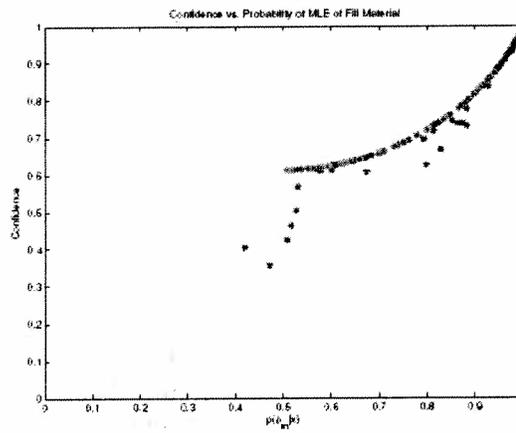
Finally, the results are tabulated and shown in tables in Section 5 of the Duke Final Report 6 for the assumption of correlated variables and for the assumption of uncorrelated variables. Each table lists the results for all the measurements corresponding to one fill material. For each measurement, the measurement number (order in the file provided by SAIC), estimated fill material, probability of the estimated fill material, and entropy-based identification confidence value are listed. The most common confusions are between diesel and gasoline, and water and ammonia. Each of these pairs has similar chemical composition.



(a) Histogram of confidence value.

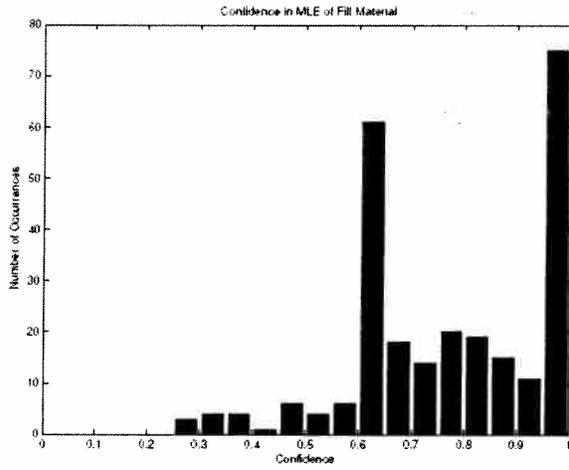


(b) Probability of correct ID.

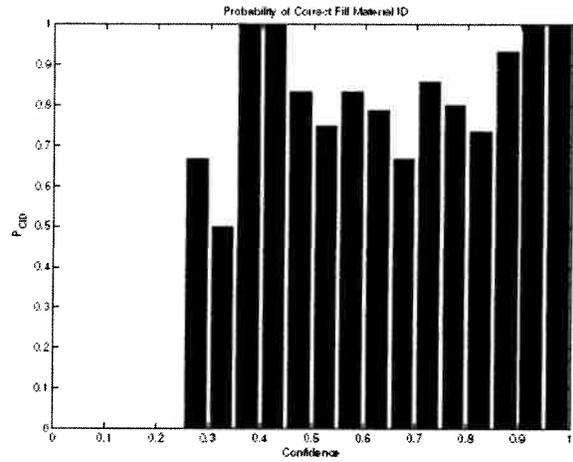


(c) Confidence vs. probability.

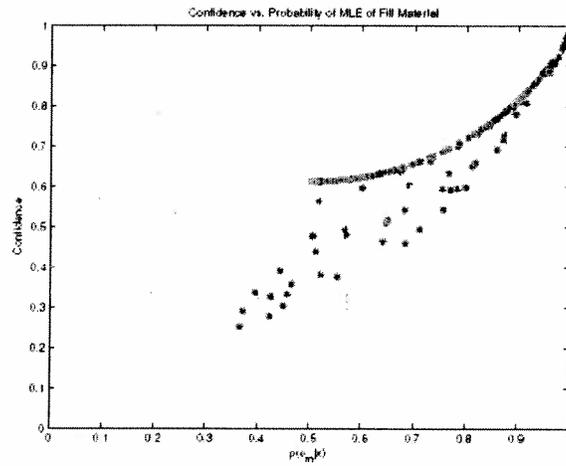
Figure 2.4-1: Identification results for ML estimates of fill material when the correlation between the element counts is included in the algorithm.



(a) Histogram of confidence value.



(b) Probability of correct ID.



(c) Confidence vs. probability.

Figure 2.4-2: Identification results for ML estimate of fill material when the element counts are assumed to be uncorrelated in the algorithm.

		Estimated Fill					
		ANFO	Bleach	Gasoline	Diesel	Ammonia	Water
True Fill	ANFO	1.000					
	Bleach	0.021	0.958	0.021			
	Gasoline			0.750	0.250		
	Diesel			0.209	0.791		
	Ammonia					0.744	0.256
	Water					0.111	0.889

Figure 2.4-3: Identification results for ML estimate of fill material when the correlation between the element counts is included in the algorithm.

		Estimated Fill					
		ANFO	Bleach	Gasoline	Diesel	Ammonia	Water
True Fill	ANFO	0.905					0.095
	Bleach	0.021	0.875	0.021			0.083
	Gasoline			0.705	0.292		
	Diesel			0.395	0.605		
	Ammonia	0.026	0.051			0.539	0.385
	Water	0.089	0.067			0.044	0.800

Figure 2.4-4: Identification results for ML estimate of fill material when the element counts are assumed to be uncorrelated in the algorithm.

2.4.3 Summary

Two approaches for determining decision confidence have been presented: a probability-based confidence measure and an entropy-based confidence measure. The probability-based confidence measure provides a rigorous manner in which to assign a confidence to a classification or identification decision. However, since it utilizes the pdfs of the algorithm output, it is necessary to estimate pdfs, and the estimation of M-dimensional pdfs for identification confidence may not be practical. Therefore, this approach is only truly viable for classification decision confidence. The entropy-based confidence measure is not as rigorously derived as is the probability-based measure, but it is easily calculated and does not require estimating pdfs. In addition, its values range from 0 to 1, rather than 1/S to 1 as does the probability-based metric. Thus, the range of the entropy-based metric is independent of the number of hypotheses, whereas the minimum value of the probability-based metric, 1/S, depends on the number of hypotheses. If the scale of the probability-based metric is more intuitively appealing, the entropy-based metric can be converted to the same scale as the probability-based metric. The probability-scaled entropy-based confidence measure provides an easily computed approximation to the probability-based metric. The entropy-based identification confidence measure was applied to SPIDER element counts provided by SAIC for chemical data.

2.5 Spectral Analysis With PCA

2.5.1 Effects of Background Subtraction

2.5.1.2 Simulations

The background response measured by the PELAN system is usually much larger than the elemental responses of the target. Thus, the background response usually masks, or nearly masks, the target response. Background subtraction has been investigated as a means to eliminate the masking effect of the background response. However, since the background itself is not precisely known, this technique has the potential to introduce more noise into the measured signal and, consequently, may degrade performance. In addition, systematic error may be introduced by background subtraction if the assumed background response subtracted from the measured signal is different from the background response present in the measured signal. The additional noise and/or error resulting from subtracting a background response may adversely impact detection and identification performance.

One method used to analyze PELAN-measured spectra is PCA. PCA is a technique wherein a set of orthogonal basis functions, termed principal components (PCs), are determined from the data. Thus, the model for the measured spectral response, $M(c)$, using PCA is

$$M(c) = \bar{M}(c) + \sum_{k=1}^K W_k B_k(c)$$

where $M(c)$ represents the mean of the data, $B_k(c)$ are the basis functions, or PCs, and W_k are the weighting coefficients associated with each PC. The coefficients W_k are calculated by projecting

the measured spectra onto the PCs and provide a concise set of features for classification and/or identification.

Simulations are performed to evaluate the impact of background subtraction on explosive detection when a GLRT is applied to the PC coefficients, W_k . The results indicate that subtracting a background response prior to applying PCA degrades explosive detection performance, and theoretical analysis explains the reasons this occurs. Details of the simulations described here are found in the Duke Final Report. We provide a summary of the investigation and its results here.

The signals utilized in the simulations are based on elemental and background spectral responses provided by SAIC. Each of the elemental spectral responses is quite distinct with respect to both the other elements and the background response. However, the background response is three orders of magnitude larger than the elemental responses.

The background response provided by SAIC is not associated with any particular background environment. In order to investigate the effects of the target background environment on explosive detection performance, additional simulated background responses were generated for different types of backgrounds based on the ESTCP 2003 data set. This data set provided background spectra for each of the target measurements. All of the background responses provided for each type of background (gravel, sand, soil, table, and wet soil) were averaged to create a simulated background response for each type of background. The average responses are shown in Fig. 16 in the Duke Final Report, along with the background response provided by SAIC. The inset contains a magnified view of the responses for channels 100 through 250. With the exception of the table background, the variability of the background responses is the same order of magnitude as the variation created by the target response.

Using these simulated spectra, the effects of subtracting a local background measurement prior to PCA were investigated. The background responses utilized for training and testing may be the same, or they may be different. Both scenarios are considered in these simulations.

PCA is applied to the simulated training data to determine the principal components and the principal component coefficients. For these simulations, five principal components are utilized for detection. The statistics for the GLRT are determined from the training data principal component coefficients. The testing data is processed by first determining the principal component coefficients corresponding to the principal components found for the training data. The GLRT designed using the training data is then applied to the testing data principal component coefficients. The decision statistic produced by the GLRT is utilized to assess performance through ROC curves.

In summary, the results indicate that subtracting a background measurement prior to applying PCA effectively lowers the signal-to-noise ratio (SNR) by 3dB, and consequently, performance is degraded. In addition, the simulation results indicate the performance is insensitive to the background utilized for training and testing, so it may not be necessary to ensure that the training and testing backgrounds are identical.

2.5.1.2 Evaluation Using PELAN Data

The results of the simulations were tested using PELAN IV data collected at Indian Head in December 2003 and April 2004 on explosive- and inert-filled shells. Additional data on inert-filled shells taken at SAIC-San Diego were also included. The spectra data set (all_shells_spectr.txt) was provided to Duke University to conducting this analysis. A 25% "Don't Know" was used in a tertiary decision. Training was conducted with data collected on a table, then a test was run for data collected on the table. Also, training was conducted on a table and then tested on data collected on other surfaces (sand, soil, and asphalt). For each of these combinations, the following approach was used:

- Repeat the following 20 times
 - Randomly select 80% of table data and train GLRT parameters
 - Randomly select a new 80% of table (or sand/asphalt/soil) data and test GLRT
 - Generate a ROC assuming 25% "Don't Know"
- Average 20 ROCs to determine the average ROC

The PCA parameters were determined with and without the background spectra subtracted from the target spectra. A total of 296 table measurements and 269 non-table measurements were used in this investigation. The training used all shell sizes; that is, no separate parameters were generated for each size group. Also, as was done for the simulations described in the previous section, only the gamma spectra from fast neutron reactions were used. Three principle components were used to train the GLRT (no model parameters). For baseline comparison, the energy using the SPIDER element counts and the energy from the measured spectra to which PCA is applied were calculated and compared to the results. The results are shown in the following figures.

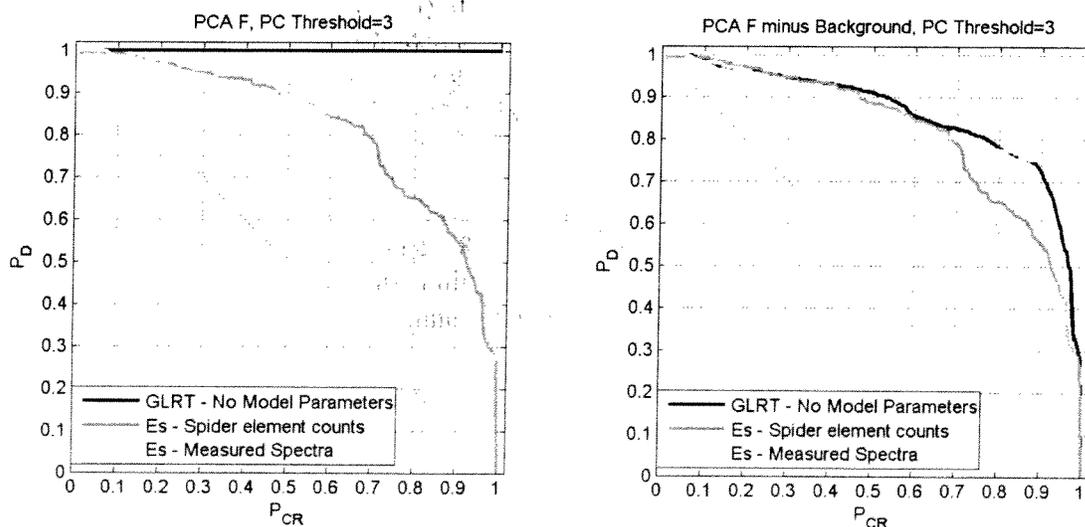


Figure 2.5.1-1. ROC of PCA trained with data taken on a table, then tested on a table. The left ROC plot is with no background subtraction, and the right ROC plot is with background subtracted.

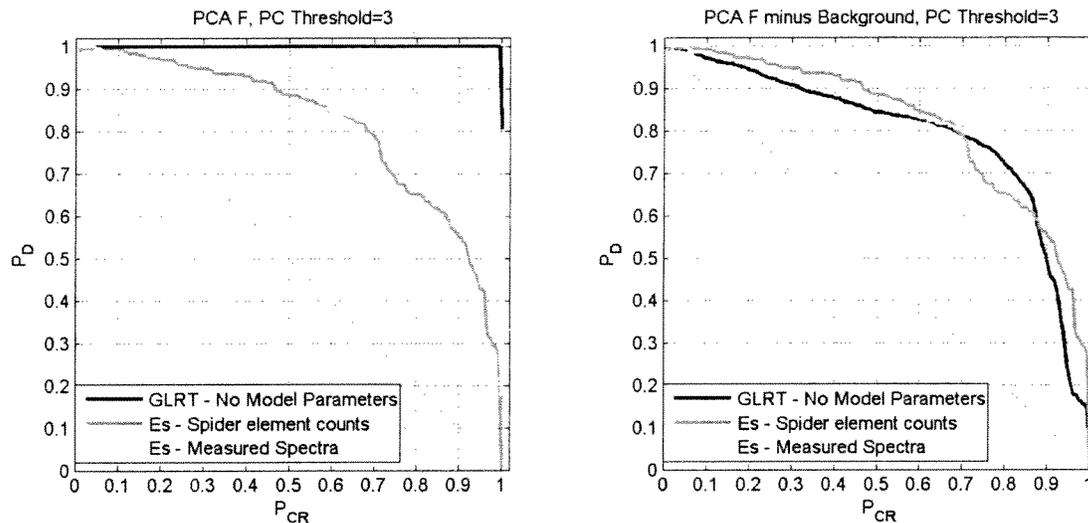


Figure 2.5.1-2. ROC of PCA trained with data taken on a table then tested on sand, asphalt, and soil. The left ROC plot is with no background subtraction, and the right ROC plot is with background subtracted.

The results of this exercise show that subtracting the background spectra greatly reduces the performance when PCA is used to analyze the spectra. Furthermore, using data trained on one environment (metal table) can be used to predict the outcome of data taken on another environment (soil, sand, asphalt) with little change in the performance. This is consistent with the results of the simulations described above. The outcome of this result is that a background run may not be required prior to the target run, saving time and eliminating the need to have an empty shell for a background run.

2.5.2 Variables Affecting Cluster Formation

2.5.2.1 Introduction

PCA is applied to PELAN data for shells containing both inert and explosive fill materials in this work. The far-reaching purpose of this work is to contribute to the understanding of how explosive fill materials can be differentiated from inert fill materials using PC analysis of PELAN signals. PCA is applied with two specific goals in mind: classification of the fill materials (explosive versus inert) and identification of the individual fill materials.

The data for explosive fill materials was collected at NAVEODTECHDIV, Indian Head, December 2003 and April 2004. The data for inert fill materials was collected at SAIC, Rancho Bernardo, spring 2004.

One of the purposes of the PC analysis is to sort data into clusters that can be visualized in three-dimensional plots. The formation of the clusters is dependent upon many variables. For the purpose of classification of fill material, it is desirable to have the PC analysis sort the data into

two clusters, one for explosive materials and one for inert materials. For the purpose of identification of fill materials, it is desirable to have smaller sub-clusters form, one for each fill material.

There are many variables besides the fill material of the shell that affect the formation of clusters in PC analysis. The variables fall into three general categories: data collection variables, data preprocessing variables, and PCA post-processing variables.

1. *Data collection variables*

- The distance from the shell to the PELAN unit
- The size of the shell
- The composition of the background (e.g., the shell may rest on soil or cement or a table)
- The overall environment (e.g., data collection may occur indoors or outdoors)

2. *Data preprocessing variables*

- The selection of PELAN channels to include in the PC analysis
- The subtraction of a background signal
- Mean centering of the data
- Autoscaling of the data

3. *PCA post-processing variables*

- The number of principal components to include in the analysis

It should be emphasized that the subject of this work is not to determine whether a shell's fill material is explosive or inert. The actual decision making is left to prediction algorithms such as GLRT.

In summary, the purpose of this study was to preprocess, process, and analyze PELAN data for a particular data set, using PC analysis, with the goal of differentiating explosive materials from inert materials. Furthermore, we wanted to understand the effect of the three types of variables described above on the analysis and formation of PCA clusters.

2.5.2.2 Description of Variables Affecting Cluster Formation

In the introduction, it was noted that there are three types of variables that affect cluster formation: data collection variables, data preprocessing variables, and PCA post-processing variables. In this section, these variables are described in detail.

Data Collection Variables

Data collection variables include those variables that come into play as the data is being collected. Four of these variables are described below.

1. *The distance of the shell to the detector:* Almost all shells in the data set were placed 2 inches from the detector, with the exception of a few that were placed at a distance of 1 inch from the detector. To eliminate this variable, the shells placed at 1 inch were excluded from the PC analysis.
2. *The size of the shells.* The shells are labeled by size according to their diameter measured in millimeters. The shells are described as small, medium, and large:

Small < 90mm,
90mm <= Medium < 120mm,
Large >= 120mm.

3. *The composition of the background.* The majority of the shells were placed on either an aluminum table or soil for collection of the PELAN data. Some of the shells were placed on cement, sand, grass, wet grass, wet asphalt, dry dirt test bed, and wet sand test bed. These background variables are labeled in this study as
 - Table
 - Soil
 - Other
4. *The overall environment.* Some of the data were collected indoors and some outdoors. The data collected indoors tends to have a larger background signal.

Preprocessing the Data

After the data has been collected but before it is sent to the PC analysis algorithm, it may be subjected to preprocessing. There are four possible preprocessing steps:

1. *Channel selection.* The PELAN unit collects 1,024 channels of data. Certain channels map to atomic elements of interest for explosives detection. Other channels are not of interest. For this study, the effect of channel selection on PC cluster formation was not considered, and the channel selection was fixed to be the range of channels 50 to 450 and channels 557 to 962.
2. *Subtraction of background signal.* The background signal is the signal obtained when the shell is not present. PC analysis was performed on data both with and without the background subtracted. This is an important variable to consider, since the collection of background signal is time-consuming in the field. One issue with background subtraction is noise. The background signal has noise associated with it, and by subtracting, it could effectively double the noise in the sample data. Another issue is that the background signal data is not always collected at the same time as the shell data and, thus, may not be an accurate reflection of the true background.

3. *Mean centering.* Mean centering is performed on the data by calculating the mean for each channel, and subtracting. Mean centering can be important to efficient PC analysis. PCA was performed on the data both with and without mean centering.
4. *Autoscaling.* Autoscaling is the application of both mean centering and variance scaling to the data. Variance scaling is performed by dividing each channel by the standard deviation of that channel. Autoscaling is typically applied to data so that scale does not dominate the analysis. Autoscaled data is unitless, and autoscaling is typically used when the data is known to be of different types (units) or of greatly different ranges. In this study, PC analysis was performed on the data both with and without autoscaling.

Post-Processing the Data

Post-processing of the data occurs after PC analysis has been performed. At this point, the number of principal components to include in the cluster formation must be decided. Typically, the first few components are included in the analysis. For the work in this report, the first three principal components are displayed. There is a section of the report dedicated to post-processing, where the effect of additional components is explored.

2.5.2.3 Overview of PELAN Data and the Interpretation of PCA Plots

It is assumed that the reader is familiar with PCA techniques. However, before proceeding further, a very brief overview of the PELAN data, the PC analysis techniques, and interpretation of PCA plots is in order.

For each shell, fill material, and background, the PELAN unit generates a data sample point. The sample point consists of 1,024 channels of data, the units of the data being "counts." In this study, as a preprocessing step, the number of channels is reduced to 807, and the other channels are ignored.

Suppose there are a total of 500 samples in the data set. For the PCA, all of the data is placed in a 500 by 807 matrix, called the data matrix. Each row of the matrix corresponds to a sample point. Preprocessing steps are applied to this matrix. As an example of preprocessing, to mean center the data, the mean of each column of the matrix is calculated and then subtracted from each element in that column.

PCA begins with a singular value decomposition on the data matrix. The principal components are the singular vectors of the data matrix. The principal components have 807 elements, which is equal to the number of data channels, and the principal components span what is called the sample space. Each principal component is associated with a singular value; the principal components are listed in descending order according to the magnitude of the singular value.

The principal components have the property that they are orthogonal to one another, and their direction describes the variance in the data. The first principal component accounts for the

largest percentage of variance in the data. The plots in this study are comprised of the first three principal components.

Each PCA plot in this report is a display of the orthogonal projection of the sample points in the data set onto the first three principal components. The clustering of the sample points gives a visual indication of how close the sample points are to one another.

The units on the PCA plots axes are “counts,” the same as the PELAN data. The exception is for the case where the data has been autoscaled, then the principal component axes are unitless. Sometimes the counts appear as negative numbers; this is due to the mean centering.

Three-dimensional visualization is important to this study. To aid in visualization, stem plots are employed at times. A stem plot is a three-dimensional plot of the sample points, (x, y, z), where each point has a tail. The end of the tail always touches the xy plane.

2.5.2.4 Preprocessing Studies

The goal of the preprocessing studies is to determine the best combination of background subtraction, mean centering or autoscaling to perform on the data before applying the PC analysis. The following six combinations of preprocessing techniques are considered.

1. *No preprocessing*
2. *Background subtraction*
3. *Mean centering*
4. *Autoscaling*
5. *Mean centering with background subtraction*
6. *Autoscaling with background subtraction*

This preprocessing study was restricted to large shells (≥ 120 mm diameter) since they were expected to produce the best signal-to-noise ratio of all the shells. All possible backgrounds (table, soil, other) were allowed so that the effect of background subtraction could be determined effectively. The ultimate goal of PCA applied to PELAN data is to separate explosive fill materials and inert fill materials into separate clusters, thus, the effect of the six combinations on the PCA clustering of inert and explosive fill materials is used as a benchmark of comparison.

The following stem plots show the large shell data in sample space for each of the six preprocessing combinations. Sample points for empty shells and shells filled with inert materials are green and sample points for shells filled with explosive materials are red.

1. No preprocessing

Data points representing the empty/inert fill materials (green) cluster separately from the data points for explosive fill materials (red) with the exception of three empty 152mm shells on the bottom right of Figure 2.5.2 - 1.

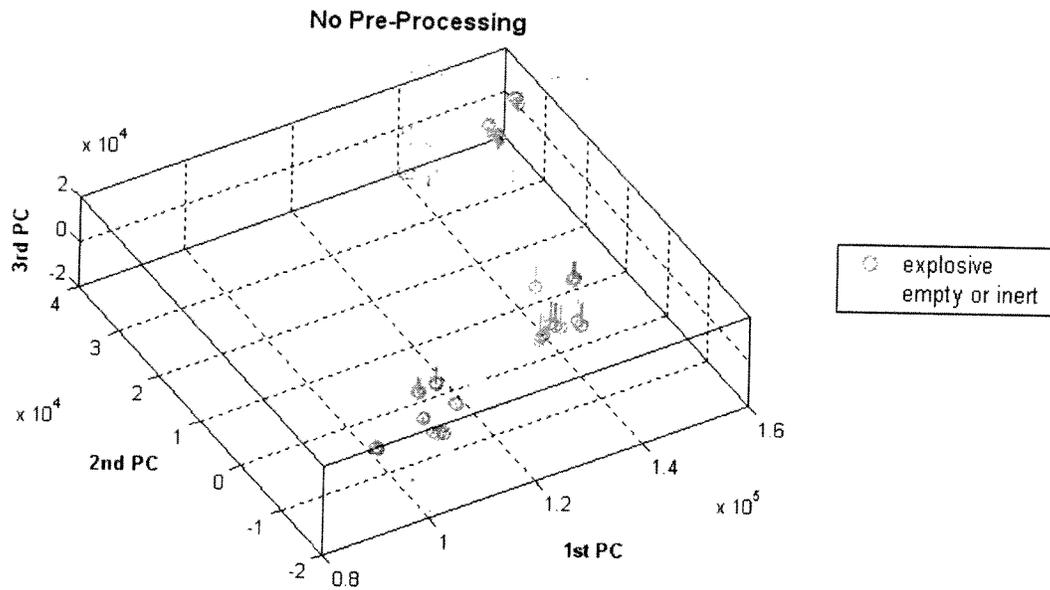


Figure 2.5.2 – 1: Stem plot of the first three principal components for large shells on any background. There is no preprocessing of the data.

2. Background subtraction

Data points representing the inert fill materials (green) cluster separately from the data points for explosive fill materials (red), although the clusters are not as distinct as they are in the no preprocessing case (No. 1). It appears that background subtraction alone may not be as effective as others for solving the identification problem, since the individual clusters, which correspond to individual fill materials, are not as distinct. However, the method may be suitable for the classification problem since the red and green clusters are distinct.

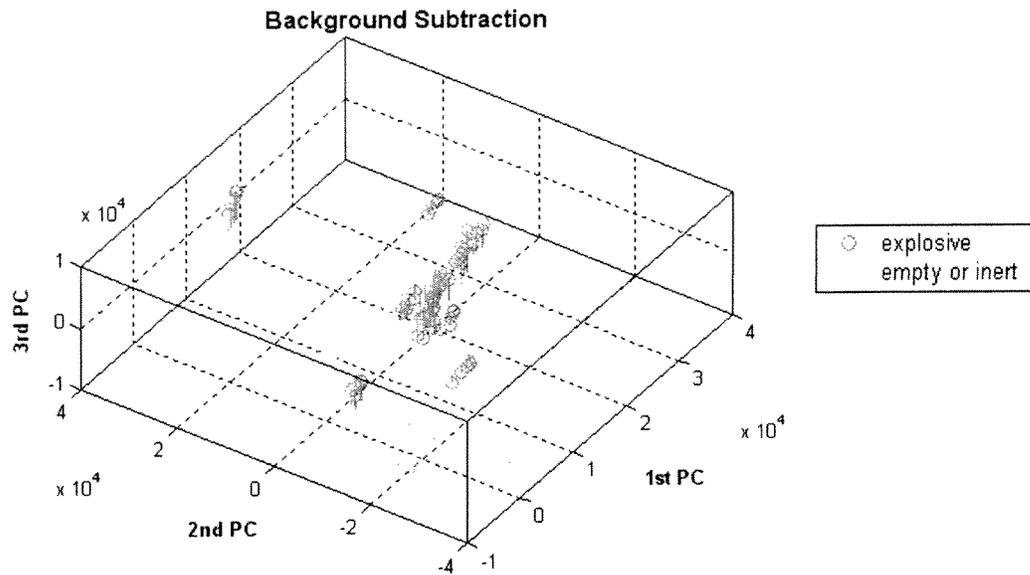


Figure 2.5.2 – 2: Stem plot of the first three principal components for large shells on any background. Background signal is subtracted from the data.

3. Mean centering

Data points representing the inert fill materials (green) cluster separately from the data points for explosive fill materials (red), again, with the exception of three 152mm empty shells, as seen in case No. 1. Note that there are two large inert clusters, one largely consisting of data taken with soil as the background and the other with a table as the background. This method of preprocessing appears suitable for both the classification and identification problems since the individual clusters are so clearly defined.

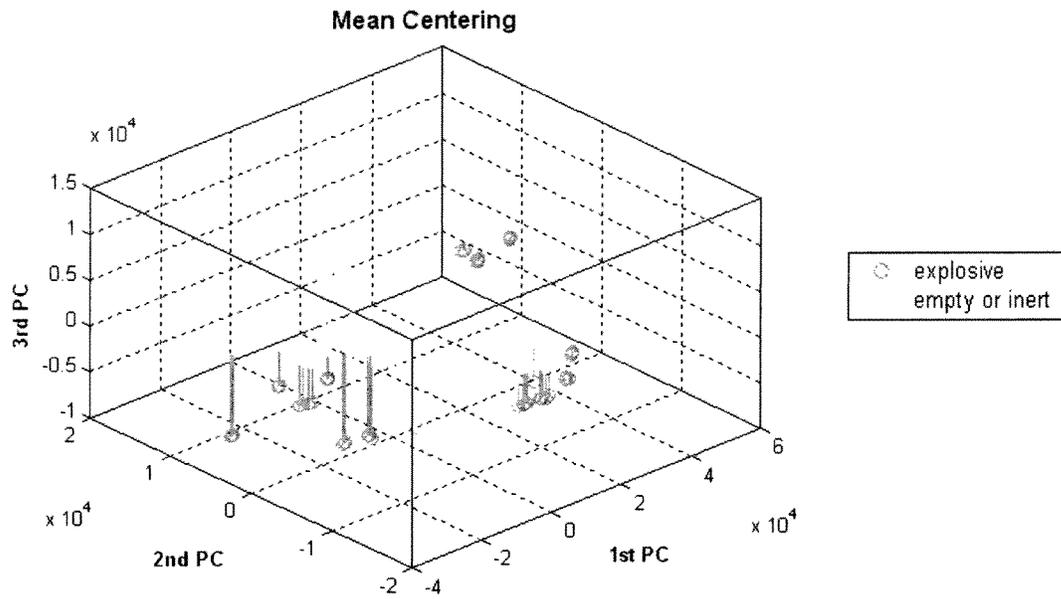


Figure 2.5.2 – 3: Stem plot of the first three principal components for large shells on any background. The data is mean-centered.

4. Autoscaling

Data points representing the inert fill materials (green) cluster separately from the data points for explosive fill materials (red), again, as seen in cases No. 1 and 3, with the exception of three empty 152mm shells near the center of the plot.

Note that the two large inert clusters are present, as in the mean-centering case (No. 3), but are even more distinct, one largely consisting of data taken with soil as the background and the other with a table as the background.

This method of preprocessing appears suitable for both the classification and identification problems since the individual clusters are very clearly defined.

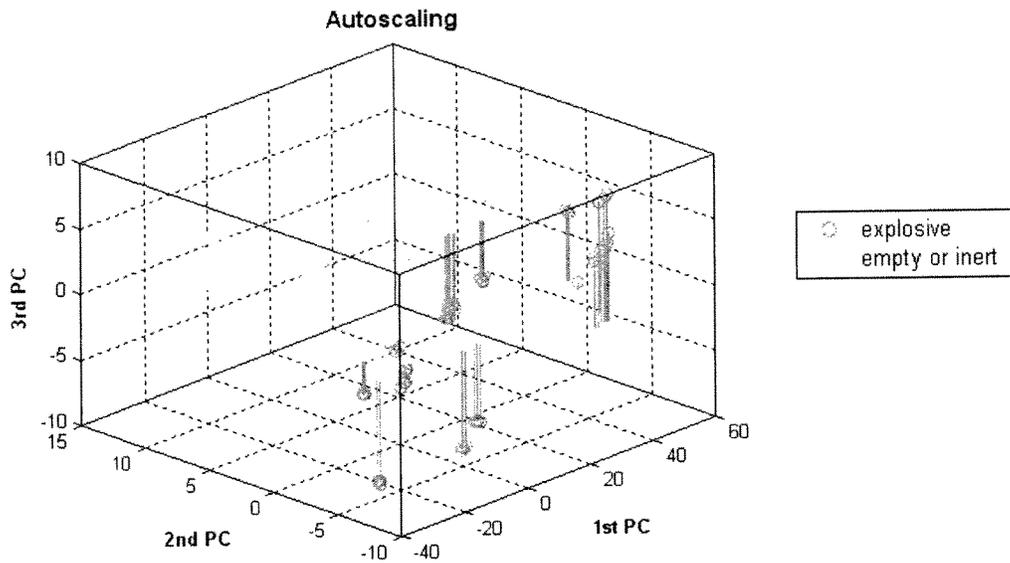


Figure 2.5.2 – 4: Stem plot of the first three principal components for large shells on any background. The data is autoscaled.

5. Mean centering with background subtraction

Data points representing the inert fill materials (green) cluster separately from the data points for explosive fill materials (red). This case is similar to the background subtraction case (No. 2).

As in case No. 2, it appears that this preprocessing method may not be as effective as others for solving the identification problem, since the individual clusters, which correspond to individual fill materials, are not very distinct. However, the method may be suitable for the classification problem since the red and green clusters are distinct.

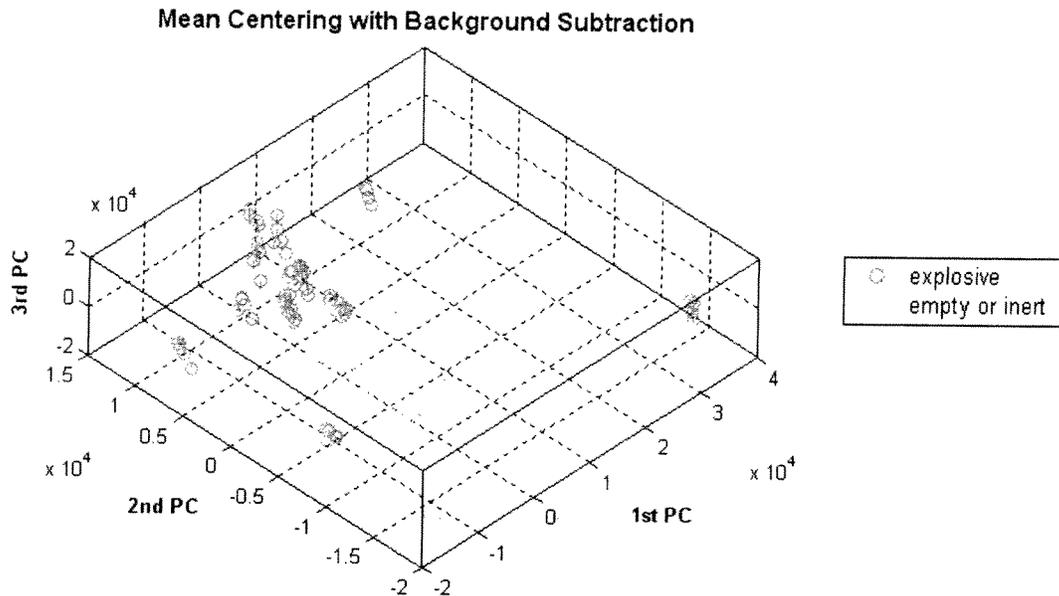


Figure 2.5.2 – 5: Stem plot of the first three principal components for large shells on any background. The data is mean-centered with background subtraction.

6. *Autoscaling with background subtraction*

Autoscaling with background subtraction (No. 6) appears to give the best separation of explosive fill materials from inert fill materials for this data set. Autoscaling removes the effect of signal strength, and all of the inert fill materials cluster tightly together. Note that the 152mm empty shells are in the inert cluster. This method of preprocessing the data appears to be the best for both classification and identification of the fill materials.

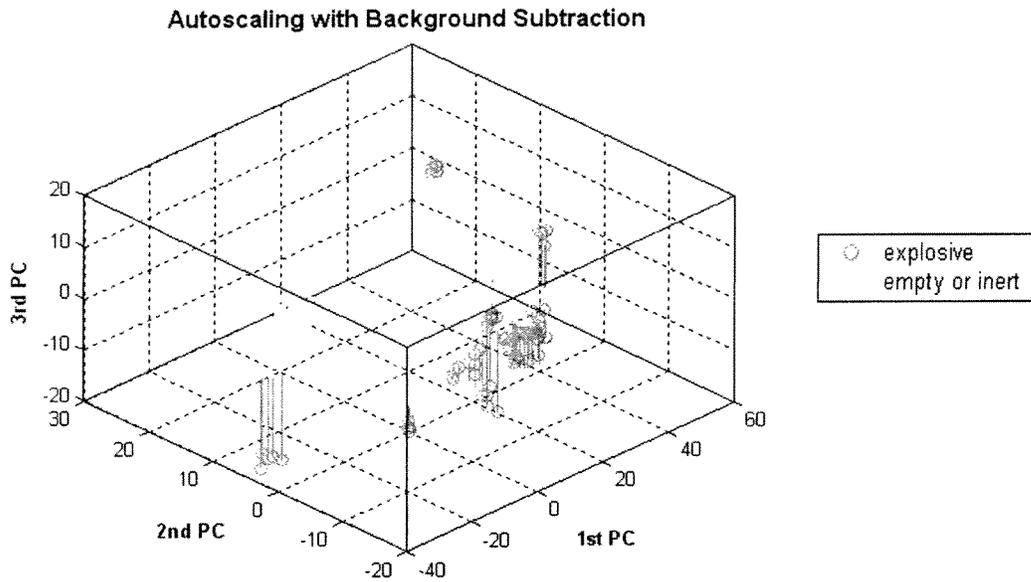


Figure 2.5.2 – 6: Stem plot of the first three principal components for large shells on any background. The data is autoscaled with background subtraction.

Summary of the Six Preprocessing Combinations

For cases No. 1, 3 and 4, where there is no background subtraction, a good separation of explosive and inert fill materials occurs, with the exception of three data points representing inert materials that fall near the explosives. These three data points correspond to empty 152mm shells.

For cases No. 2, background subtraction, and No. 5, mean centering with background subtraction, there is separation of explosive and inert fill materials, which is useful for the classification problem, but the individual clusters may not be distinct enough for the identification problem.

Case No. 6, autoscaling with background subtraction, provides the best separation of inert materials from explosive materials and preserves individual clusters for the identification problem. Since autoscaling tends to remove the element of signal size, it may eliminate the effect of shell size on cluster formation, which is a key variable in the analysis.

Autoscaling with background subtraction is worthy of further investigation, however, that work is not pursued here. There are two main issues to consider with background subtraction: first, the background must be accurate, since the background signal dwarfs the shell signal; and second, the collection of the background signal slows down work in the field.

These six examples show that subtracting local background may not be necessary in order to separate explosive fill materials from inert fill materials. This is an important finding because ignoring the local background greatly simplifies the data collection process in the field.

Elaboration of Case No. 3 (Mean centering without background subtraction)

All further studies in this work are conducted with mean centering as in case No. 3. This method is selected for several reasons: the good separation of explosive and inert fill materials (with the exception of 152mm empties), the tight clustering of the inert data points, the simplification of the data collection process because background signal is not required, mean centering allows efficient PC analysis.

The last reason needs elaboration: Mean centering is important to efficient PC analysis. When the data is not mean-centered, the first principal component vector describes the direction from the origin to the cloud of data. The second principal component is constrained to be orthogonal to the first and cannot orient itself along the length (maximum variation) of the cloud of data. With mean centering of the data, the data cloud is shifted to the origin, and the first principal component effectively describes the length of the cloud of data.

In addition, previous studies at Duke University (Duke University, Final Report for SERDP, Part II) have indicated that mean centering without background subtraction is a promising preprocessing combination for classifying explosive and inert materials.

2.5.2.5 Studies on Empty Shells

A study of all the empty shells was performed in order to determine the role of shell size and environment in the PC analysis. First, the background environment was fixed for these studies as soil or metal table. The PC analysis shows that for a fixed environment (soil or metal table), the shells clustered in sample space according to their size. Then the data for the soil and metal table backgrounds were combined in a single data matrix and analyzed together.

Empty Shells on Soil Background

The first case studied was empty shells of any size on soil background. Figure 2.5.2 – 7 clearly demonstrates that the shells cluster according to size.

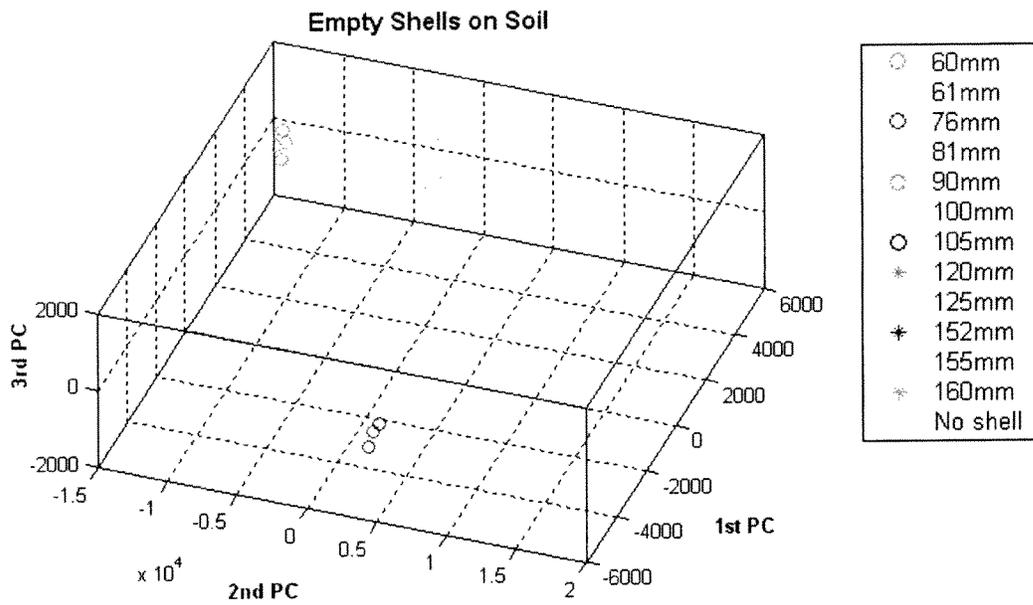


Figure 2.5.2 – 7: Graphical display of the first three principal components for empty shells on soil background. The data is mean-centered.

Empty Shells on Table Background

The second case studied is empty shells of any size on table background. Figure 2.5.2 – 8 clearly demonstrates that the shells cluster according to size, with the exception of overlap of the 61mm on the 81mm shells. In addition, some of the shells of the same size form more than one cluster, the blue 76mm shells for example; this may be due to the data collection process, where some of the data on the table was collected indoors and some outdoors.

Notice that “No Shell” data is included in the plot. This data consists of just the table background.

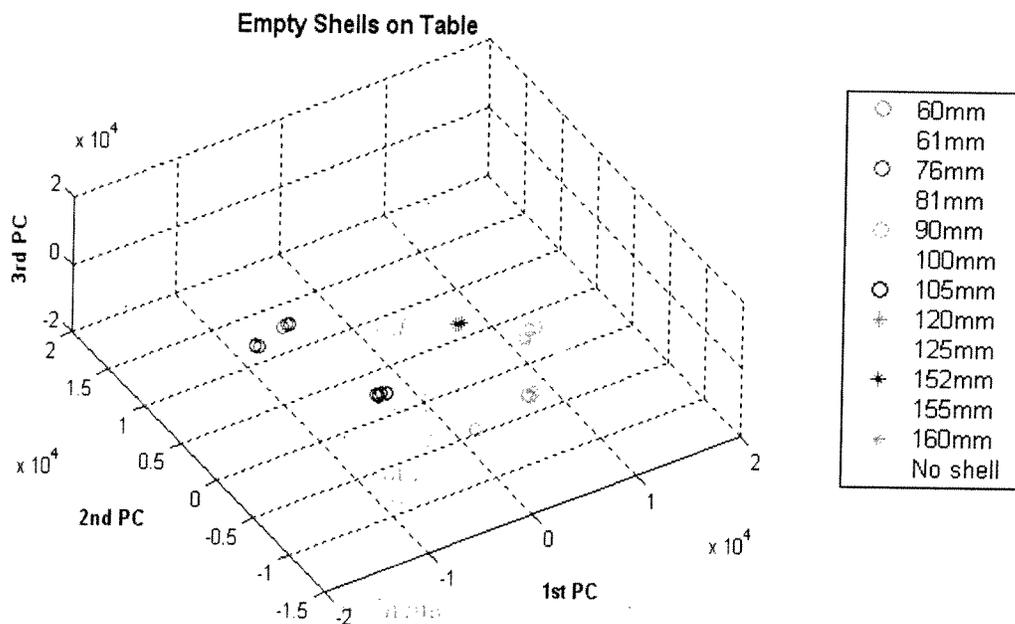


Figure 2.5.2 – 8: Graphical display of the first three principal components for empty shells on table background. The data is mean-centered.

Empty Shells on Soil or Table Background

The third case studied is empty shells of any size on both soil and table backgrounds. Figure 2.5.2 – 9 again clearly demonstrates that the shells cluster according to size, with the exception of overlap of the 61mm and the 81mm shells. In addition, some of the shells of the same size form more than one cluster; this may be due to the data collection process, where some of the data on the table was collected indoors and some outdoors.

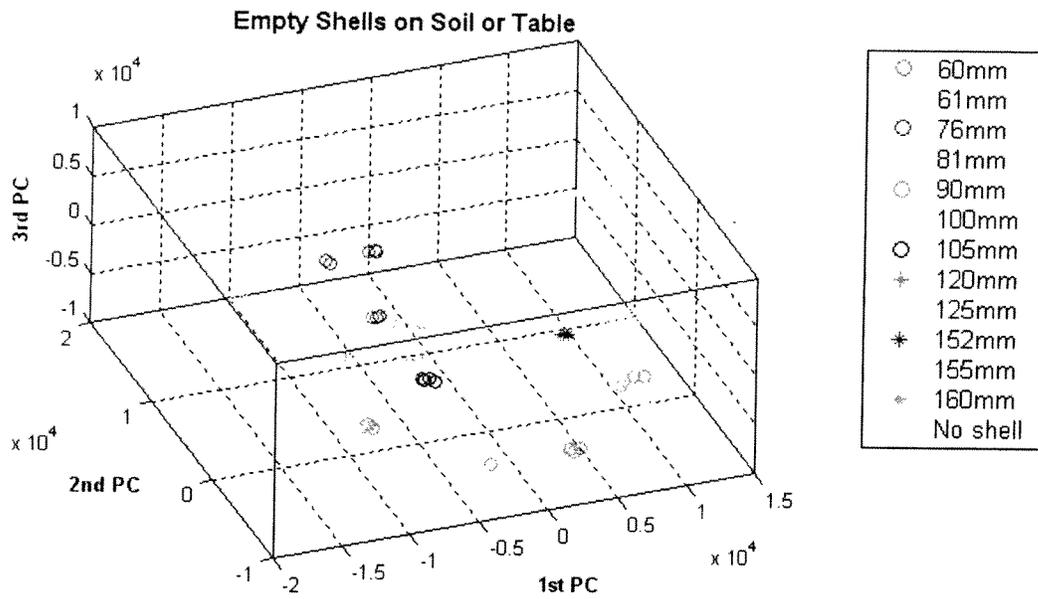


Figure 2.5.2 – 9: Graphical display of the first three principal components for empty shells on table or soil background. The data is mean-centered. Color-coding is according to shell size.

To further investigate the relative effect of background on empty shell clusters, Figure 2.5.2 – 10 was produced. This Figure 2.5.2 – 10 can be superimposed on Figure 2.5.2 – 9. Figure 2.5.2 – 10 shows the shells displayed by background type, either soil or table.

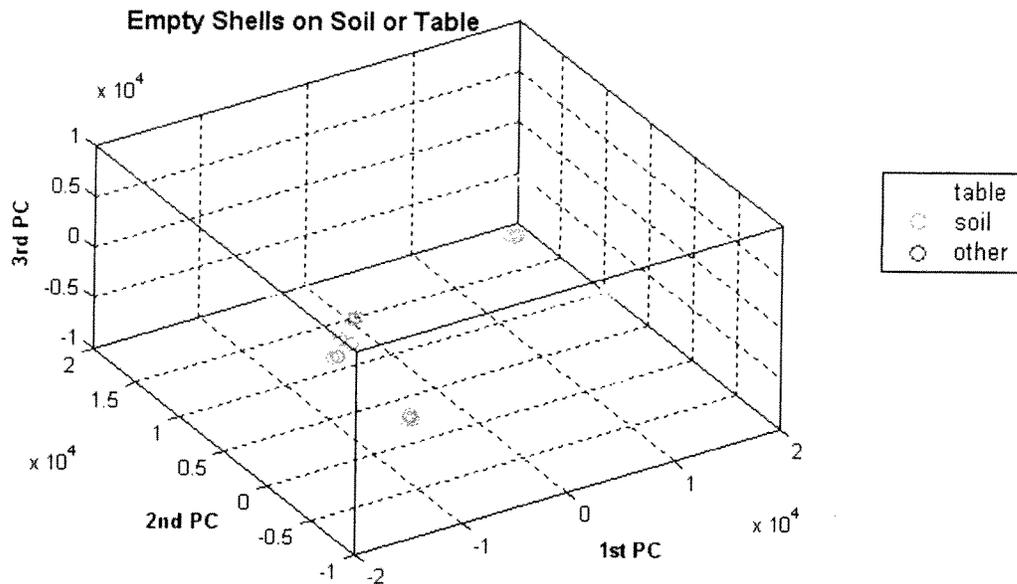


Figure 2.5.2 – 10: Graphical display of the first three principal components for empty shells on table or soil background. The data is mean-centered. Color-coding is according to background.

This empty shell study is inconclusive as to the dominance of background over shell size in the PC analysis. No obvious conclusions are drawn from the two figures and further investigation is warranted. For example, the effect of indoor versus outdoor data collection could be taken into account.

2.5.2.6 Studies on Inert Fill Materials and Empty Shells

Empty shells together with inert fill materials are studied to determine how they cluster in sample space relative to one another. The inert fill materials included in the study are sand, cement, plaster of Paris, paraffin, and wax. Soil and table backgrounds are considered. Shells of all sizes are included.

Three PCA plots are given to illustrate the relative impact of background, shell size, and fill material on the formation of the PCA clusters.

First, Figure 2.5.2 – 11 shows the separation of the sample points according to the two backgrounds, soil and table. Figure 2.5.2 – 11 shows the sample space divided into two distinct regions, soil (magenta) and table (cyan). There are a few exceptions: three 90mm empty shells in the region corresponding to soil (magenta) that appear in the region corresponding to table (cyan), and one 61mm empty shell on table background that appears in the soil region.

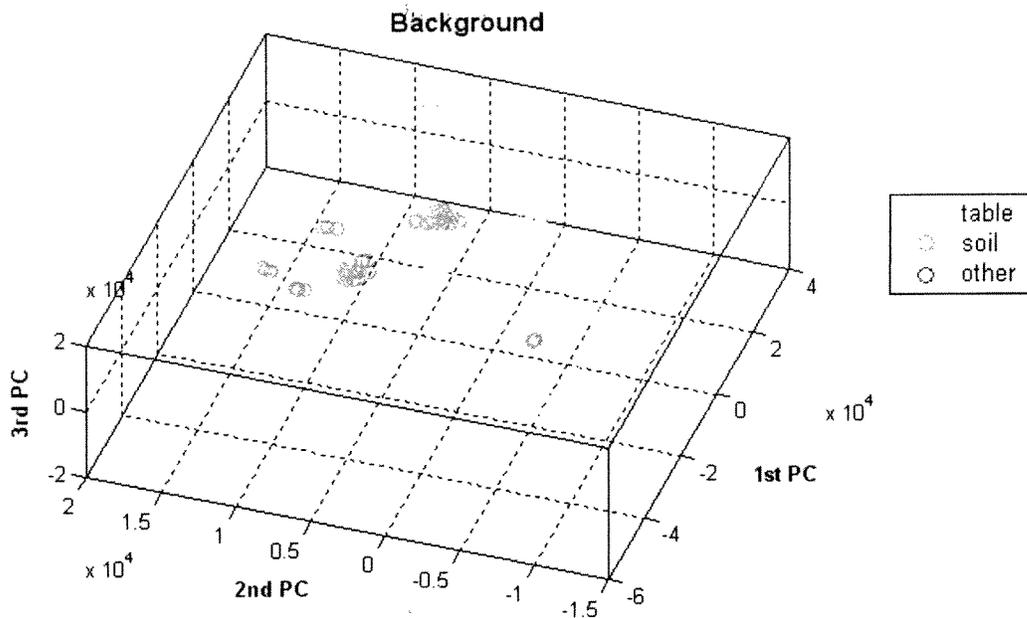


Figure 2.5.2 – 11: Graphical display of the first three principal components for empty shells and shells filled with inert materials on table or soil background. The data is mean-centered. Color-coding is according to background type.

Second, Figure 2.5.2 – 12 shows the separation of the sample points according to shell size. These are the identical sample points that are plotted in Figure 2.5.2 – 11, but they are now color-coded according to shell size. Figure 2.5.2 – 12 shows that within the two distinct regions corresponding to background, the shells tend to cluster according to size. There are some exceptions: the 61mm and 81mm do not always form separate clusters.

Note that most of the shell sizes form two clusters. Take the 155mm shells designated by yellow asterisks, for example, near the top of the plot. They form two separate clusters, and by comparing to Figure 2.5.2 – 11, it is evident the one cluster is on a soil background and the other on a table background.

From Figures 2.5.2 – 11 and 2.5.2 – 12, it may be concluded that the background takes precedence over shell size in cluster formation for empty/inert shells.

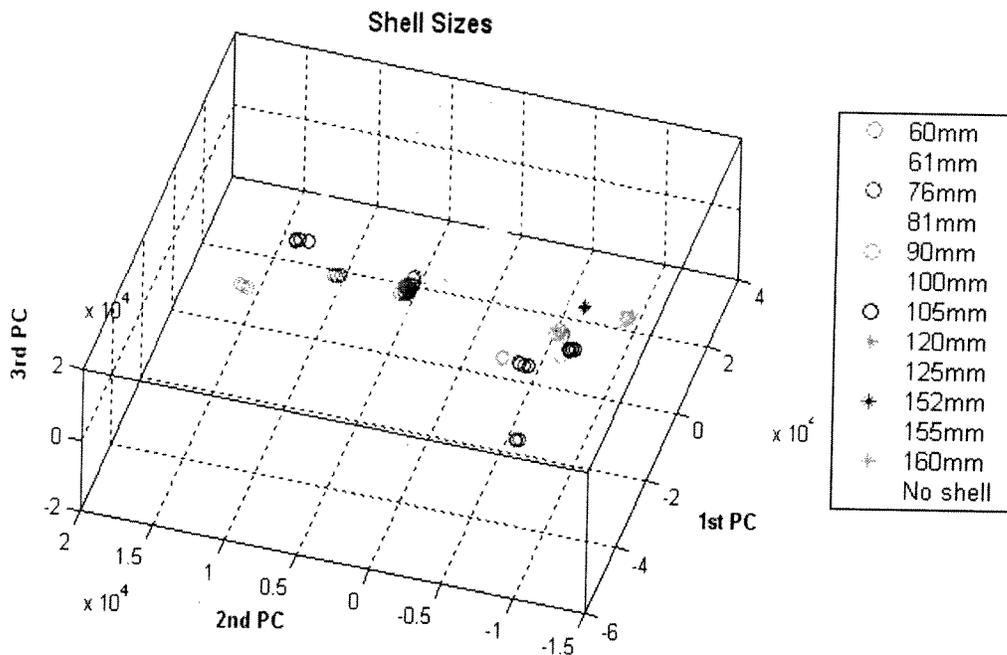


Figure 2.5.2 – 12: Graphical display of the first three principal components for empty shells and shells filled with inert materials on table or soil background. The data is mean-centered. Color-coding is according to shell size.

Third, Figure 2.5.2 – 13 shows the separation of the sample points according to fill material. These are the identical sample points that are plotted in the previous two figures, but they are now color-coded according to fill material. Not all of the sample points are plotted, due to software limitations. Figure 2.5.2 – 13 shows that the sample points do not cluster according to inert fill material; each cluster is comprised of multiple fill materials. This is an important result. It implies that inert fill materials cannot be distinguished from one another with this PCA methodology.

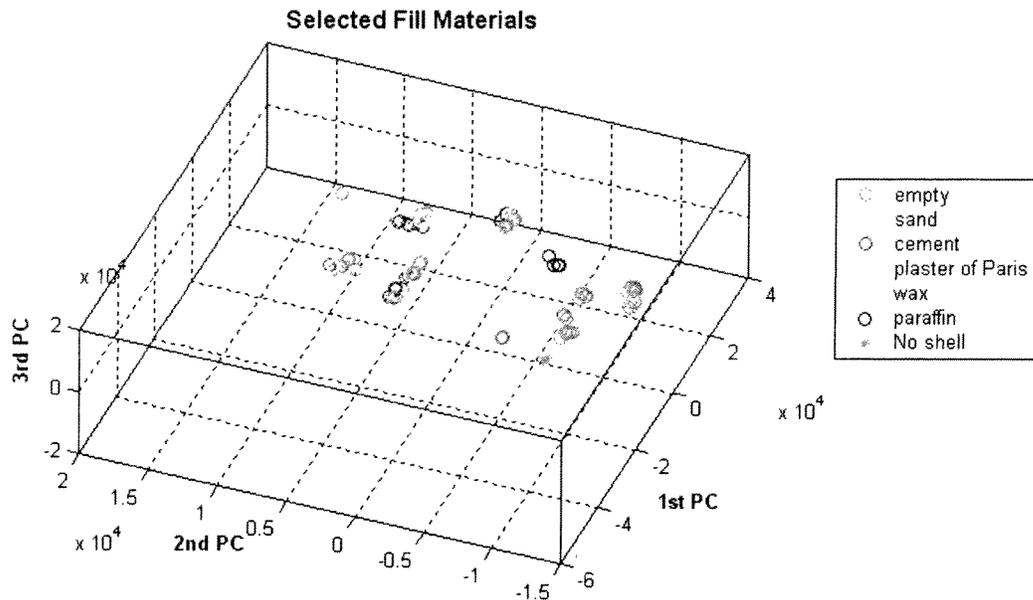


Figure 2.5.2 – 13: Graphical display of the first three principal components for empty shells and shells filled with inert materials on table or soil background. The data is mean-centered. Color-coding is according to fill material. Not all sample points are displayed, due to software limitations.

In summary, comparing Figure 2.5.2 – 13 with Figures 2.5.2 – 12 and 2.5.2 – 11 shows that the sample points divide into two regions according to background and then, within these regions, cluster according to size. The inert fill materials do not affect the formation of the clusters, and (with the exceptions noted for Figure 2.5.2 – 12) the shells cluster according to size. From these two figures, it may be concluded that the background takes precedence over size, and size takes precedence over empty/inert fill material.

In other words, for this data set, with few exceptions, shells filled with sand, cement, plaster of Paris, paraffin and wax are similar to the response from empty shells of the same size and background. This similarity may be due to the small C/H ratio (<1) for inerts compared to large C/H ratios (>1) typical for explosives.

2.5.2.7 Studies of Explosive Fill Materials versus Empty/Inert Fill Materials

The ultimate goal of this work is to understand how to best apply PC analysis to PELAN data so that explosive fill materials may be differentiated from inert fill materials. In this section, PC analysis is applied to classify the data into explosive and inert fill materials. It was determined that shell size is the largest factor in successfully classifying the shells into explosive and empty/inert fill materials. Shells described as large ($\geq 120\text{mm}$ in diameter) were classified with the most success, and this success is attributed to the greater signal-to-noise ratio. Classification was studied for three cases: large shells; large and medium shells; and large, medium, and small shells. All backgrounds are included (soil, metal table, grass, wet grass,

metal table, wet asphalt, dry dirt test bed, wet sand test bed) as well as all inert fill materials (sand, cement, plaster of Paris, wax, paraffin) and all explosive fill materials.

Large Shells

Shells in the data set described as large ($\geq 120\text{mm}$ in diameter), on any background, are successfully separated into two very distinct regions by PCA: one region corresponding to explosive fill materials and one corresponding to inert fill materials together with empty shells. There are three data points that do not separate, which are 152mm empty shells.

Figure 2.5.2 – 14 shows this very promising result for the classification problem.

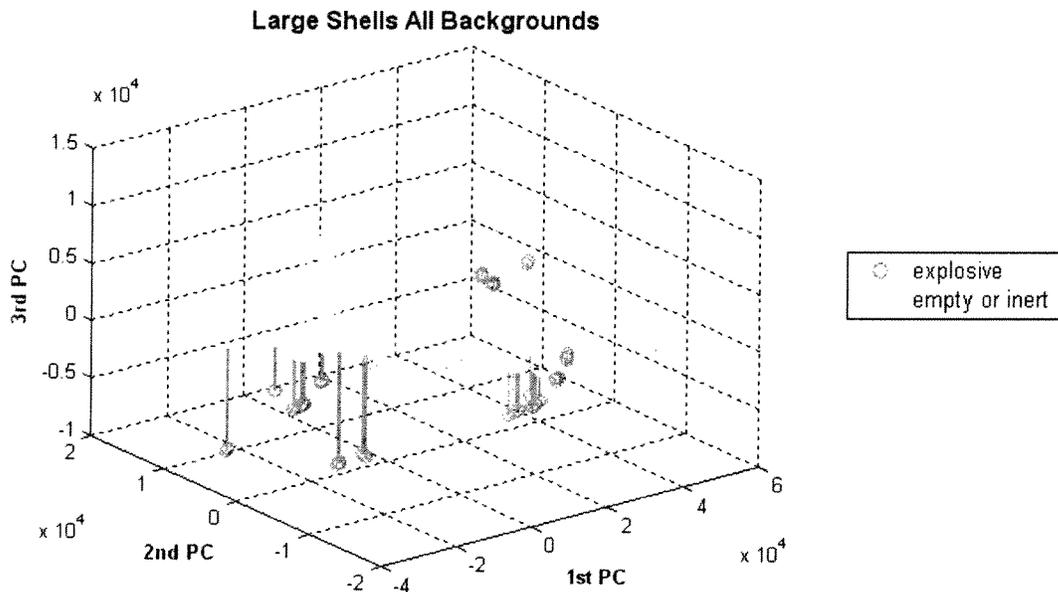


Figure 2.5.2 – 14: Stem plot of the first three principal components for large shells on any background. The data is mean-centered.

Medium and Large shells

Shells in the data set described as large and medium ($\geq 90\text{mm}$ in diameter), on any background, are successfully separated into two distinct regions by PCA, with a few exceptions.

Figure 2.5.2 – 15 shows the PCA plot for medium and large shells. The 152mm empty shells are in the explosive region as described above, and in addition, three new clusters of inert shells have

formed, each corresponding to 152mm shells, one for wax-filled shells on a table, one for empty shells on a table, and one for empty shells on a dry dirt test bed.

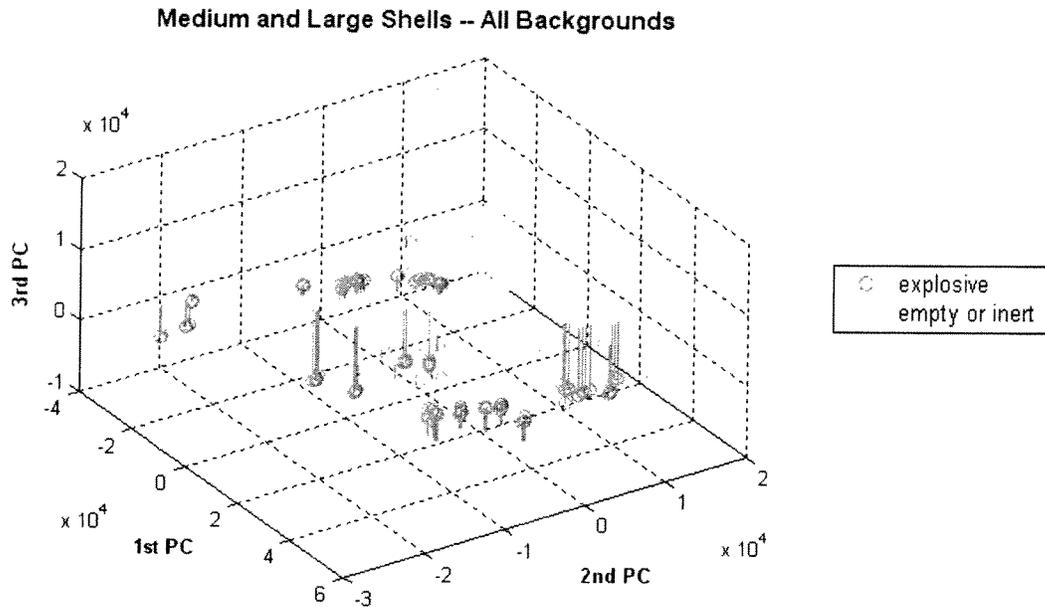


Figure 2.5.2 – 15: Stem plot of the first three principal components for large and medium shells on any background. The data is mean-centered.

Small, Medium and Large Shells

Shells in the data set described as large, medium, and small (≥ 60 mm in diameter), on any background, do not successfully separate into two distinct regions by PCA with this methodology.

There is much interleaving of the inert and explosive fill materials, as is shown in Figure 2.5.2 – 16. This failure to separate is most likely due to the small signal generated by the smaller sized shells.

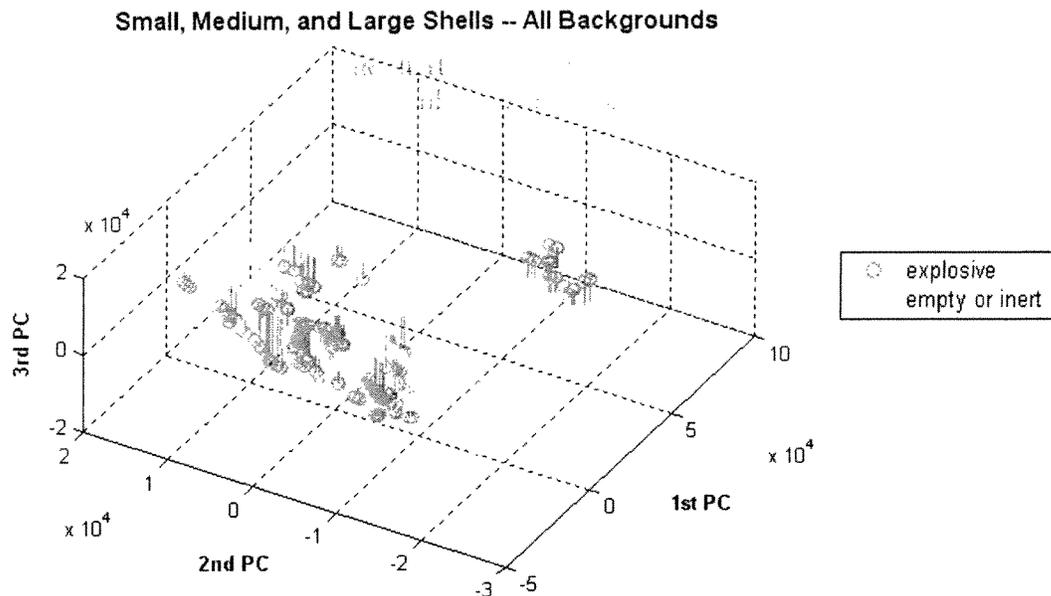


Figure 2.5.2 – 16: Stem plot of the first three principal components for large, medium, and small shells on any background. The data is mean-centered.

2.5.2.8 Studies on the Number of Principal Components to Include in PC Analysis

In this section, an example is given that demonstrates the effect of adding an additional component to the PC analysis. The example is of medium and large shells on any background, with the goal of separating explosive fill materials from empty/inert fill materials.

Figure 2.5.2 – 17 is a plot of the first three principal components. Note that there are two large inert clusters (green), and that subjectively, the lower cluster appears “close” to the explosive cluster (red).

Figure 2.5.2 – 18 is a plot of the second, third, and fourth principal components. In this plot, the two large inert clusters are clearly separate from the explosive region.

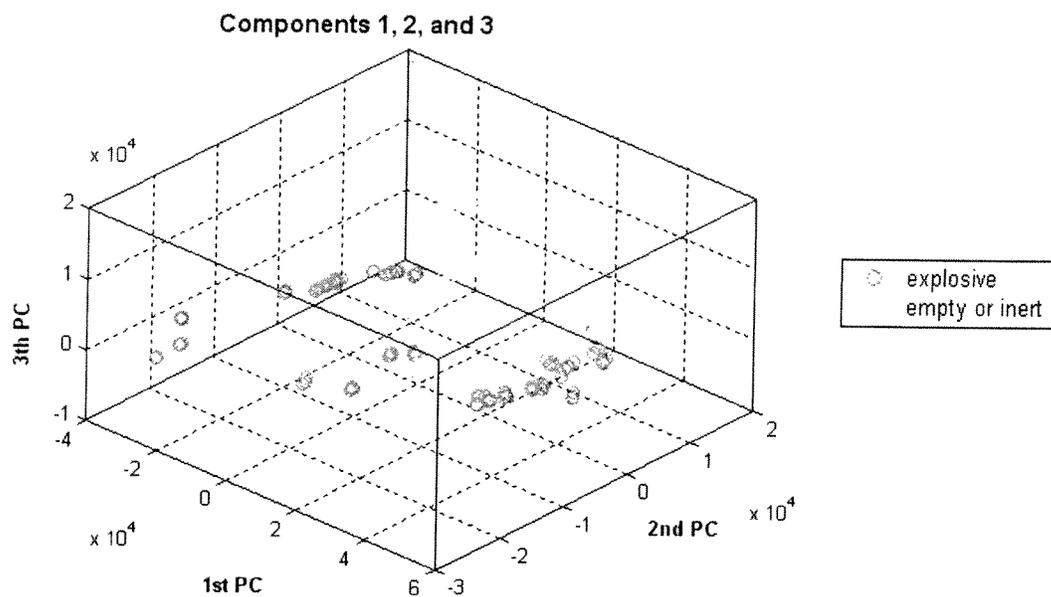


Figure 2.5.2 – 17: Graphical display of the first three principal components for large and medium shells on any background. The data is mean-centered.

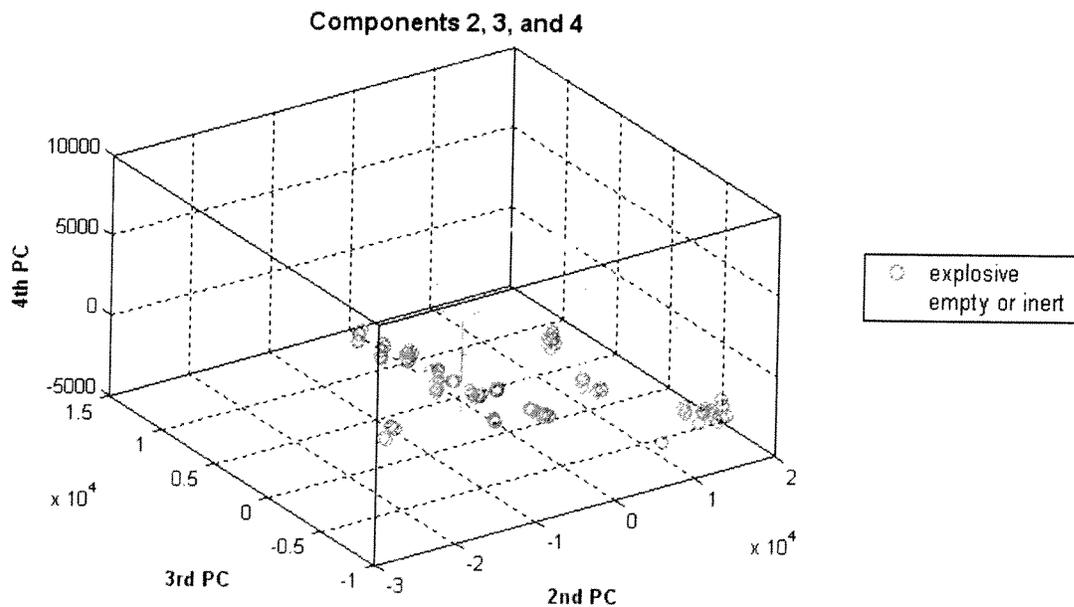


Figure 2.5.2 – 18: Graphical display of second, third and fourth principal components for large and medium shells on any background. The data is mean-centered.

The chart below shows that the fourth principal component is responsible for only 1.22 percent of the variance in the data. However, Figures 2.5.2 – 17 and 2.5.2 – 18 show that, at least visually, the component is important for cluster separation.

Percent Variance Captured by PCA Model

Principal Component Number	Eigenvalue of Cov(X)	% Variance Captured This PC	% Variance Captured Total
1	6.60e+008	83.50	83.50
2	8.48e+007	10.73	94.22
3	2.87e+007	3.64	97.86
4	9.61e+006	1.22	99.08
5	1.97e+006	0.25	99.33
6	1.09e+006	0.14	99.46

2.5.2.9 Further Work

Several areas in this study are worthy of further investigation:

- The explosive fill materials form individual clusters and investigation of the content of these clusters will provide additional understanding of the explosive-fill identification problem.
- Autoscaling with background subtraction provided excellent separation of explosive and empty/inert fill materials and warrants further investigation, especially for smaller sized shells, which were not successfully differentiated with the mean-centering technique.
- The distance between clusters can be quantified.
- The problem of predicting the fill material of a shell can be addressed using well-known techniques that are compatible to PCA, such as Soft Independent Modeling of Class Analogies (SIMCA).

2.5.2.10 Summary and Conclusions

PCA techniques are effective at classifying explosive and inert fill materials in large and medium sized shells ($\geq 90\text{mm}$) on a multitude of backgrounds for this data set. In addition, the sample space can be divided into two distinct regions, explosive and empty/inert.

It was determined that mean centering of data is an effective preprocessing technique and that background subtraction is not necessary for separating the explosive and inert fill materials for large and medium sized shells. This is an important result because the collection of background signal can be time-consuming in the field.

An accurate background signal may be necessary to apply PCA techniques to small shells since the small shells produce a low signal-to-noise ratio.

It appears that the data for large and medium shells form individual clusters according to background and shell size.

It appears that shells of all sizes with inert fill materials are indistinguishable from empty shells of the same background and size.

For this data set, three principal components are sufficient to separate the explosive and inert fill materials in sample space for large and medium sized shells.

2.6 Spectral Estimation

Fill material classification and identification performance are dependent on the quality of the measured spectra. This may also be true of methods that operate directly on the measured spectra as well as on methods that extract a set of features, such as the SPIDER Element Counts or principal component coefficients, from the measured spectra. The theoretical model for the measured spectra indicates that it should consist of spectral peaks corresponding to the constituent elements. The resolution of the measured spectra, however, may not be sufficient to resolve closely spaced peaks.

Frequency estimation methods were investigated to improve the resolution of the measured spectra. These methods are well suited to estimating spectra that contain sinusoidal components and, therefore, are appropriate for this application where the spectral response contains peaks due to individual elemental responses. It was anticipated that detection performance should improve if the resolution of the spectral peaks due to the elemental responses in the measured spectra can be improved. All of the approaches considered here are based on eigenanalysis of the autocorrelation matrix. Only a summary of the methods and results are presented in this section. Please see Section III of the Duke Final Report for more details.

The relationship between the power spectral density (PSD) and the autocorrelation function (ACF) of a wide-sense stationary (WSS) random process is a familiar one, the Fourier transform. [2] The PSD is simply the Fourier transform of the ACF, and similarly, the ACF is the inverse Fourier transform of the PSD. An important property of the ACF of a WSS process is that it is conjugate symmetric. In the special case of a real-valued random process, the ACF is a real-valued even function according to this property. Two important properties of the PSD are that it is real-valued because the ACF is conjugate symmetric, and it is non-negative. Again, a real-valued random process is a special case for which the PSD is an even function because the ACF is real and even.

The measured PELAN spectra satisfy the two properties of a PSD (it is real-valued and nonnegative) and, therefore, may be interpreted as a PSD for positive frequencies. Assuming the underlying random process is real-valued, so that the PSD is an even function, the corresponding ACF, $r(t)$, for each measurement is generated by reflecting the measured spectra about the $f = 0$ axis and then taking the real part of the inverse Fourier transform. The real part of the inverse Fourier transform is taken because the ACF of a real-valued random process is a real-valued

even function. Once the ACF has been determined, the autocorrelation matrix (ACM), R , can be generated.

Once the ACM corresponding to the measured PELAN data has been formed, standard eigenanalysis approaches may be applied to decompose the ACM into a set of orthogonal vectors, termed eigenvectors. Corresponding to each eigenvector is an eigenvalue, λ , which is a complex-valued scalar satisfying $\mathbf{R}\mathbf{v}=\lambda\mathbf{v}$ for the eigenvector \mathbf{v} .

The parametric spectral estimation, or more precisely, frequency estimation, techniques considered here are all based on the eigenanalysis of the total ACM, which is composed of two distinct ACMs, the signal ACM and the noise ACM. [7] The theory behind these eigenanalysis approaches is that the p principal eigenvectors of the total ACM, which are the same as the p principal eigenvectors of the signal ACM, may be used to extract the sinusoidal components of the signal. Once the principal eigenvectors have been obtained, they are transformed to the frequency domain, and the spectral (frequency) estimate is obtained by summing the frequency domain representations of the eigenvectors. It is important to note that these methods do not actually provide true spectral estimates, but rather, estimates of distinct frequencies present in the signal.

Several approaches were used that differ primarily in the criteria applied to select the eigenvectors to retain for the frequency estimation and how the retained eigenvectors are combined to form the spectral estimate. In addition, those that retain only a subset of the eigenvectors all share the common challenge of choosing the correct model order, or the correct number of eigenvectors, to retain for the frequency estimation. These approaches are

- Auto-Regressive (AR) Frequency Estimation
- Minimum Variance Frequency Estimation
- Bartlett Frequency Estimation
- Multiple Signal Classification (MUSIC)
- Eigenvector Frequency Analysis

The eigenanalysis spectral estimation algorithms were applied to the chemical data provided by SAIC. This data set contains six chemical compounds: ammonium nitrate (AN), bleach (BL), gasoline (GS), diesel fuel (DS), ammonia (AM), and water (WA).

Each of the eigenanalysis spectral estimation techniques assumes the model order, p , is known. The model order p determines the number of complex exponentials assumed to exist in the spectrum. Since a single sinusoid is the sum of two complex exponentials, the number of sinusoids assumed to exist in the spectrum is $p/2$. Selecting the model order has proven to be a very challenging task, particularly when a background measurement has been subtracted, and to-date, no method for selecting the model order has been found to be universally appropriate. Thus, model order selection for the eigenanalysis spectral estimation algorithms remains an area for continuing work.

Frequency analysis of the autocorrelation matrix eigenvectors using the chemical compounds has produced some promising initial results. Analysis was performed with and without background subtraction. See Sections 15-17 of the Duke Final Report for some sample results.

In summary, alternative spectral estimation techniques have the potential to improve fill material classification and identification performance by improving the quality of the measured spectra. Initial efforts focused on frequency estimation methods, which employ eigenanalysis of the autocorrelation matrix. Unfortunately, the question of model order selection hindered the spectral estimates and remains an area of open research. However, an alternative approach, also based on the eigenanalysis of the autocorrelation matrix, did produce promising results. The eigenvector frequency analysis differs from the spectral estimation methods in that, rather than summing the magnitude spectra of a selected subset of the eigenvectors, the pattern of the magnitude spectra for all eigenvectors is considered. This approach considers the order in which the frequencies appear in the eigenvectors but does not consider the relative contribution of each frequency as identification algorithms operating on the measured spectra would.

2.7. Processing of PELAN IV Data

2.7.1. Introduction

SPIDER element count (SEC) data were collected at Indian Head and at SAIC using the PELAN IV system from December 2003 to December 2004 and were processed to assess system performance for fill material classification (non-explosive versus explosive), where the non-explosive class included both empty shells and shells with an inert fill material. The data contains an extensive set of fill materials; however, for this study, the measurements for the chemical and miscellaneous fills are discarded and only the empty, inert, and explosive fill materials are retained. The explosive fill materials retained are TNT, RDX, HMX, and CompB, and the inert fill materials retained are cement, sand, plaster of Paris, and wax. The measured data has been parameterized according to general shell size (small, medium, and large) as well as background environment. Although there are four distinct background environments, (sand, asphalt, soil, and a metal table), the first three environments (sand, asphalt, and soil) were grouped together to create a common environment so, in practice, only two environment parameterizations are considered (common and metal table). The SECs were determined with an empty shell in the background measurement as well as without an empty shell in the background measurement. The data distribution, excluding the December 2004 data, for the data taken with an empty shell in the background is given in Table 2.7-1, and the data distribution, again excluding the December 2004 data, for the data taken without an empty shell in the background follows in Table 2.7-2. These tables show that although this is a rather large data set (546 measurements with an empty shell and 494 measurements without an empty shell), it does not contain enough data for all the subset parameterizations, such as for empty large shells measured with a common background, to reliably evaluate all the detection algorithms that have been developed. It will be shown in the following sections how the paucity of data affects evaluation of the detection algorithms.

With Empty Shell in Background

		Environment						All Environments		
		Common (Sand, Asphalt, Soil)			Metal Table					
Fill Material		EX	IN	EM	EX	IN	EM	EX	IN	EM
Shell Size	SM	47	43	11	42	74	29	89	117	40
		54		103		157				
	MD	41	27	10	47	18	20	88	45	30
		37		38		75				
	LG	55	18	3	39	12	10	94	30	13
		21		22		43				
All Shell Sizes		143	88	24	128	104	59	271	192	83
			112			163			275	

Table 2.7-1. Data distribution (546 total) for PELAN IV data (excluding December 2004 at Indian Head) taken with an empty shell in the background.

Without Empty Shell in Background

		Environment						All Environments		
		Common (Sand, Asphalt, Soil)			Metal Table					
Fill Material		EX	IN	EM	EX	IN	EM	EX	IN	EM
Shell Size	SM	42	38	11	32	78	34	74	116	45
		49		112		161				
	MD	36	23	10	39	13	25	75	36	35
		33		38		71				
	LG	51	18	3	22	12	7	73	30	10
		21		19		40				
All Shell Sizes		129	79	24	93	103	66	222	182	90
			103			169			272	

Table 2.7-2. Data distribution (494 total) for 2004 data taken without an empty shell in the background.

2.7.2. Analysis Algorithms

The baseline algorithm is the energy detector. This algorithm simply computes the energy in the SEC vector for each measurement,

$$E_s = \sum_{c=1}^C S_c^2$$

where S_c is the SEC value for element c and there is a total of C counts in the SEC vector.

The WKU decision tree results are also shown as a baseline performance measure. However, they are probably not reliable, since they were designed for a previous generation of the PELAN system and have not been optimized for this system.

The algorithms considered for fill material classification are all variants of a generalized GLRT. The GLRTs vary in the choice of model parameterization. The GLRTs considered here range from a simple GLRT, in which there is no model parameterization, to a more complex GLRT, in which the model is parameterized by both shell size and background environment.

The simplest GLRT under consideration is the one for which there is no model parameterization. This GLRT simply aggregates all the explosive data to determine a single set of statistics for that class (H_1) and the non-explosive data to determine a single set of statistics for that class (H_0). Assuming the data, \mathbf{x} , follows a Gaussian distribution, the closed form expression for the GLR after simplification is

$$\lambda(\mathbf{x}) = \frac{|\mathbf{C}_0|^{1/2}}{|\mathbf{C}_1|^{1/2}} e^{\frac{1}{2}[(\mathbf{x}-\mu_0)^T \mathbf{C}_0^{-1}(\mathbf{x}-\mu_0) - (\mathbf{x}-\mu_1)^T \mathbf{C}_1^{-1}(\mathbf{x}-\mu_1)]}$$

where μ_i is the mean of the distribution for hypothesis i , and \mathbf{C}_i is its covariance. For computational efficiency, constants are typically absorbed into the threshold, and the natural logarithm of the likelihood ratio is usually computed. Since the natural logarithm is a monotonic function, it does not alter the detection results. Consequently, the GLR is often expressed as

$$\lambda'(\mathbf{x}) = (\mathbf{x}-\mu_0)^T \mathbf{C}_0^{-1}(\mathbf{x}-\mu_0) - (\mathbf{x}-\mu_1)^T \mathbf{C}_1^{-1}(\mathbf{x}-\mu_1)$$

This form of the GLR assumes there are no uncertain parameters, other than the underlying hypothesis, on which the signal depends. The signal model utilized in the GLR (or LR) can be made as general or precise as desired. In general, the trade-offs considered are computational complexity, model accuracy, and ability to estimate the necessary parameters or otherwise train the algorithm. Using a more precise model usually requires more extensive training data and increases the complexity of the GLR computation, but in return, it can provide improved performance. Usually, studies are conducted to determine how performance depends on the signal model so that the simplest model with the best performance can be chosen for implementation.

Considering the dependence of the data on the signal parameters when forming the GLR often improves performance, even when the values taken by the parameters are not known a priori. However, if some, or all, of them are known a priori, performance often improves even more, sometimes dramatically. The degree of improvement depends on the characteristics of the data, and how strongly each parameter influences the data.

In this study, the signal model parameters chosen for investigation are the fill material, the target size, and the background environment. Thus, the measured signal, $s(\phi, \tau, \zeta)$, is a function of the fill material, ϕ , the target size, τ , and the background environment, ζ . Expressions of $\lambda(\mathbf{x})$ as a function of these parameters are found in the Duke Final Report.

2.7.3. Performance Results

The performance of the GLRTs on the PELAN IV data, excluding the December 2004 data, is compared to the baseline energy detector and WKU decision tree performance. The GLRTs implemented for this study are

- No model parameters
- Size known
- Size unknown
- Background known
- Background unknown
- Size unknown and background unknown
- Size unknown and background known
- Size known and background unknown
- Size known and background known

In addition, each of the above GLRTs was implemented for three different assumptions regarding the fill material for the null hypothesis, H_0 :

- Inert and empty fills aggregated to determine H_0 statistics ($M = 1$)
- Inert and empty fills kept separate to determine two groups of H_0 statistics ($M = 2$)
- Only inert fills are utilized to determine the H_0 statistics ($M = 1$)

In all of the GLRTs, all explosive fills were aggregated to determine H_1 statistics ($M = 1$).

The ROCs were generated by taking the average of 100 independent training/testing set realizations in which 90% of the data was utilized for training and the remaining 10% was utilized for testing. No effort was made to assure the training and testing sets were evenly matched. For instance, the training set consisted of 90% of the total data, not the combination of 90% of the explosive fill material data, 90% of the inert fill material data, and 90% of the empty (no fill material) data. Finally, performance was assessed for 0%, 15%, and 25% "Don't Know" in order to evaluate the impact of the "Don't Know" declaration on the overall performance level.

ROC curves generated as part of this analysis are provided in the Duke Final Report. Only a selection of the ROC curves is shown here for aiding the discussion. Each figure shows a comparison of ROCs for a given “Don’t Know” level. The comparisons show the performance for each of the three assumptions regarding H_0 as well as the effects of excluding the cross-correlation from the covariance matrix. It is important to note the scale on the axes when comparing ROCs since many of them have limits of 0.5 to 1, rather than the conventional limits of 0 to 1, in order to more clearly show the different curves.

Figure 2.7-1 shows performance when an empty shell is included in the background measurement. For these curves and those in the Duke report, the data includes all three shell sizes. The ROCs indicate that, with sufficient training data, the GLRTs outperform the baseline energy detector, and generally there is not much performance difference between grouping the empty and inert fills together or integrating over the H_0 fill material. They also show that for all three parameterizations, the GLRTs with unknown parameters often perform better than the same GLRTs with the parameters known. This result is an artifact of insufficient training data for the large shells because the statistics cannot be reliably estimated. Neglecting the correlation generally degrades performance, but sometimes it corrects the problem of the parameter-uncertain GLRT, outperforming the parameter-known GLRT. This occurs because it is easier to estimate variances than the full covariance matrix.

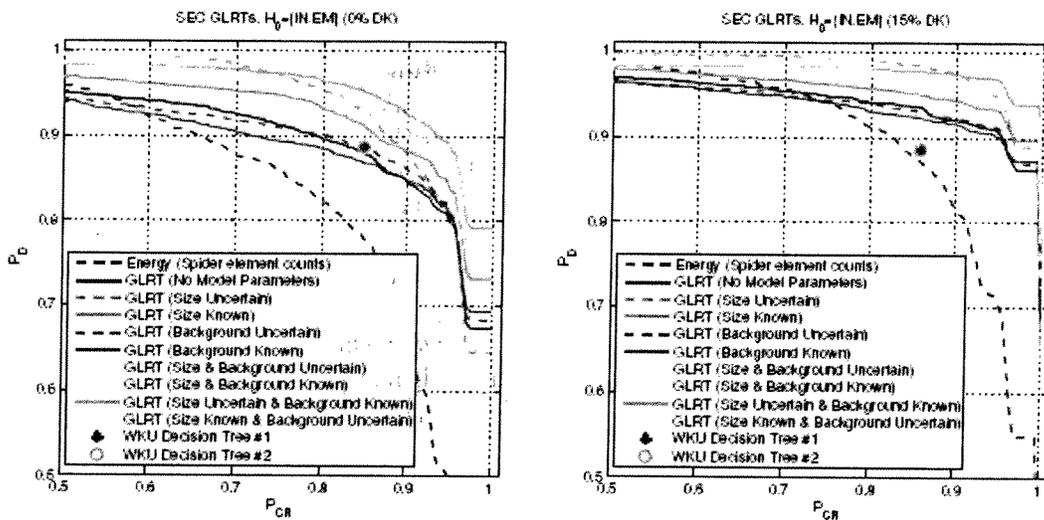


Figure 2.7-1. All shell sizes, with correlation, 0% Don’t Know (left) and 15% Don’t Know (right), with an empty shell in the background run.

To test the hypothesis that there is insufficient large shell data to adequately train the GLRTs, the performance is also evaluated for only the small and medium shells. These ROCs are represented in Figure 2.7-2. Once the large shells are excluded from the data set, the size-known GLRT outperforms the size-unknown GLRT, which supports the hypothesis that the previous results showing the size-uncertain GLRT outperforming the size-known GLRT are a result of insufficient training data for the large shells. However, the other parameter-uncertain GLRTs, which include the background environment in their parameterizations, still often outperform their respective parameter-known GLRTs. This is likely an unreliable result, due to inadequate training data. There is consistency with the previous set of results in that generally there is not

much performance difference between grouping the empty and inert fills together or integrating over the H0 fill material, and neglecting the correlation generally degrades performance, but sometimes it corrects the problem of the parameter-uncertain GLRT, outperforming the parameter-known GLRT because it is easier to estimate variances than the full covariance matrix.

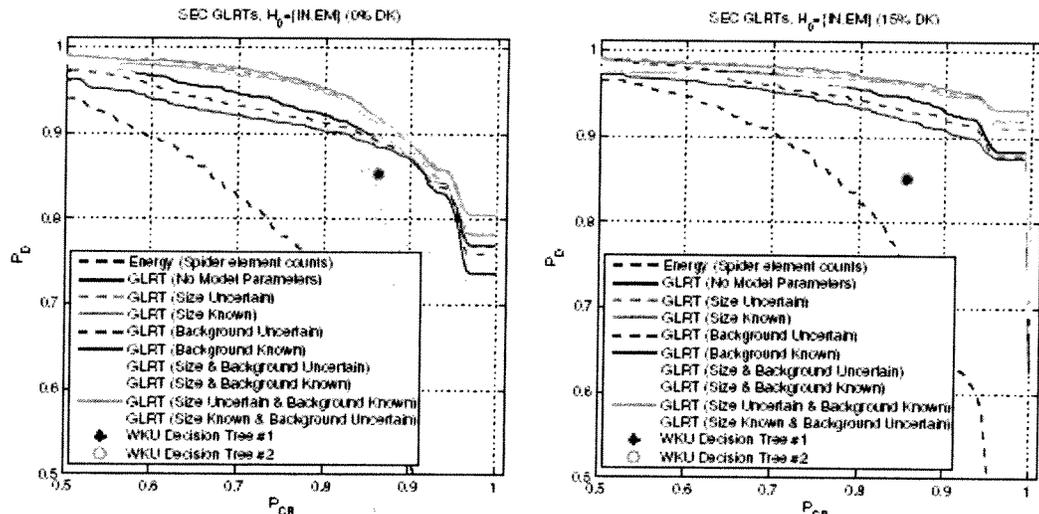


Figure 2.7-2. Small and medium size shells, with correlation, 0% Don't Know (left) and 15% Don't Know (right), with an empty shell in the background run.

ROCs are also generated for small shells and medium shells as represented in Figure 2.7-3. The performance for medium shells is better than for small shells. This is due to the increased volume of fill material in the larger shell. In addition, the performance for medium shells generally exhibits very good performance. These ROCs must be viewed with caution, however, since they were generated with randomly selected training data sets which may not accurately reflect the characteristics of the testing data set due the small amount of available data.

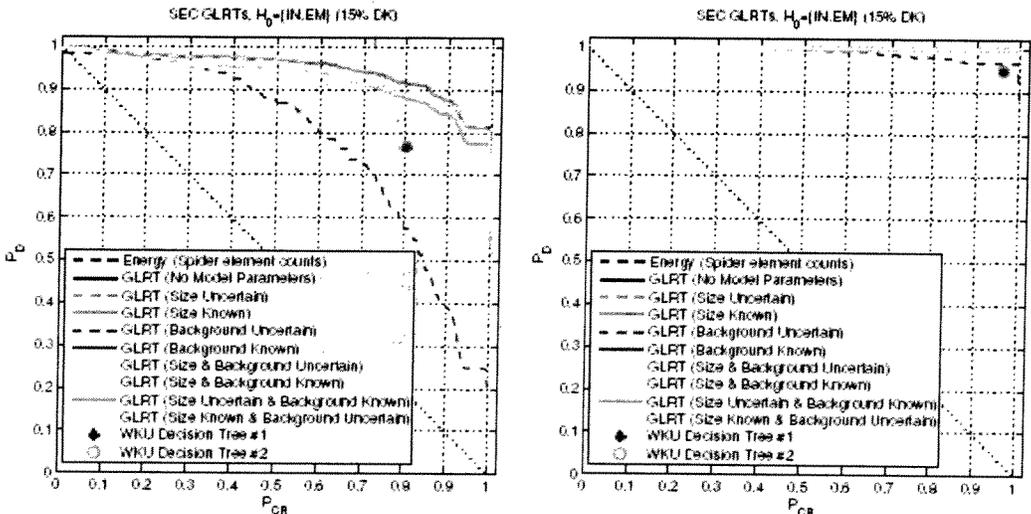


Figure 2.7-3. Small (left) and medium (right) shells, with correlation, each 15% Don't Know, with an empty shell in the background run.

Finally, the data without an empty shell in the background measurement is considered. Performance using the full data set, with all three shell sizes, could not be evaluated due to insufficient training data; however, once the large shells were removed, it could be evaluated, and the ROCs are represented in Figure 2.7-4. Some general trends are the same as the previous sets of results. The GLRTs are better than the energy detector, provided sufficient training data is available, and neglecting correlation generally degrades performance. One difference is that the parameter-known GLRTs outperform the parameter-unknown GLRTs. This could be due to sufficient training data now being available, or the Gaussian assumption is now more appropriate than it was for the other two cases. Compared to the data for which an empty shell was included in the background measurement, the performance is degraded for no model parameterization, size parameterization, and background parameterization. For size and background parameterization, however, the performance without an empty shell in the background is comparable to the performance with an empty shell in the background. One additional difference is that performance is slightly better for the GLRT which integrates over the uncertainty in the H0 fill material ($H_0 = \{IN \text{ or } EM\}$) than the GLRT which aggregates the H0 fill materials into a single group ($H_0 = \{IN, EM\}$).

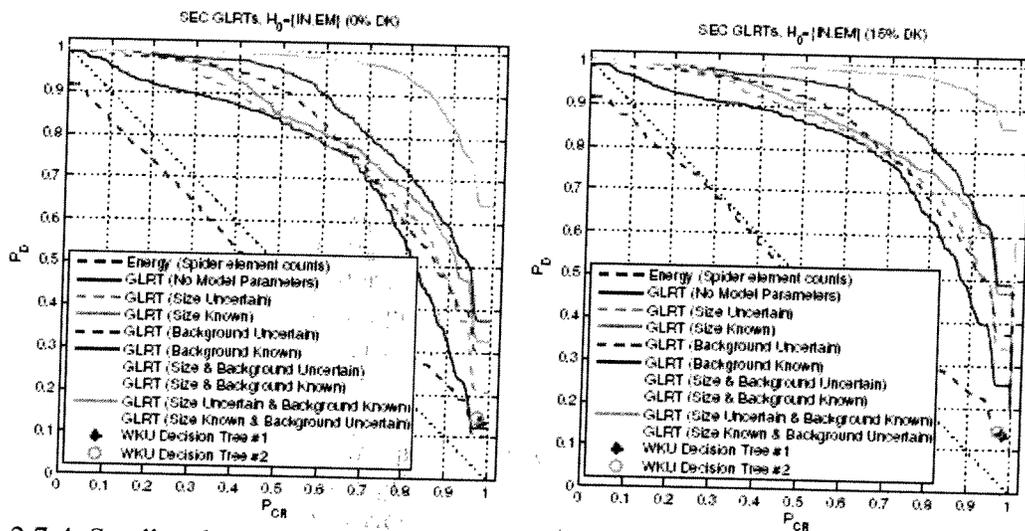


Figure 2.7-4. Small and medium shells together, with correlation, 0% Don't Know (left), 15% Don't Know (right), with NO empty shell in the background run.

The small and medium shell data without an empty shell in the background were also evaluated individually. The ROCs for the small and medium shells separately are represented in Figure 2.7-5. Again, these ROCs should be viewed with caution since the amount of available data is rather small and the training data sets were selected randomly.

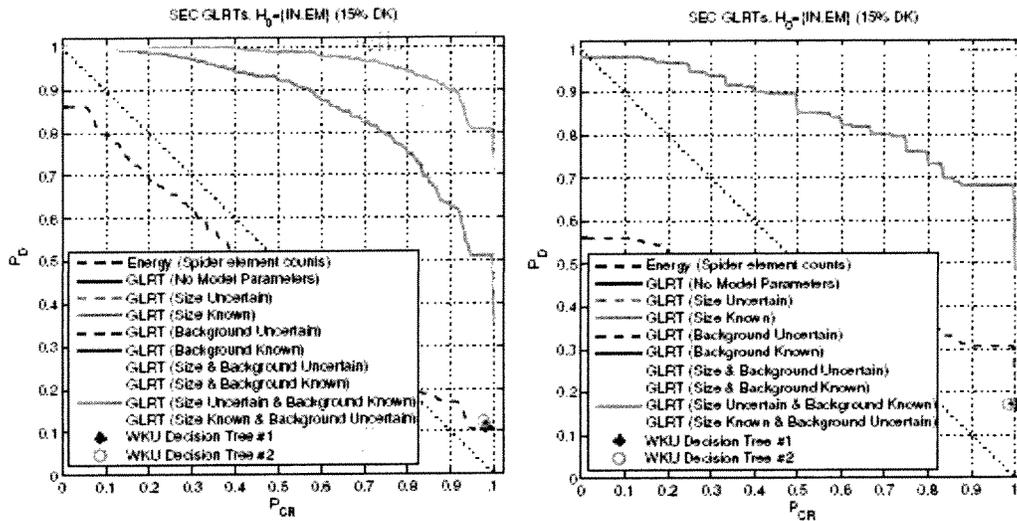


Figure 2.7-5. Small (left) and medium (right) shells analyzed separately, with correlation, 15% Don't Know, with NO empty shell in the background run.

To mitigate some of the problems associated with using random training sets with limited data – namely, not being guaranteed to have both H_0 and H_1 data in the testing sets – matched training sets are also evaluated. For these results, the training sets are composed of 90% of the H_0 data and 90% of the H_1 data. Otherwise, the performance analysis is the same as for the random training sets.

ROCs were generated for small shells with an empty shell in the background measurement and for medium shells. Also, ROCs are shown with results for small and medium shells without an empty shell in the background. The performance trends seen earlier for the random training sets continue here. Performance generally improves as more parameters are included in the model and generally improves when the parameters are known, rather than uncertain. Performance generally improves when correlation is included, though it may degrade if there is not enough data to estimate the correlations. Integrating over the H_0 hypotheses (empty and inert) also slightly improves performance.

Overall, the performance is promising. As expected, the medium shells show better performance than the small shells, and the data with the empty shell in the background measurement shows better performance than the data without the empty shell in the background measurement.

2.7.4. Performance Results of December 2004 Data

The December 2004 SEC data has been analyzed using GLRTs. The GLRTs were trained using all PELAN IV data with the empty shell in the background measurement analyzed in the preceding section, and then applied to the December 2004 data. The December 2004 data consists of 70 measurements, 61 of which have an empty shell in the background measurement. The remaining nine measurements do not have an empty shell in the background measurement.

The data were analyzed in two ways. First, all the PELAN IV data is retained, including the December 2004 data. In this case, there potentially is a mismatch between the training data and the nine test measurements that do not have an empty shell in the background. For this reason, only the measurements that include an empty shell in the background measurement are also evaluated. Some of the results for all the data and for only data with an empty in the background are shown in Figure 2.7-6. The data without an empty shell in the background could not be analyzed independently because the only fill material for which measurements were taken was "Empty." Therefore, ROCs cannot be generated, because there is only H_0 data and no H_1 data.

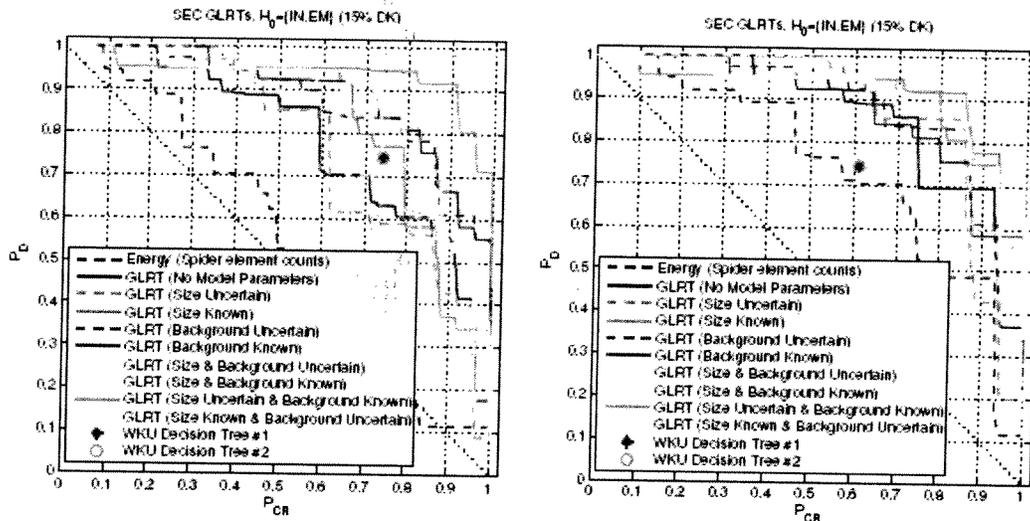


Figure 2.7-6. All December 2004 data (left) and December 2004 data with an empty shell in the background (right), with correlation, 15% Don't Know

There does not seem to be a significant difference between performance for all the data and performance for only the data with an empty shell in the background. In addition, the trends seen previously follow through here. Increased model specificity generally improves performance, as does knowledge of the parameter values.

2.7.5. Comparison with Neural Net Analysis

In a parallel effort supported by SERDP under a separate contract, NAVEODTECHDIV investigated the application of neural nets to SEC for target classification (explosive versus inert). The same SEC data sets used in the GLRT analysis described above were provided to NAVEODTECHDIV, and the same analysis approach was used to generate ROC curves. Several ROCs were provided to SAIC for comparison with the neural net results. More details of the neural net investigation can be found in NAVEODTECHDIV's final report (April 2005).

The baseline neural network parameters used in the analysis are as follows:

- Three-layer neural network

- Four inputs (unless otherwise specified), the elemental counts of carbon, hydrogen, nitrogen and oxygen. In certain cases, shell size and background were added as input variables.
- Two internal neurons in hidden layer
- One output neuron
- Transfer function sequence of tansig, tansig, logsig

As in the GLRT analysis, the neural network randomly selected 90% training data and the remaining 10% was used for testing. Therefore, no testing data was used for training. No “Don’t Knows” were used in the training. The background (environment) and shell size were not used as inputs in this training. Training and testing occurred 100 times, and the false positives and probabilities of detection were stored in memory and were averaged.

At the time the analysis with GLRT was performed on the SEC for comparison with the neural network (NN) results, the December 2004 data was not yet available. The results with GLRT here are done without the 70 runs from December 2004, while the NN approach included these data. Because the December 2004 data is close to only 10% of the total runs used in the training, the differences in the ROC results without December 2004 data are minor.

This network is the “control” neural network to which all other networks are compared, and it is represented in black on the ROC curve graphs provided by NAVEODTECHDIV. Two test groups were provided to SAIC for comparison with GLRT: the first, in which size and background are added as input variables (in addition to C, H, N, and O) and the second, in which the three differently sized shell groups (small, medium, large) are trained and tested separately.

Test Group 1: Addition of shell size and background as input variables

Initially, the number of inputs to the neural network were four (i.e., just the elemental counts from C, H, N, and O). During this testing, the number of inputs was increased to five when shell size was added as an input variable and again when background was added as an input variable. The number of inputs was increased to six when both shell size and background were added as input variables. This made for a total of three tests. The NN results are shown in Figures 2.7.5-1 and 2.7.5-2, where the black curve in each figure is the control result (shell size and background not included as input variables) and the colored curves are with an added input for shell size, background, and shell size plus background.

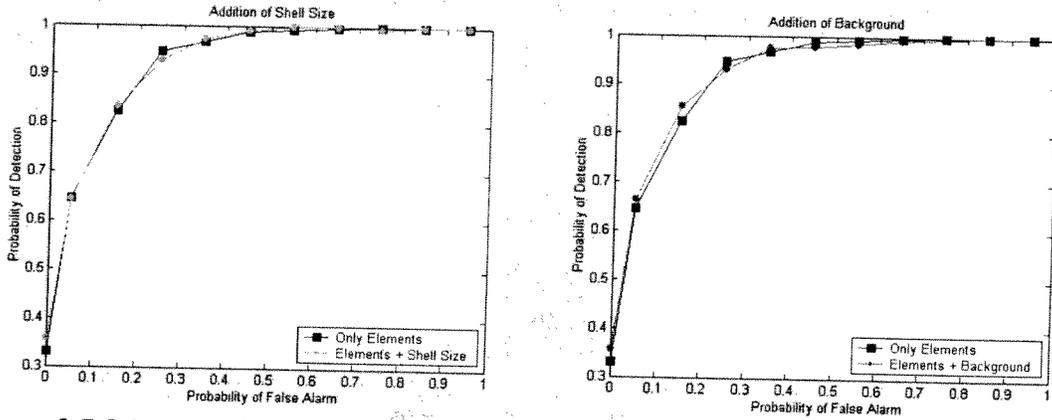


Figure 2.7.5-1. Neural net results of all PELAN IV data with shell size as input (left plot) and with background as input (right plot). The black curve is the baseline curve as described in the text.

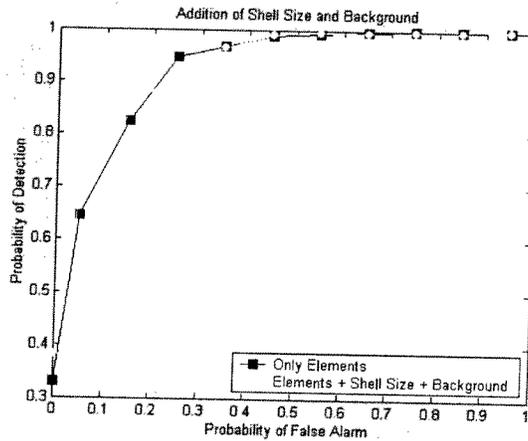


Figure 2.7.5-2. Neural net results of all PELAN IV data with baseline curve (no inputs) and with shells size and background as inputs.

In Figures 2.7.5-3 and 2.7.5-4, we show the NN results plotted along with the GLRT results for training, with no separation in shell size or background, with shell size known only, with background known only, and with both shell size and background known. The black curves in each plot are the NN results using the same data set and training parameters.

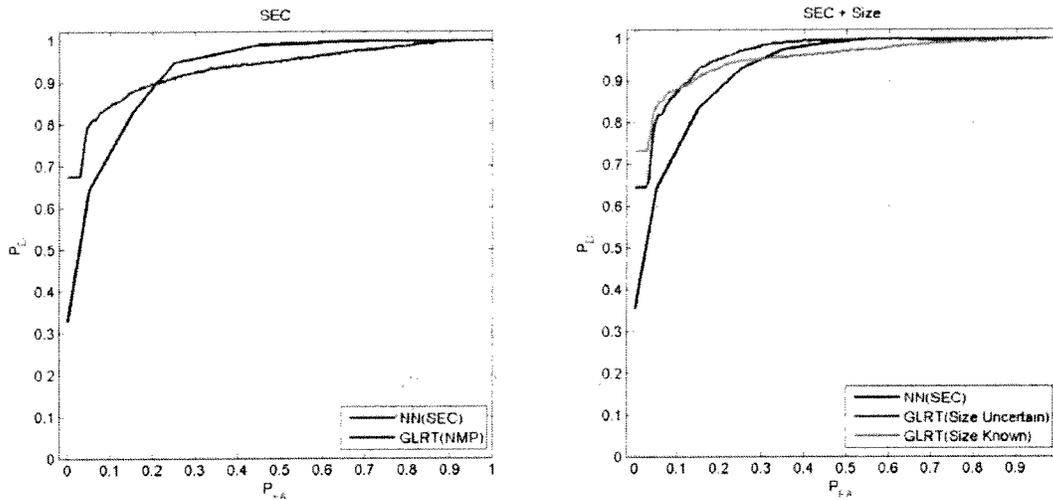


Figure 2.7.5-3. GLRT results of all PELAN IV data with only SEC input (left plot) and with size as input (right plot). The black curve is for the equivalent analysis using the neural net.

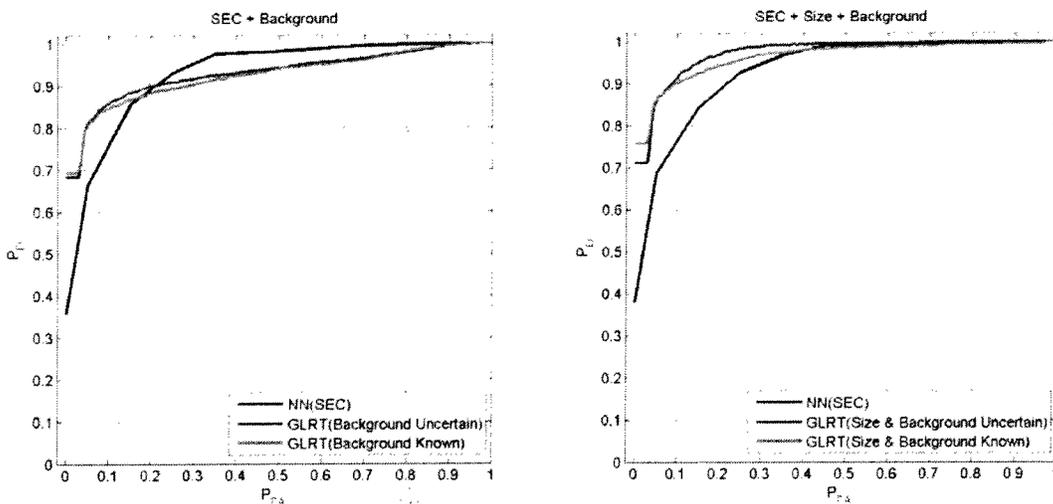


Figure 2.7.5-4. GLRT results of all PELAN IV data with background as input (left plot) and with size and background as input (right plot). The black curve is for the equivalent analysis using the neural net.

Test Group 2: Separation of Data by Shell Size

All the data was separated into three groups by its shell size: small, medium, and large. Each group, separately, was used to train the neural network using the procedure outline above and the results were plotted graphed. Only the elemental counts from C, H, N, and O were used as inputs. Figure 2.7.5-5 shows the results of each shell size group compared to the result when all data is grouped together.

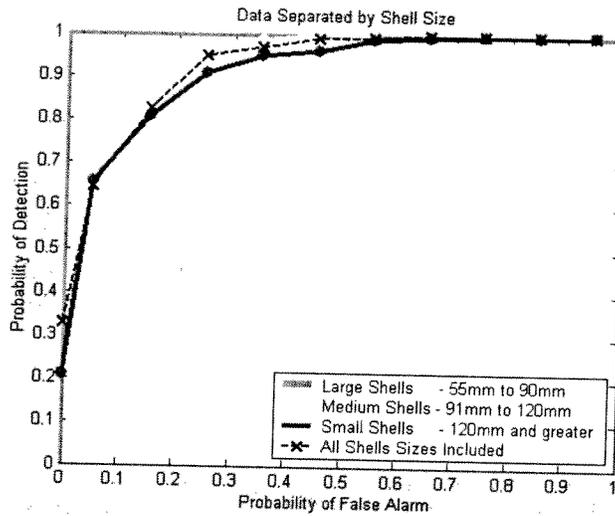


Figure 2.7.5-5. Neural network results of all PELAN IV data for each shell size group after training on each separately.

The plots in Figure 2.7.5-6 compare the results of the NN (black curves) and the GLRT (blue curves) for the small shells and the medium shells, separated in the training. The plot for the large shell with GLRT is the same as that for the NN result.

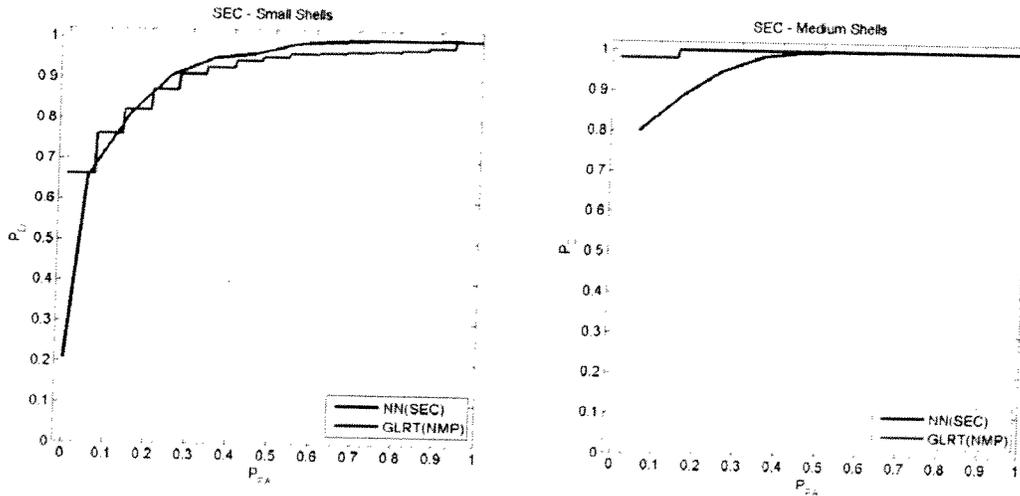


Figure 2.7.5-6. GLRT results of all PELAN IV data for small and medium sized shell size group after training on each separately. The black curves are the equivalent analysis using neural networks.

In general, the NN and GLRT give similar results, but most often, GLRT reaches higher detection probabilities at lower false alarms rates sooner than does the NN. The results of the NN and GLRT for the small shells, where training is done for each group of shell sizes, are very similar to each other, possibly due to fewer distinguishing features when the signal-to-noise ratio is lower.

2.7.6. Summary

The performance of GLRTs using SECs on the 2004 data is shown for a variety of model parameterizations, several different data subsets, and three different "Don't Know" percentages. Generally, more precise model parameterization improves performance, provided that the statistics can be reliably estimated. For this reason, performance without correlation is sometimes better than with correlation. This is contrary to the expected result of performance with correlation being better than performance without correlation. As expected, performance improves as the shell size increases, due to the increased fill material volume. In addition, the effect of including an empty shell in the background measurement is evaluated by considering data measured with and without an empty shell in the background measurement. The neural network analysis performed by NAVEODTECHDIV was compared to the GLRT analysis with SECs and showed similar performance, though in general, GLRT performed better. Generally, performance is better when an empty shell is included in the background measurement.

2.8. Implementation in PELAN IV

Because several algorithms developed with the support of SERDP demonstrated that they met the goals of improving performance, increasing robustness, and easing the trainability for multiple targets and library updates, they were implemented into the PELAN IV systems that were delivered to the Navy as part of the NFI 6.4 program. SAIC implemented the GLRT decision maker with the tertiary declaration and the entropy-based confidence level into PELAN IV systems. The SPIDER spectral analysis is still used in the system.

A test report with performance results was delivered to the Navy and, with their approval, can be provided to SERDP. The Navy conducted tests in December 2004 using these units. As of this writing, no results have been presented.

3. CONCLUSIONS

Several analysis algorithms were investigated in the SERDP contract for improving the performance of PELAN IV, increasing the robustness of the decision maker and allowing for easier training and parameter upgrades.

SAIC and Duke University collaborated on the development of advanced algorithms. During this project, several key results were obtained and are summarized below.

- GLRT was established as an effective decision-making algorithm.
 - GLRT can be used in conjunction with Least Squares (SPIDER), PCA, and other spectral analysis techniques (e.g., MUSIC).
- The tertiary declaration was added for GLRT decision making (“Don’t Know”) for explosives/inert-filled shells.
 - Improved performance occurred with addition of “Don’t Know.”
- PCA can perform better than Least Squares on shell targets.
- Background measurements may not be necessary for effective PCA analysis.
 - The need for empty shell in background is eliminated.
 - Less user input is required for recording the environment.
 - Further testing is required for verification.
- The GLRT, tertiary declaration, and entropy-based confidence level were implemented into PELAN NFI systems.
- Results using GLRT on elemental counts compared similarly in performance with the neural network results provided by NAVEODTECHDIV.
- Data collection at Indian Head was conducted December 6-22, 2004.
 - SERDP data was made available to SAIC and Duke University in February 2005.
 - Using prior training parameters, the performance results of December 2004 data were consistent with previous results.

For different sensor geometries or gamma-ray detector types (such as LaBr3), the data collected with PELAN IV cannot be directly used in the target library. However, the methods developed here using PELAN IV data can be applied to systems with these different configurations.

4. TRANSITION PLAN AND RECOMMENDATIONS

SAIC, in collaboration with Duke University and Environmental Chemical Corporation (ECC), have been selected by ESTCP to build, test, demonstrate and validate a mobile, multi-detector-based PELAN unit for the classification of UXO filler at cleanup sites. With improved classification algorithms, we can improve the reliability of the target analysis, improve the performance, and thus, provide a cheaper and safer means for environmental remediation needs and other EOD-related efforts. The improved spectral analysis and decision-making algorithms developed in this project will be implemented directly into the current PELAN systems. Along these lines, we have already implemented and started testing the tertiary GLRT and entropy-based confidence algorithms in the PELAN IV system.

Through the ESTCP project, we recommend the following steps for transitioning this technology to field applications.

- Build and test a lab system.
 - Conduct tests using modular system for determining optimal signal-to-noise
 - Using data collected in the lab, evaluate LS/GLRT and PCA/GLRT algorithms
 - Test system at Indian Head on live shells
 - Evaluate performance against algorithms
- Build and test a prototype system.
 - Build a fieldable PELAN system using multiple detectors
 - Work with ECC, a UXO cleanup contractor, for evaluating PELAN
 - Test PELAN at an ECC-supported site, such as Massachusetts Military Reserve, based on accepted criteria
 - Evaluate performance of system against selected algorithms
- Establish conditions and scale of cost benefits.
 - Use results of validation testing at the selected site to verify cost benefits and best mode of operation
 - Transition to commercial phase

Using a multiple detector system with collimation of the neutron beam and/or of the gamma-ray detector, the inspection time can be reduced and the signal-to-noise ratio increased so that detection performance of UXO targets is increased. Increasing the neutron output will also reduce the inspection time. This approach will also improve the detection rate and reduce the false alarm rate for the smaller shells (<90mm) where performance is compromised because of the smaller return signal.

Furthermore, with a multi-detector system, off-axis/off-angle detectors can be used to measure the background at the same time that target spectra are acquired. Both combinations of PCA with GLRT and Least Squares analysis (SPIDER) with GLRT have shown very good performance without the need for an empty shell in the background or, especially in the case with PCA, no need for a background spectrum at all. The algorithms examined in this project will apply

directly to the multi-detector-based system. Efforts will be carried out to determine the best way to combine the spectra before an analysis is made for the decision. Our intent is to explore further both of these approaches, and others described in this report, in the ESTCP contract and use tests to validate their robustness, trainability, and performance.

Data collection is an important part of determining an optimal system design and the most effective analysis algorithms for meeting the site remediation requirements. For small design changes, the system response changes little or can be corrected so that previous training data sets can be applied. Major changes in the sensor design, such as going with a high-resolution detector such as LaBr3, will require additional training data. The algorithms developed here can still be applied to, for example, a LaBr3-based system. We recommend investigating techniques for data transformation to map from one design change to another for preserving the performance while maintaining as much of the previous data as possible.

APPENDIX A: Test Plan

Test Plan for Data Collection at Indian Head in December 2004

Below is the test plan submitted to Kurt Hacker and Denise (Forsht) Lee on November 8, 2004. The purpose of the data collection was to support the algorithm development in the SERDP project. This plan was incorporated into one written by NAVEODTECHDIV, which can provide SERDP a copy if needed. The data was collected during the period December 6 through 22, 2004.

The issues we want to address with this testing are:

1. Study the effects of variations in target-to-detector distance and filler size and to evaluate methods to correct for these effects. Reducing this affect would also allow the training on smaller, more readily available shells, to be used for identification of larger shell fills.
2. Study and correct for changes on H signal (or any other thermal capture gamma ray) due to variations in the moisture content of the soil (or changes in neutron thermalization caused by the presence of a nearby wall).
3. Evaluate the effect on PCA results due to variation in background environment (especially dry vs. wet soil).
4. Investigate methods to eliminate the need for using an empty shell in the background measurement.
5. Improve the tertiary explosives identifier by separating the particular types of inert and explosive fills. Develop GLRT parameters separately for the separate fill clusters.
6. Acquire additional data on new targets for addition to the library.

There is overlap in these issues, and much of the PELAN 4 data on shells is being used to address some of these issues. More data is always useful because there are usually areas where too few measurements prevent adequate training. The most useful data would be on dry vs. wet environments and distance/size measurements.

Aside from the usual equipment for most PELAN tests conducted thus far, we'll need the following for these tests:

1. Shells sizes from 60mm to 155mm and larger.
2. Both explosives and inert fillers.
3. Soil, wet soil, and sand filled test beds. Ground environment is good too, but need a controlled environment.
4. Metal table for above ground measurements.
5. Moisture meter (provided by SAIC).
6. Photographer for taking photos of the various setups.

Data reporting:

At the end of each day, all Extensible Markup Language (XML) files and descriptions of the runs should be sent to SAIC. Have photos taken of representative setups.

Test procedures:

1. Each target and each background measurement will be 5 minutes long (neutrons on). If it's known that the fill contains no phosphorus (found in WP and nerve agent), then the activation can be turned off so the runs never go more than 5 min.
2. For each particular environment (i.e., sand, dry, wet soil, or metal table), record a background run without any empty shell present.
3. For each shell size, record a background with an empty shell present on the particular environment. If no empty shell is available, such as for a 500 lb bomb, use an empty shell with a similar thickness. If several distance measurements on a shell will be made at, say, 2" to 10", take the background measurement at the average distance (may even consider taking two, one closer, one further away).
4. For each target and environment, take 5 (10 preferred) target runs at different positions. Positions should vary side-to-side and front to back over a few inches, depending on the size of the shell. For small shells, like 60mm, a distance spread of 2" is sufficient. For 500 lb bombs, a distance spread of 10" is reasonable.
5. For the smaller shells (155mm and lower), take all measurements on a dry soil test bed first. Then wet the soil, let it set for a good hour, and continue with the measurements. (Measurements could be done on sand in the mean time.) Periodically (every few runs) record the moisture level. Try to maintain constant moisture level.
6. Periodically record the moisture level for each environment tested.

