

AD-A279 479



2

PL-TR-93-2221

**APPLICATIONS OF GENERALIZED LIKELIHOOD
RATIO TESTS TO SEISMIC EVENT IDENTIFICATION**

**Mark D. Fisk
Henry L. Gray
Gary D. McCartor**

**Mission Research Corporation
735 State Street
P.O. Drawer 719
Santa Barbara, CA 93102**

**DTIC
ELECTE
APR 07 1994
S E D**

23 September 1993

**Final Report
25 May 1990 - 25 August 1993**

copy **94-10526**


Approved for public release; distribution unlimited



**PHILLIPS LABORATORY
Directorate of Geophysics
AIR FORCE MATERIEL COMMAND
HANSCOM AIR FORCE BASE, MA 01731-3010**

DTIC QUALITY INSPECTED 3

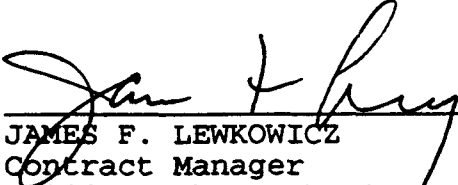
94-4-6 063

SPONSORED BY
Advanced Research Projects Agency
Nuclear Monitoring Research Office, and
Defense Sciences Office
ARPA ORDER NO 5307

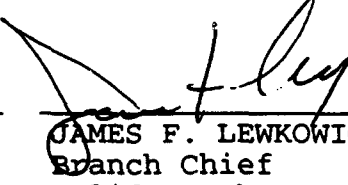
MONITORED BY
Phillips Laboratory
CONTRACT NO. F19628-90-C-0135

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Advanced Research Projects Agency or the U.S. Government.


This technical report has been reviewed and is approved for publication.



JAMES F. LEWKOWICZ
Contract Manager
Solid Earth Geophysics Branch
Earth Sciences Division



JAMES F. LEWKOWICZ
Branch Chief
Solid Earth Geophysics Branch
Earth Sciences Division



DONALD H. ECKHARDT, Director
Earth Sciences Division

This report has been reviewed by the ESD Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS).

Qualified requestors may obtain additional copies from the Defense Technical Information Center (DTIC). All others should apply to the National Technical Information Service (NTIS).

If your address has changed, or if you wish to be removed from the mailing list, or if the addressee is no longer employed by your organization, please notify PL/TSI, Hanscom AFB, MA 01731-5000. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704.0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)		2. REPORT DATE 23 September 1993	3. REPORT TYPE AND DATES COVERED Final Report 25 May 1990 - 25 August 1993	
4. TITLE AND SUBTITLE APPLICATIONS OF GENERALIZED LIKELIHOOD RATIO TESTS TO SEISMIC EVENT IDENTIFICATION			5. FUNDING NUMBERS PE 62714E PR 9A10 TA DA WU AC Contract F19628-90-C-0135	
6. AUTHOR(s) Mark D. Fisk Henry L. Gray* Gary McCartor*				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Mission Research Corporation P. O. Drawer 719 Santa Barbara, CA 93102-0719			8. PERFORMING ORGANIZATION REPORT NUMBER MRC-R-1444	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Phillips Laboratory 29 Randolph Road Hanscom Air Force Base, MA 01731-3010 Contract Manager: James Lewkowicz/GPEH			10. SPONSORING/MONITORING AGENCY REPORT NUMBER PL-TR-93-2221	

11. SUPPLEMENTARY NOTES

* Southern Methodist University, Dallas, TX

12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; Distribution unlimited.	12b. DISTRIBUTION CODE
---	------------------------

13. ABSTRACT (Maximum 200 words)
A summary of the work performed under this project is provided. It contains brief descriptions of statistical methods developed and data analysis performed during the early stages of this contract to improve yield estimation. Preliminary results of a study of discriminant transportability for Eurasia are also summarized. The main body of the report, however, focuses on development of statistical methods and their applications to seismic event identification problems relevant to monitoring a CTBT or NPT.

Classification and outlier detection methods, developed recently at MRC and SMU, are applied to the problem of seismic event identification, particularly for regional events. The methods, based on the generalized likelihood ratio (GLR), have considerable flexibility to rigorously treat a mixture of continuous and discrete features, normal or non-normal statistics, equal or unequal covariance matrices, and different distribution functions for different event types. Statistical procedures are also presented to test the validity of appropriate assumptions and, if necessary, transform the data to simplify implementation of the outlier and classification methods and to ensure that the necessary assumptions hold. Methods are also presented and applied to select the optimal subset of

14. SUBJECT TERMS Seismic Discrimination Outlier Detection Novaya Zemlya		Generalized Likelihood Ratio Classification Regional Arrays		15. NUMBER OF PAGES 96
				16. PRICE CODE
17. Security CLASSIFICATION OF REPORT UNCLASSIFIED	18. Security CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. Security CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT SAR	

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

CLASSIFIED BY:

DECLASSIFY ON:

discriminants and to treat missing data values robustly. This collection of methods provides a robust statistical framework to identify seismic events and to estimate classification probabilities rigorously under extremely general conditions.

These methods are applied to monitoring applications using data sets based on ISEIS feature data provided by Baumgardt (1992a, 1993a), for earthquakes in the Vogtland and Steigen regions, mining blasts in the Vogtland and Kola Peninsula regions, and nuclear explosions at Novaya Zemlya, recorded at combinations of the ARCESS, GERESS and NORESS arrays. The utility of the GLR outlier test for monitoring a new region with limited training data is demonstrated. A study is also presented in which training sets are contaminated intentionally with events of other types in proportions of 10-20% to determine if the GLR outlier test can detect these events and to assess the effect they can have on subsequent error rates if they are inadvertently included in training sets. Results also show that the GLR classification method leads to reduced error rates if training data are available for all relevant event classes. Event identification results are presented for the Novaya Zemlya event of 31 December 1992 and our conclusions and recommendations are discussed.

Accession For	
NTIS CRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and/or Special
A-1	

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

TABLE OF CONTENTS

List of Illustrations	v
List of Tables	ix
Executive Summary	x
Yield Determination	x
A Constrained Bayesian Approach for Testing TTBT Compliance	x
Multivariate Seismic Calibration for the Novaya Zemlya Test Site	xii
Seismic Event Identification	xiii
Discriminant Distributions and Transportability Analysis	xiv
Statistical Methods for Seismic Event Identification	xvi
Bibliography of Publications Sponsored by Contract	xviii
1. Introduction	1
2. Statistical Methods	3
2.1. Motivation for Bootstrap Generalized Likelihood Ratio Tests	3
2.2. Generalized Likelihood Ratio Outlier Test	5
2.2.1. Overview	5
2.2.2. Methodology	6
2.2.3. Specific Case	8
2.3. Generalized Likelihood Ratio Classification Test	10
2.3.1. Overview	10
2.3.2. Methodology	13
2.3.3. Specific Cases	15
2.4. Preliminary Training Set Analyses	19
2.4.1. Tests for Normality	19
2.4.2. Box-Cox Transformations	20
2.4.3. Tests for Equal Covariance Matrices	21
2.4.4. Treatment of Missing Data Values	22
2.4.5. Feature Selection	22
3. Monitoring Applications	23
3.1. Data Sets	23
3.2. Multivariate Graphical Displays: Andrews' Plots	25
3.3. Results of Preliminary Training Set Analyses	28
3.4. Monitoring a New Region	37
3.5. Visualization of Likelihood Ratio Test Results	40
3.6. Contaminated Training Set Study	45

4. Identification Analysis of the 31 December 1992 Novaya Zemlya Event	46
4.1. Introduction	46
4.2. Training Sets and Discriminants	46
4.3. Results of Preliminary Training Set Analyses	53
4.4. Discriminant Analysis Results	54
4.4.1. Outlier Test Results.....	54
4.4.2. Classification Test Results	55
4.5. Conclusions and Recommendations Regarding Event 361575	57
5. Conclusions and Recommendations	62
References	65

LIST OF ILLUSTRATIONS

1. Schematic illustrating the outlier test based on the generalized likelihood ratio and the bootstrap procedure. 6
2. Schematic illustrating classification based on the generalized likelihood ratio and the bootstrap procedure. 12
3. Locations of 10 earthquakes and 13 quarry blasts in the Vogtland region, recorded at GERESS, and 25 earthquakes in the Steigen region and 53 quarry blasts in the Kola Peninsula, recorded at ARCESS. 24
4. Scatter plots of Pn/Lg (left) and Pn/Sn (right) ratios of maximum amplitude in 9 frequency bands for the Vogtland events (upper) recorded by GEC2 and the Kola/Steigen events (lower) recorded by ARA0. 26
5. Scatter plots of spectral ratios of maximum (left plot) and rms (right plot) Lg for 3 frequency band combinations for the Vogtland data set. 27
6. Andrew's plots for 10 earthquakes and 13 quarry blasts in the Vogtland region, using the discriminants listed in the upper legend as coefficients of the Fourier series. All three discriminants are larger for the blasts than for the earthquakes. 29
7. Andrew's plots for 25 Steigen earthquakes and 53 Kola blasts, using the discriminants listed in the upper legend as coefficients of the Fourier series. All three discriminants are typically much larger for the blasts than for the earthquakes. 30
8. Histogram plots of Pn/Sn(max; 4-6 Hz) for Steigen earthquakes (left) and Kola quarry blasts (right), before (lower) and after (upper) applying the Box-Cox transformation. The legends in each frame contain the results of the tests for normality. 31

9. Quantile-Quantile plots of $P_n/S_n(\max; 4-6 \text{ Hz})$ for Steigen earthquakes (left) and Kola quarry blasts (right), before (lower) and after (upper) applying the Box-Cox transformation. The legends in each frame contain the results of the tests for normality. Data sets that are approximately normal lie closely along the straight line.	32
10. Same as Figure 8 but for $P_n/S_n(\max; 5-7 \text{ Hz})$	33
11. Same as Figure 9 but for $P_n/S_n(\max; 5-7 \text{ Hz})$	34
12. Same as Figure 8 but for $P_n/L_g(\max; 5-7 \text{ Hz})$	35
13. Same as Figure 9 but for $P_n/L_g(\max; 5-7 \text{ Hz})$	36
14. Quantile-Quantile plots of $P_n/L_g(\max; 8-16 \text{ Hz})$ for Vogtland earthquakes (left) and quarry blasts (right), before (lower) and after (upper) applying the Box-Cox transformation. The legends in each frame contain the results of the tests for normality. Data sets that are approximately normal lie closely along the straight line.	38
15. Same as Figure 14 but for $P_n/S_n(\max; 8-16 \text{ Hz})$	39
16. Graphical representation of the GLR outlier tests for the Vogtland and Kola/Steigen data sets. The distributions shown are smoothed histograms of the likelihood ratio (LR) obtained by bootstrapping. The LRs for the events that were tested are depicted by the triangles. The training and test event sets are listed in the upper legends in each plot. The vertical lines represent critical values of the tests for various significance levels listed in the lowest legend. Events whose LRs fall to the left of the critical value are rejected as belonging to the training set population at that significance level. Because the variances of the Kola quarry blast discriminants are so large, none of the Steigen earthquakes are outliers of the Kola quarry blast population.	42

17. Graphical representation of the GLR classification tests for the Vogtland and Kola/Steigen data sets. The distributions shown are smoothed histograms of the likelihood ratio (LR) obtained by bootstrapping. The vertical lines represent critical values of the test for various significance levels listed in the lowest legend. The thick vertical line corresponds to the critical value which minimizes the total misclassification rate. The LRs for the events that were tested are depicted by the triangles. 44

18. Locations of the seismic events and arrays used in the identification analysis of the Novaya Zemlya event on 31 December 1992 (ORID=361575). Only data from ARA0 were used in the analysis 48

19. Scatter plot of Pn/Sn ratios of maximum amplitude in five frequency bands for all events used in the analysis of the Novaya Zemlya event on 31 December 1992. Event 361575 is listed as RU for regional unknown. 49

20. Andrews' Fourier plots for all events used in the analysis of event 361575. The discriminants used as coefficients of the Fourier series are given in the lefthand legend and the line types associated with the various event types are given in the righthand legend. 50

21. Andrews' Fourier plots for all events used in the analysis of event 361575 except the Kola quarry blasts. 51

22. Andrews' Fourier plots for event 361575 and the Kola quarry blasts. 52

23. Graphical representation of the four GLR outlier tests applied to NZ event 361575. The distributions shown are smoothed histograms of the likelihood ratio (LR) obtained by bootstrapping. The LR for event 361575 is depicted by the triangle. The vertical lines represent critical values of the test for various significance levels listed in the lowest legend. The training sets and discriminants used for each test are listed in the upper and middle legends, respectively. 56

24. Graphical representation of the two GLR classification tests applied to NZ event 361575. The distributions shown are smoothed histograms of the LR obtained by bootstrapping. The LR for event 361575 is depicted by the triangle. The vertical lines represent critical values of the test for various significance levels listed in the lowest legend. The thick vertical line corresponds to the critical value which minimizes the total misclassification rate. The training sets and discriminants used for each test are listed in the upper and middle legends, respectively. 58

25. Scatter plot of Pn/Sn ratios of maximum amplitude in five frequency bands for 3 nuclear explosions at the Novaya Zemlya test site that were recorded in common by the ARCESS (ARA0) and NORESS (NRA0) arrays. The Pn/Sn values for the events recorded by NRA0 are considerably larger than those recorded by ARA0, apparently due to greater Sn attenuation relative to Pn over roughly double the propagation distance from NZ to NRA0 as compared to the NZ to ARA0 path. 61

LIST OF TABLES

1. Results of normality and equal variance tests for ranked features.	53
--	----

EXECUTIVE SUMMARY

At the outset of this project a major focus of the ARPA Nuclear Monitoring Program was to provide accurate yield estimates of underground nuclear tests in order to monitor the Threshold Test Ban Treaty (TTBT) as effectively as possible. Thus, the work we performed during the first portion of this project was directed at using all available useful information to obtain optimal estimates of calibration parameters for seismic magnitude-yield relations. As the Soviet Union collapsed, monitoring emphasis shifted from that of yield determination at a relatively small number of sites to global non-proliferation monitoring associated with the Comprehensive Test Ban Treaty (CTBT) and the Non-Proliferation Treaty (NPT), both yet to be negotiated at the time of this writing. Thus, our focus also shifted to providing robust statistical methods for seismic event identification, particularly of small ($m_b \sim 2.5$) regional events. In this summary we describe the key studies we performed during this project.

Yield Determination

Our yield determination work was split into two main projects: (i) Development of a Bayesian procedure for testing TTBT compliance which, among other information, incorporates seismic magnitude data for which the associated yields are unavailable; (ii) Estimation of multivariate magnitude-log yield calibration parameters at Novaya Zemlya (NZ). The objective of the first project was to improve yield estimates for situations in which seismic magnitude data for underground nuclear tests are relatively abundant, while corresponding yield data are not. The objective of the second project was to estimate the calibration parameters for NZ, first in the absence of explicit yield information and later using yield data given to ARPA by an official of the nuclear test program of the former Soviet Union. This project was motivated by the expectation at the time that future Soviet underground nuclear tests were likely to occur at NZ. Brief descriptions of these projects follow.

A Constrained Bayesian Approach for Testing TTBT Compliance

At the time of this work there was a growing interest in applying Bayesian techniques to the problem of testing TTBT compliance (e.g., Shumway and Der, 1990; Nicholson et al., 1991). During that same time it had been pointed out that previous magnitude data does furnish some information about the unknown parameters in the yield estimation problem even though the associated yields are unknown. For example, for

equal slope parameters, $m_b - m_{Lg}$ does not depend on yield; hence, one can estimate the mean and variance of the difference from data even though the yields are unknown. Typically the size of such data sets is large enough to provide excellent estimates of these parameters. This information can then be used to improve estimates of future yields using a unified magnitude approach (Nicholson et al., 1991).

Our Bayesian approach is a modification of the approach suggested by Shumway and Der (1990). They showed that *a priori* information, supplied for example by a panel of experts, can be useful for estimating log yield confidence intervals, particularly when little or no calibration data (i.e., magnitudes with accompanying yields) are available. Our approach also allows expert opinion to be used as *a priori* information, as well as including available calibration data. In addition, it includes information gained from data for which yields are unknown, to further improve yield estimates. This information is included in the joint prior distribution in the form of constraints on the difference of the mean and variance parameters for m_b and m_{Lg} .

We investigated the impact of the constraint information on the test of TTBT compliance by comparing power curves and F-numbers of two tests based on Bayesian procedures, with and without the constraints. We also compared the Bayesian results to those of two tests based on classical statistics. Our formulation treats the case in which intercepts and the covariance matrix of the random seismic errors are unknown. (Gray et al., 1992, extended the formulation to also treat unknown slope parameters.) To obtain closed form expressions for the power and F-numbers, however, we considered the special case in which the intercepts are treated as unknown, while the random error covariance matrix was treated as known. The comparisons showed that the probability of detecting a TTBT violation using the constrained Bayesian approach is always at least as great as when using the unconstrained approach. Also, the power of the constrained Bayesian approach is greater than that of the classical test of hypothesis using only calibration data and assuming unknown intercept. The constrained approach is particularly useful when only small amounts of calibration data are available and there is more uncertainty in one of the intercepts, e.g., the intercept for m_{Lg} . We also showed that the constrained approach is far more robust than the unconstrained approach when poor *a priori* information is furnished. A detailed presentation of the methodology and the results of the power comparison study were provided by Fisk et al. (1991a).

Multivariate Seismic Calibration for the Novaya Zemlya Test Site

Prior to the availability of yield data for tests conducted at NZ, it was noted that 12 NZ events had very similar magnitudes ($m_b \sim 5.7$), suggesting that devices of similar yield had been tested. Using magnitude data published by Ringdal and Fyen (1991), we obtained estimates of the random error covariance matrix in the linear multivariate magnitude-log yield model and an estimate of the variance of the unified yield estimator. Estimates were also obtained using 16 NZ events. If the yields are all the same, these estimates are unbiased. If they are not the same, we showed that the estimates are actually estimates of upper bounds. A robustness study was performed to determine how the upper bounds vary as a function of departure of the yields from equality. It was shown that the upper bounds of the expected values do increase as the spread in the yields increase, but this just leads to more conservative estimates which are still better than none at all. The details of this study were provided by Fisk et al. (1991b).

At a meeting in Norway in September 1991, Victor Mikhailov presented Soviet yield data for 42 underground nuclear tests that occurred at the Novaya Zemlya test site between 1964 and 1990. The original yield data were provided in the form of a bar graph which we converted to numerical values. We compared these yields to previous estimates by Burger et al. (1986), Nuttli (1988) and Sykes and Ruggi (1988). Their estimates were based on observed seismic magnitudes and magnitude-log yield relations transported from other test sites. While there is some agreement among the four sets of yields, overall there is considerable variation in the values from one set to another. The yields estimated by Nuttli (1988) are, in general, in closest agreement to the Soviet-furnished values.

Using the yield data and published seismic magnitude data, based on NORSAR Lg and P coda, Gräfenberg Lg, and a world-wide m_b (Ringdal and Fyen, 1991) for 18 NZ events, and a similar set of Soviet network magnitudes (Israelsson, 1992), we computed estimates of the multivariate calibration parameters of the linear magnitude-log yield relations. Our analysis treated the slopes, intercepts and random error covariance matrix as unknown parameters to be estimated from data. We also presented a classical confidence interval to estimate future yields, a definition of the F-number, and a hypothesis test of TTBT compliance for future events at NZ. F-number estimates were obtained by applying the jackknife procedure to the data. This study was presented by Fisk et al. (1992). In that report we also noted several discrepancies regarding the

original Soviet-supplied yield graph which were first pointed out to us by Richards (1992) and later substantiated by Adushkin et al. (1992). Thus, our results were presented with the caveat that they are contingent on the accuracy of the original yield data.

Seismic Event Identification

The focus of our work for the past one and one-half years has been on developing statistical methods and applying them to the problem of seismic event identification, particularly for regional events. This work is directed at providing robust statistical methods to perform outlier detection, script-matching and classification, to assess associated error rates, and to address statistical issues regarding transportability of discriminants.

Under the current plan by the ARPA Nuclear Monitoring Research Office to concentrate CTBT monitoring efforts on regions of interest (Ryall, 1993a), there are many regions at present, and we should expect other regions in the future, for which relevant seismic data are not available, particularly for small underground nuclear explosions. A situation that may arise is one in which we need to begin monitoring a region of interest immediately. If a seismic station or array is installed within regional distances, a monitoring procedure is needed that can be applied as soon as the station starts recording events. It is likely in this situation that the type of events being observed will be unknown. Artificial intelligence techniques that require training data for all relevant event types have limited utility in this situation where sufficient training data are lacking. Discrimination using a rule-based approach may not apply in this situation until it is determined how to transport discrimination rules from one region to another.

However, under the assumptions that only a small number of nuclear tests would be performed in a region relative to the number of earthquakes or conventional mining blasts and that we have at least one seismic characteristic that discriminates in that type of region, even though we may not know the details of how the discriminant is distributed for the various types of events there, we can test whether the events observed are "business as usual" in that region or not. This is the key motivation behind our outlier test. (Note that if the first assumption fails, i.e., if a region is completely aseismic and with no mining activity, any new event would be flagged as an outlier warranting further investigation.)

A procedure is also needed to optimize event identification accuracy if more information is available. If training data are available for more than one class of events or as we learn how to transport discriminants from one region to another, a robust classification method is needed that can use this information rigorously to minimize identification error rates. Together the outlier and classification tests allow monitoring to be performed under the range of training data availability that will be encountered and do not rely on fixed discrimination rules that may not apply to a new region.

Even so, it is clear that additional useful information regarding seismic discriminants for all event types will lead to more events being identified correctly. Thus, it is vital that we understand why discriminants work in particular regions and determine how they can be transported from regions for which data are abundant to those where data are lacking.

Based on these motivating factors, this portion of the project has included two distinct efforts: (i) Analysis of discriminant distributions and their transportability; (ii) Development and event identification applications of statistical methods for outlier detection and classification.

Discriminant Distributions and Transportability Analysis

We are in the process of performing a statistical study of the distributions of seismic amplitude ratios and their transportability using data for a set of 114 nuclear tests conducted at 7 Soviet test sites and recorded at 11 seismic stations within the former Soviet Union. The majority of tests were conducted at 4 Semipalatinsk test sites (Balapan NE, Balapan SW, Degelen and Murzhik) and the Matochkin Shar site at NZ. The 11 Soviet stations are at Apatity (APA), Arti (ARU), Bodaybo (BOD), Chusal (CHS), Fruenze (FRU), Norilsk (NRI), Novosibirsk (NVS), Obninsk (OBN), Talaya (TLY), Tupik (TUP) and Uzhgorod (UZH). Soviet seismic data are also available for explosions at Azghir, two locations in Central Siberia and one location near Lake Baikal (Israelsson, 1992); however, we focused on the tests at Semipalatinsk and NZ.

RMS amplitude data for P^* and L_g , in frequency bands at 0.5-1.0, 1.0-2.0, 2.0-4.0 and 4.0-6.0 Hz, were provided to us by Schwartz (1992). (P^* denotes the RMS amplitude in the first 20 second of the P wave. Also, amplitudes in the three highest frequency bands were provided for only a small number of NZ events.) Waveforms at each station were normalized for instrument response to a common instrument at OBN. This is a

similar data set to the one analyzed by Israelsson (1992). Most of these events were in the mb magnitude range of 4.8-6.0. Although they are considerably larger than those expected to stress the limits of CTBT or NPT monitoring capabilities, this is a unique data set that warrants full investigation simply because it involves such a large number of nuclear explosions, paths and stations.

The objectives of this analysis are to determine (1) what the discriminant distributions are so that identification error rates can be estimated on a region by region basis, (2) on what distance scales and for what geological characteristics do regions need to be divided in order to monitor them effectively, and (3) what information is needed to transport discriminant distributions from one site-station combination to another so that new events can be identified and error rates estimated in regions lacking data. To do this we looked at each phase (P* and Lg in this study) used in discriminants such as amplitude ratios, normalized by an independent measure of source size, namely, ISC world-wide mb. Looking at each phase individually was done to determine how each is subject to path variability. We then tested the variances for equality for various paths, as well as the means before and after applying distance and station corrections. The purpose of the corrections is to remove biases in the means due to propagation and near-station geology effects; the corrections do not affect the variances.

Following Israelsson (1992), we used a linear model to separate log RMS amplitudes into source, distance, station and random error contributions. It is important to note that the model we used assumes that distance corrections depend only on distance for western Eurasia and not individual receiver-source geology. A single attenuation coefficient for Eurasia is estimated from the data. For P amplitudes we used empirical Slunga-Veith-Clawson distance corrections for Eurasia (GSE/SW/62, 1988) which depend on distance but not on the specific path. For Lg amplitudes an analytic expression for the geometrical spreading (Nuttli, 1986) was used and the attenuation coefficient was treated as one of the unknown parameters estimated from data.

Jih and Lynnes (1992) discuss an alternative model in the context of estimating source size, in which each path has its own attenuation coefficient which is solved for simultaneously to solving for the other unknown parameters of the model. This model is more appealing from the standpoint of treating the path differences in more detail and removing path biases from the means. However, for purposes of understanding transportability, it is hoped that the linear model with a single attenuation coefficient for

Eurasia is adequate. If not, we will need to know attenuation coefficients accurately over a relatively fine global grid, requiring either considerable data or *a priori* knowledge. This may tend to defeat the whole program of transporting discriminants.

Using the linear model with a single attenuation coefficient (for each phase in each frequency band) for Eurasia, tests of hypothesis based on the F-distribution for the variances and Student's t-distribution for the means were applied to the path and station corrected amplitudes. Since there was very little data for NZ in the higher frequency bands we concentrated on the 0.5-1.0 Hz band. At 0.05 significance level we found that roughly 78% (60%) of the variances and 41% (53%) of the means for log RMS P* (Lg) amplitudes were accepted by the tests as equal. Uniform distance corrections for Eurasia and the assumption of similar source coupling at NZ and Semipalatinsk do not remove all of the path bias but it remains to be seen whether the remaining biases for explosions are significant as compared to the differences in earthquake and explosion means. We are in the process of acquiring earthquake data for similar paths and more high frequency data for NZ events. We intend to complete this study under a follow-on contract and report our results in detail at a later time.

Statistical Methods for Seismic Event Identification

The main thrust of our event identification work and the body of this report are focused on algorithm development and application to statistical discriminant analysis of seismic events. We have developed outlier detection (Baek et al., 1992) and classification (Baek et al., 1993; Gray et al., 1993) methods to perform event identification under a wide range of training set availability conditions and to treat the statistical distributions of discriminants rigorously. Among other applications, we show how the outlier test can be used to detect peculiar events in data sets lacking ground-truth, to "script-match" mining blasts to associate them with particular mines (e.g., Baumgardt, 1987), or to monitor new regions for which training events of only one type may be available. The classification test is based on a similar methodology but provides more accurate event identification for cases in which training data are available for more than one class of events.

We have also implemented a number of preliminary algorithms, tests and data transformations to be applied to training sets to (1) ensure that appropriate assumptions of the statistical event identification tests hold, (2) treat missing data values in the training set, and (3) select the optimal set of discriminants which minimize misclassification

errors. The combination of these techniques provides a robust statistical framework to identify seismic events and estimate error rates rigorously under extremely general or limited conditions. We are combining these algorithms into a single software package with an X Window graphical user interface under a separate but strongly related project.

We have applied these methods to identification problems using features that were extracted from seismic waveform analysis performed by Baumgardt (1993a) using the Intelligent Seismic Event Identification System (ISEIS) (e.g., Baumgardt et al., 1991). Data sets used in our analysis include earthquakes and quarry blasts in the Vogtland region in Germany, quarry blasts in the Kola Peninsula, earthquakes in the Steigen region in Norway and in the direction of the Mid-Atlantic Ridge relative to ARCESS, and nuclear explosions at NZ. These events were recorded at combinations of ARCESS, GERESS and NORESS.

The main body of this report includes descriptions of the outlier and classification tests, as well as the supporting algorithms to optimize their use and rigor. Results of applications to regional data sets are presented below to illustrate how these methods can be used for CTBT or NPT monitoring applications under less than ideal circumstances. Among analyses performed, we show how the outlier test can be used to monitor successfully new regions with limited training data and that the classification test improves event identification accuracy as more information becomes available. A study was also performed in which training sets are contaminated intentionally with a few events of other types to determine if the outlier test can detect them. Last, we applied the tests to event identification of a regional event of unknown type recorded at ARCESS which occurred on Novaya Zemlya on 31 December 1992. This application has also been included as a paper by Fisk and Gray (1993) in a volume of papers by several sets of authors compiled by Ryall (1993b) on this event.

BIBLIOGRAPHY OF PUBLICATIONS SPONSORED BY CONTRACT

The following papers, technical reports and notes have been sponsored by this contract:

J. Baek, H. L. Gray, W. A. Woodward and M. D. Fisk (1993). A Bootstrap Generalized Likelihood Ratio Test in Discriminant Analysis, submitted to *J. Am. Stat. Assoc.*, Southern Methodist University, Dallas, TX.

Fisk, M. D., R. W. Alewine, H. L. Gray, G. D. McCartor (1992). Multivariate Seismic Calibration for the Novaya Zemlya Test Site, PL-TR-92-2251, Phillips Laboratory, Hanscom AFB, MA, ADA261725.

Fisk, M. D., and H. L. Gray (1993). Event Identification Analysis of the Novaya Zemlya Event on 31 December 1992 using Outlier and Classification Likelihood Ratio Tests, MRC-R-1449, Mission Research Corp., Santa Barbara, CA. Also presented as a paper in *The Novaya Zemlya Event of 31 December 1992 and Seismic Identification Issues* (Ed. A. Ryall), Supplemental volume of papers to Proceedings of the 15th Annual Seismic Research Symposium, 8-10 September 1993, Vail, CO, PL-TR-93-2160, ADA271458.

Fisk, M. D., H. L. Gray, G. D. McCartor and G. L. Wilson (1991a). A Constrained Bayesian Approach for Testing TTBT Compliance, PL-TR-91-2170, Phillips Laboratory, Hanscom AFB, MA, ADA253288.

Fisk, M. D., H. L. Gray, G. D. McCartor and G. L. Wilson (1991b). Estimates of the Covariance Parameters for Novaya Zemlya and Robustness of Point Estimates for Unequal Yields, MRC-N-940, Mission Research Corp., Santa Barbara, CA.

1. INTRODUCTION

Among the crucial statistical issues associated with seismic event identification are (1) availability of useful training sets and (2) complicated statistics of discriminants. Data sets for new regions of interest may be extremely limited. Path differences often prohibit direct application of training sets or discrimination rules to new regions. Ground-truth data sets are being established (e.g., Grant et al., 1993) but it is a very tedious process. Consequently, regions may have to be monitored before applicable ground-truth data sets can be established.

In addition, discriminants rarely have normal distributions, and the distribution type and covariance structure typically differ from one event type to the next (e.g., Baumgardt, 1992b; Baumgardt et al., 1992). Some very useful discriminants are in fact discrete rather than continuous and should be treated accordingly. In many cases some of the discriminant values are missing for particular events. Missing values may be due to blockage or strong attenuation of particular seismic phases, poor signal-to-noise, interference by other events, or instrument malfunction, for example. Techniques are needed to make use of partial data for events in training sets rather than simply discarding particular discriminants or events with missing values.

Combinations of these and other issues have limited useful application of many statistical discriminant analysis methods in the past. It is vital, however, for monitoring applications that these issues are all addressed with statistical rigor so that classification accuracy can be optimized and so that estimated identification probabilities and error rates have precise meaning. Improper treatment of discriminant statistics can lead to unnecessary misclassifications and biased estimates of error rates.

To address these issues MRC and SMU have developed and applied statistical methods for outlier detection (Baek et al., 1992) and classification (Baek et al., 1993). The methods, based on the generalized likelihood ratio (GLR) and bootstrap techniques, have considerable flexibility to rigorously treat a mixture of continuous (e.g., amplitude and spectral ratios, spectral and cepstral variances, etc.) and discrete (e.g., presence of cepstral peaks, seismicity, deep/shallow, offshore/onshore, etc.) features, normal or non-normal statistics, equal or unequal covariance matrices, and even different distribution functions for different event types. The methods are presented in detail in Section 2. There is also extensive literature on likelihood ratio and bootstrap techniques, although

not used together nor with as much generality. For reviews see Anderson, 1984; Seber, 1984; Johnson and Wichern, 1988).

The GLR outlier test is particularly useful for situations in which training data is limited to one class. Here a hypothesis test is performed to determine whether an event belongs to the same population as the training data or not. For example, when monitoring a new region there may only be a handful of previously recorded events. In fact, we may not even know what the event type is for a particular set. The GLR outlier test can still be applied to flag peculiar events that are not "business as usual." A related application of the outlier test is to test training sets which lack ground-truth for possible contamination. In this way, peculiar events can be flagged for further expert analysis or corroboration with bulletins, etc. Thus, the GLR outlier test allows monitoring to be performed under conditions for which minimal or contaminated information is available.

For situations where more information is available the GLR classification method is an improved procedure to optimize event identification accuracy. Here training data for all available classes are used and the event in question is allocated into one of two or more classes. (Our classification routine treats only two populations at present, so we adopt a sequential testing procedure for cases of more than two classes.) In addition to proper treatment of discriminant statistics, the Bootstrap GLR method allows the overall cost of misclassification to be minimized or, alternatively, the false alarm rate associated with misclassifying a particular event type (e.g., nuclear explosions) to be controlled. The latter approach is particularly important since for most monitoring applications we want to limit the number of undetected nuclear explosions to a very small number.

There are known methods for controlling the false alarm rate for the case where normal distributions with the same covariance matrix are assumed. Our approach, however, can be applied to not only this case but to the case of normal distributions with different covariance matrices and the case of a mixture of discrete and continuous variables. The method can, in fact, be applied to any distribution for which the maximum likelihood estimates exist.

The remainder of this report is organized as follows. In Section 2 we describe the statistical discriminant analysis methods, as well as preliminary training set analyses which are applied to ensure that appropriate assumptions, employed by the discriminant analysis methods, hold. In Section 3 we present applications to various monitoring

scenarios using data sets with a high degree of ground-truth. We describe the regional training sets and discriminants used and present novel multivariate graphical displays of discriminants as a means to better visualize comparisons of event characteristics as a whole. We also describe results of preliminary training set analyses and specific monitoring applications. In Section 4 we present results of identification analysis applied to the seismic event that occurred on Novaya Zemlya on 31 December 1992. In Section 5 we provide our conclusions and recommendations for further work.

2. STATISTICAL METHODS

2.1. Motivation for Bootstrap Generalized Likelihood Ratio Tests

Both the outlier detection and classification approaches are based on the same methodology of the likelihood ratio and the bootstrap. Likelihood ratio tests are well-established statistical techniques that have been used for a variety of applications to test statistical hypotheses. For example, they have been used to test for equality of means, variances, correlations and covariance matrices, independence of random variables, linearity of regression relations, etc. (For reviews of these and other applications, see Anderson, 1984; Seber, 1984). They have also been used in the past to test for outliers or to classify events, usually under the assumption of multivariate normality. They are widely used because they can in principle be applied to test any hypothesis for which the likelihood functions exist and they usually have good power characteristics.

The basic idea behind the likelihood ratio test is to form the ratio of likelihoods

$$LR = \frac{\max L(\text{under null hypothesis})}{\max L(\text{under alternative hypothesis})}$$

where the numerator is computed under the null hypothesis that is being tested, given the data, and the denominator is computed under the alternative hypothesis, given the data. Since the true parameters are unknown, maximum likelihood estimates (MLEs) are computed, subject to the constraints of the particular hypothesis, and inserted into the likelihoods. This yields the ratio of maximized likelihoods. Intuitively, small values of LR indicate that the null hypothesis should be rejected.

For example, the null hypothesis might be that the means of two populations are equal ($\mu_1 = \mu_2$) and the alternative is that they are not ($\mu_1 \neq \mu_2$). Given two samples of data, the MLE of a only one mean is estimated from the pooled samples under the null hypothesis and inserted into the likelihood, while under the alternative each mean is estimated separately from the two samples.

For either the outlier or classification tests, the likelihood ratio test statistic is a mathematical mapping of the multivariate discriminants for the all of the training events and the new event in question into a single univariate variable. Given the discriminant values for the training data and new event, the value of the likelihood ratio is simply a number that measures how likely it is that the new event belongs to a particular class of events versus the alternative that it does not. In general the likelihood ratio is a randomly distributed variable just as the training data and new event are considered to be random samples. We need to know the distribution of the likelihood ratio in order to know how to set the critical value of the test of hypothesis.

Even if the discriminants have normal distributions, the distribution of the likelihood ratio is quite complicated in general. Anderson (1973) and Kanazawa (1979) obtained asymptotic forms of the likelihood ratio distribution for large sample sizes from multivariate normal distributions with equal covariance matrices assumed. Application to discriminants with unequal covariance matrices, arbitrary distributions, or the inclusion of continuous and discrete discriminants makes determination of even asymptotic forms of the likelihood ratio distribution intractable for most cases.

If numerous samples of the training data and the new event to be tested were available, the distribution could be estimated empirically. Unfortunately this is almost never the case. This situation has arisen in many statistical problems in the past, which has led to development of the jackknife and bootstrap (Efron, 1979, 1982) resampling techniques. Here we use the parametric bootstrap to numerically generate random samples of data using the distribution assumed for the likelihood with parameters estimated from the original data. Inserting the bootstrapped data in the likelihood ratio for many samples allows the distribution of the likelihood ratio to be estimated. Given the bootstrap distribution of the likelihood ratio, the critical value is set so that the test has a specified significance level. The parametric bootstrap algorithms for the outlier and classification tests are described explicitly in Sections 2.2.2 and 2.3.2, respectively.

The key innovation of our approach is the combination of the likelihood ratio and the bootstrap techniques. This permits application to any distribution for which the MLEs exist, while controlling the false alarm rate of a particular type of event. The bootstrap also allows classification to be performed in the conventional sense in which a loss function associated with misclassifying all types of events is minimized. Thus discriminant statistics can be treated rigorously regardless of how complicated they are and we can specify how we want to control the misclassification probabilities.

2.2. Generalized Likelihood Ratio Outlier Test

2.2.1. Overview

A statistic commonly used for outlier detection is Wilks' statistic (Wilks, 1963), which is based on the likelihood ratio under the assumptions of multivariate normality and equal covariance matrices. It is easily shown that Wilks' statistic is equivalent to Hotelling's T^2 statistic (Anderson, 1984) which is generally used for testing the equality of the means of two normal populations. Caroni and Prescott (1992) showed that a hypothesis-testing approach with the likelihood ratio statistic can be used successfully for outlier detection for multivariate normal distributions. Here we extend the formalism to treat a combination of discrete and continuous variables with arbitrary distributions.

The statistic on which the outlier test is given by the ratio of the maximized likelihood computed under the null hypothesis that the event in question belongs to the same class as the training set, to the maximized likelihood computed under the alternative hypothesis that it does not belong to this class. The bootstrap technique is used to set the critical value of the test such that the false alarm rate is fixed. Figure 1 shows a schematic of this procedure, where the distribution of the likelihood ratio is obtained by bootstrapping. The critical value is set such that there are only $100\alpha\%$ of the earthquakes, in this example, in the tail that would be rejected falsely, where α is the significance level of the test. Values of the likelihood ratio less than the critical value are rejected as belonging to the same population.

It is important to understand what the outcome of this test means at a given significance level. Intuitively, the outlier test determines whether an event is similar or significantly different than events in the training set, based on the discriminant values of the event being tested relative to the distribution for the training data. If the null

hypothesis is rejected, the outlier test makes no statement as to which other class of events the tested event might belong. The significance level is the probability that an event is rejected when it should not be, i.e., when it actually belongs to the same population as the training sample.

Outlier Test Likelihood Ratio

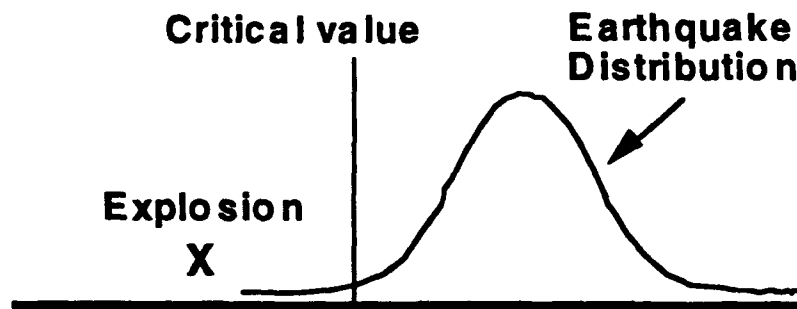


Figure 1. Schematic illustrating the outlier test based on the generalized likelihood ratio and the bootstrap procedure.

For example, if an event is rejected as belonging to the class of nuclear explosions at the 0.025 significance level, then we can say that this event has been rejected as a nuclear explosion with a 2.5% probability of falsely rejecting explosions. Thus, if an event is rejected at a small significance level, the probability that a mistake was made is also small. Unfortunately, the significance level of the test does not have a simple relation to the probability that the event was accepted when it should not have been. This probability depends on unknown distributions for other types of events.

2.2.2. Methodology

Following Baek et al. (1992), suppose for each event the vector of discriminants $\mathbf{V} = (\mathbf{Z}, \mathbf{X})$ is used to characterize the event type, where $\mathbf{Z} = (Z_1, \dots, Z_r)'$ are discrete variables and $\mathbf{X} = (X_1, \dots, X_p)'$ are continuous ones. (The primes denote the vector transpose.) Let $f(\mathbf{v}|\theta^{(i)})$ be the probability or probability density function of \mathbf{V} evaluated at \mathbf{v} , given the set of parameters $\theta^{(i)}$. Suppose also that a training sample for N events, $\mathbf{v}_1, \dots, \mathbf{v}_N \in \pi_1$, are available from a single population π_1 and that a new observation \mathbf{v}_{N+1} must be classified as whether or not it belongs to the same population as

the training sample. Allocation of \mathbf{v}_{N+1} as an outlier or not is accomplished by testing the hypothesis

$$H_0: \mathbf{v}_1, \dots, \mathbf{v}_{N+1} \in \pi_1 \text{ versus } H_1: \mathbf{v}_1, \dots, \mathbf{v}_N \in \pi_1, \mathbf{v}_{N+1} \in \pi_2. \quad (1)$$

We use the generalized likelihood ratio to construct a test. The likelihood of the training sample is given by

$$L(\theta^{(1)} | \mathbf{v}_1, \dots, \mathbf{v}_N) = \prod_{j=1}^N f(\mathbf{v}_j | \theta^{(1)}). \quad (2)$$

The likelihood of a new observation is $L_2(\theta^{(2)} | \mathbf{v}_{N+1}) = f(\mathbf{v}_{N+1} | \theta^{(2)})$. Testing the hypothesis in Equation (1) is equivalent to testing $H_0: \theta^{(1)} = \theta^{(2)}$ versus $H_1: \theta^{(1)} \neq \theta^{(2)}$. Using these expressions for the likelihoods, the generalized likelihood ratio is given by

$$\lambda = \frac{L_2(\hat{\theta}_0^{(1)} | \mathbf{v}_{N+1}) L(\hat{\theta}_0^{(1)} | \mathbf{v}_1, \dots, \mathbf{v}_N)}{L_2(\hat{\theta}_1^{(2)} | \mathbf{v}_{N+1}) L(\hat{\theta}_1^{(1)} | \mathbf{v}_1, \dots, \mathbf{v}_N)}, \quad (3)$$

where $\hat{\theta}_0^{(1)}$ is the MLE of $\theta^{(1)}$ under the null hypothesis and $\hat{\theta}_1^{(1)}$ and $\hat{\theta}_1^{(2)}$ are the MLEs of $\theta^{(1)}$ and $\theta^{(2)}$ under the alternative hypothesis.

It intuitively follows that small values of λ provide evidence against H_0 and thus the generalized likelihood ratio test is: reject H_0 if $\lambda \leq \lambda_\alpha$, where λ_α is the critical value such that the test has α significance level. That is, if we know the distribution of λ given that the null hypothesis is true, $f_\lambda(\lambda | H_0)$, the critical value λ_α is set such that

$$P[\lambda \leq \lambda_\alpha | H_0] = \int_{-\infty}^{\lambda_\alpha} d\lambda f_\lambda(\lambda | H_0) = \alpha \quad (4)$$

This is the false alarm rate or Type I error rate of the test. It is the probability that an event is rejected as belonging to π_1 when in fact it does belong.

In most cases it is difficult to obtain a closed form expression for the distribution of λ . The distribution, however, can be approximated by using the bootstrap method. Since the form of the probability density function $f(\mathbf{v} | \theta^{(1)})$ is assumed known, the bootstrap samples can be obtained from the density function. This is called the

parametric bootstrap (Efron, 1979) and we employ it here. The algorithm consists of performing the following steps:

1. Compute $\hat{\theta}$ from $\mathbf{v}_1, \dots, \mathbf{v}_N$
2. Generate random samples $\mathbf{v}_1^*, \dots, \mathbf{v}_{N+1}^*$ from $f(\mathbf{v}|\hat{\theta})$
3. Compute λ^* ($= \lambda$ computed from $\mathbf{v}_1^*, \dots, \mathbf{v}_{N+1}^*$) (5)
4. Repeat steps 2 and 3 M times
5. Set λ_α from empirical (bootstrap) distribution of λ^*
 $\lambda_\alpha^* = \text{int}(\alpha(M+1))$ smallest value of $\{\lambda_j^*\}_{j=1}^M$.

The critical value is set in step 5 such that 100 α % of the test events, bootstrapped under H_0 , are rejected (MacLachlan, 1987). This provides an approximate α significance level test.

2.2.3. Specific Case

To clarify this formalism it is useful to illustrate the method for a particular case. Suppose that there is a single discrete variable, $Z = 0$ or 1 , which has a Bernoulli distribution and the continuous variables have a multivariate normal distribution. For example, the discrete variable could be "presence of cepstral peaks" indicative of a ripple-fired mining blast. To simplify the following expressions for the sake of illustration, assume that the continuous variables are independent of the discrete one. (See Baek et al., 1992, for treatment of correlated discrete multinomial and continuous multivariate normal variables.) Also, assume that there is a single covariance matrix since in this setting there is insufficient evidence to suggest otherwise. In this case the likelihoods are given by

$$L(\theta^{(1)}|\mathbf{v}_1, \dots, \mathbf{v}_N) = p_1^m (1-p_1)^{N-m} \{(2\pi)^p |\Sigma|\}^{-N/2} \exp\left\{-\frac{1}{2} \sum_{i=1}^N (\mathbf{x}_i - \mu^{(1)})' \Sigma^{-1} (\mathbf{x}_i - \mu^{(1)})\right\} \quad (6)$$

$$L_2(\theta^{(2)}|\mathbf{v}_{N+1}) = p_2^I (1-p_2)^{1-I} \{(2\pi)^p |\Sigma|\}^{-1/2} \exp\left\{-\frac{1}{2} (\mathbf{x}_{N+1} - \mu^{(2)})' \Sigma^{-1} (\mathbf{x}_{N+1} - \mu^{(2)})\right\}, \quad (7)$$

where m is the number of training events whose discrete value is 1, $I = 0$ or 1 is the discrete value for the new event, and p_1 and p_2 are the probabilities of obtaining an observation with a discrete value of 1 for the populations from which the training set and

the new event, respectively, were sampled. These probabilities are equal if the null hypothesis is true.

Under the null hypothesis H_0 the MLEs ($\hat{\theta}_0^{(1)} = \hat{\theta}_0^{(2)}$) are given by

$$\begin{aligned}\hat{p}_{10} &= \hat{p}_{20} = (m + I) / (N + 1) \\ \hat{\mu}_0^{(1)} &= \hat{\mu}_0^{(2)} = \frac{1}{N + 1} \sum_{i=1}^{N+1} \mathbf{x}_i \\ \hat{\Sigma}_0 &= \frac{1}{N + 1} \sum_{i=1}^{N+1} (\mathbf{x}_i - \hat{\mu}_0^{(1)})(\mathbf{x}_i - \hat{\mu}_0^{(1)})',\end{aligned}\tag{8}$$

and under the alternative hypothesis H_1 the MLEs ($\hat{\theta}_1^{(1)} \neq \hat{\theta}_1^{(2)}$) are given by

$$\begin{aligned}\hat{p}_{11} &= m / N, \quad \hat{p}_{21} = I \quad (= 0 \text{ or } 1) \\ \hat{\mu}_1^{(1)} &= \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i, \quad \hat{\mu}_1^{(2)} = \mathbf{x}_{N+1} \\ \hat{\Sigma}_1 &= \frac{1}{N + 1} \left[\sum_{i=1}^N (\mathbf{x}_i - \hat{\mu}_1^{(1)})(\mathbf{x}_i - \hat{\mu}_1^{(1)})' + (\mathbf{x}_{N+1} - \hat{\mu}_1^{(2)})(\mathbf{x}_{N+1} - \hat{\mu}_1^{(2)})' \right]\end{aligned}\tag{9}$$

Using Equations (6)-(9) and performing some simplifying algebra, the likelihood ratio defined in Equation (3) is given by

$$\lambda = \frac{\left(\frac{m+I}{N+1}\right)^{m+I} \left(1 - \frac{m+I}{N+1}\right)^{N+1-m-I}}{\left(\frac{m}{N}\right)^m \left(1 - \frac{m}{N}\right)^{N-m}} \cdot \left| \frac{\hat{\Sigma}_0}{\hat{\Sigma}_1} \right|^{(N+1)/2}\tag{10}$$

Note that in the absence of the discrete Bernoulli variable, the part of the expression in Equation (10) involving the covariance matrix estimates is equivalent to Wilks' likelihood ratio statistic (Wilks, 1963) which is commonly used for outlier detection for the case of a multivariate normal distribution. To set the critical value of the test we use the algorithm expressed in (5) where the bootstrap samples are generated from Bernoulli and multivariate normal distributions. This example illustrates how the likelihood ratio can be generalized to include discrete variables. In fact, any set of discrete and continuous variables can be included, provided their joint probability density function is defined and the MLEs exist.

Baek et al. (1992) investigated the performance of the test based on the likelihood ratio defined in Equation (10) with simulations. They examined how well the bootstrap estimate of the critical value λ_α^* approximates the true critical value λ_α as functions of the Bernoulli probabilities p_1 and p_2 and separations of the means μ_1 and μ_2 . Since the parameters were known for the simulation, a very precise estimate of the true critical value was obtained using a Monte Carlo procedure. They found that for a training sample size of 100 the bootstrap estimate of the critical value is nearly identical to the true value.

2.3. Generalized Likelihood Ratio Classification Test

2.3.1. Overview

Classical approaches for discriminant analysis in two populations rely on the ratio of the probabilities or probability density functions. The classification rule based on the ratio is optimal in the sense that it minimizes the total probability of misclassification (Welch, 1939) or the total cost of misclassification (Anderson 1984, Chapter 6). Under the assumptions of normality, equal covariances, and unknown parameters for the variables, Anderson (1951) derived a classification rule based on the linear discriminant function, which is known as Anderson's W statistic, by substituting estimates for the parameters in the ratio. When the covariance matrices are not equal, replacing each parameter by its estimate gives the classical quadratic discriminant function (Seber, 1984, p297; Anderson, 1984, p235).

Among other classification rules is a hypothesis-testing approach which is derived by obtaining the generalized likelihood ratio. This rule, based on the assumption of normal distributions with equal covariance matrices, was proposed by Anderson (1958), studied by John (1963), and has become known as John's Z statistic. Krzanowski (1982) extended this approach to mixed discrete and continuous variables. For more discriminant procedures in the mixture case, see Knoke (1982), Krzanowski (1975, 1979, 1980), and Tu and Han (1982).

Most of these classical classification rules allocate the event to be classified to one of the populations if the ratio is less than a cut-off point, and to the other otherwise. The cut-off point is usually based on the probabilities of drawing an observation from the individual populations and/or the costs of misclassification. Associated with these

procedures are the resulting misclassification probabilities. However, when one of the probabilities of misclassification is extremely important, one may want to determine the cut-off point to allow this probability of misclassification to be prespecified. When the p -dimensional characteristic variable is multivariate normal and the covariance matrices of two populations are equal, Anderson (1973) and Kanazawa (1979) obtained asymptotically normal expansions of the distribution of the W and Z statistics, respectively, which are used to find the cut-off point for a fixed value of the particular misclassification probability. In other cases (e.g., for unequal covariance matrices, non-normal distributions, or a mixture of continuous and discrete variables, etc.) it is very difficult to derive the asymptotic distribution of the classification rules analytically since they are usually nonlinear.

The objective here is to develop a classification rule for which one of the misclassification probabilities can be controlled, and that can also be applied to any mixture of continuous and discrete variables for which the likelihood function is defined. The proposed classification procedure is constructed by applying the bootstrap to the generalized likelihood ratio. Although the approach in which we fix one of the misclassification probabilities is actually a test of hypothesis, we show how the bootstrap GLR method can also be used for classification in the classical sense, whereby minimizing the expected cost of misclassification.

The test statistic for classification into one of two groups is similar to the likelihood ratio for the outlier test except that now it is given by the ratio of the maximized likelihood computed under the null hypothesis that the event being tested belongs to one of the classes, to the maximized likelihood computed under the alternative hypothesis that the event belongs to the other class. Small values of the likelihood ratio indicate that the event does not belong to the first class. The bootstrap technique is used here to set the critical value of the test at fixed significance level. Figure 2 shows a schematic of this procedure. The distribution on the right is obtained by bootstrapping, where the new event is sampled from the explosion population, in this example, while the distribution on the left is obtained by sampling the new event from the earthquake population. The critical value is set such that there are only $100\alpha\%$ of the explosions in the tail that would be rejected falsely, where α is the significance level.

The classification test is similar to the outlier test in that it determines whether the event in question should be rejected as belonging to a particular class of events. In this

case, however, the test makes use of training data for two event types, allowing the event to be classified preferentially as one or the other type of event. The significance level, now set with respect to one of the two classes, is still the probability that a new event is rejected as belonging to that particular class when it should not be rejected.

Classification Likelihood Ratio

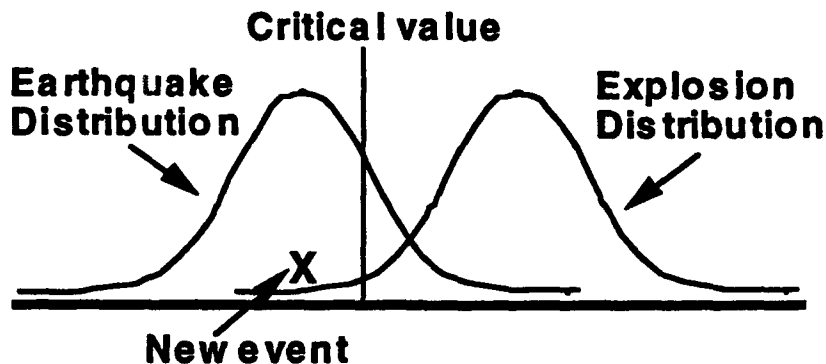


Figure 2. Schematic illustrating classification based on the generalized likelihood ratio and the bootstrap procedure.

For example, suppose we use both nuclear explosion and earthquake training data to classify an event in question. If the event is rejected as belonging to the nuclear explosion class, with the significance level set at 0.025 with respect to that class, then we can say that it was rejected with only 2.5% probability of falsely rejecting an explosion. Thus the event is classified in this case as looking more like an earthquake than an explosion with only 2.5% probability of it actually being an explosion. As before, the significance level does not have a simple relation to the probability that the event is correctly identified as an earthquake given that it actually was an earthquake. However, this probability can be estimated straightforwardly, based on the critical value of the test and the distribution of the discriminants for the earthquake class.

Furthermore, a rejection that it was an explosion does not mean that the event could not be of some other type not included in the training set, only that its classification as an earthquake is preferred to classification as an explosion in this example. However, the test does make a strong statement about the event not being an explosion. So in this

sense, rejections of hypothesis at fixed significance level are much more meaningful in order to say that an event is not of a given type than it is to say that the event belongs to a particular class.

2.3.2. Methodology

Following Baek et al. (1993), suppose for each event the vector of discriminants $\mathbf{V} = (\mathbf{Z}, \mathbf{X})$ is used to characterize the event type, where $\mathbf{Z} = (Z_1, \dots, Z_r)'$ are discrete variables and $\mathbf{X} = (X_1, \dots, X_p)'$ are continuous ones. Let $f_i(\mathbf{v}|\theta^{(i)})$ be the joint probability or probability density function of \mathbf{V} evaluated at \mathbf{v} , if \mathbf{v} comes from population π_i , given the set of unknown parameters $\theta^{(i)}$. Suppose also that we have training samples $\{\mathbf{v}_1^{(1)}, \dots, \mathbf{v}_{N_1}^{(1)}\}$ of size N_1 and $\{\mathbf{v}_1^{(2)}, \dots, \mathbf{v}_{N_2}^{(2)}\}$ of size N_2 from populations π_1 and π_2 , respectively, and that a new observation \mathbf{v} must be classified as from either π_1 or π_2 . Classification of \mathbf{v} is accomplished by testing the hypothesis

$$\begin{aligned} H_0: \mathbf{v}, \mathbf{v}_1^{(1)}, \dots, \mathbf{v}_{N_1}^{(1)} \in \pi_1; \mathbf{v}_1^{(2)}, \dots, \mathbf{v}_{N_2}^{(2)} \in \pi_2 \\ \text{versus} \\ H_1: \mathbf{v}_1^{(1)}, \dots, \mathbf{v}_{N_1}^{(1)} \in \pi_1; \mathbf{v}, \mathbf{v}_1^{(2)}, \dots, \mathbf{v}_{N_2}^{(2)} \in \pi_2. \end{aligned} \quad (11)$$

We use the generalized likelihood ratio to construct a test. The likelihood of the training sample is given by

$$L(\theta^{(1)}, \theta^{(2)} | \mathbf{v}_1^{(1)}, \dots, \mathbf{v}_{N_1}^{(1)}; \mathbf{v}_1^{(2)}, \dots, \mathbf{v}_{N_2}^{(2)}) = \prod_{j=1}^{N_1} f_1(\mathbf{v}_j^{(1)} | \theta^{(1)}) \prod_{k=1}^{N_2} f_2(\mathbf{v}_k^{(2)} | \theta^{(2)}). \quad (12)$$

If the new event to be classified is included with the training sample from π_i , the likelihood is multiplied by $L_i(\theta^{(i)} | \mathbf{v}) = f_i(\mathbf{v} | \theta^{(i)})$. Thus the generalized likelihood ratio for classification is given by

$$\lambda_{\text{class}} = \frac{L_1(\hat{\theta}_0^{(1)} | \mathbf{v}) L(\hat{\theta}_0^{(1)}, \hat{\theta}_0^{(2)} | \mathbf{v}_1^{(1)}, \dots, \mathbf{v}_{N_1}^{(1)}; \mathbf{v}_1^{(2)}, \dots, \mathbf{v}_{N_2}^{(2)})}{L_2(\hat{\theta}_1^{(2)} | \mathbf{v}) L(\hat{\theta}_1^{(1)}, \hat{\theta}_1^{(2)} | \mathbf{v}_1^{(1)}, \dots, \mathbf{v}_{N_1}^{(1)}; \mathbf{v}_1^{(2)}, \dots, \mathbf{v}_{N_2}^{(2)})}, \quad (13)$$

where $\hat{\theta}_0^{(i)}$ is the MLE of $\theta^{(i)}$ under H_0 and $\hat{\theta}_1^{(i)}$ is the MLE of $\theta^{(i)}$ under H_1 , for $i = 1, 2$.

It intuitively follows that small values of λ_{class} provide evidence against H_0 and thus the generalized likelihood ratio test is: reject H_0 if $\lambda_{\text{class}} \leq \lambda_\alpha$, where λ_α is the

critical value such that the test has α significance level. That is, if we know the distribution of λ_{class} given that the null hypothesis is true, $f_{\lambda_{\text{class}}}(\lambda_{\text{class}}|H_0)$, the critical value λ_α is set such that $P[\lambda_{\text{class}} \leq \lambda_\alpha | H_0] = \alpha$. This is the significance level or Type I error rate of the test. It is the probability that H_0 is rejected when it should be accepted.

As for the likelihood ratio for outlier detection, it is difficult in most cases to obtain a closed form expression for the distribution of λ_{class} . Thus, we estimate the distribution using the parametric bootstrap method (Efron, 1979). The algorithm consists of performing the following steps:

1. Compute $\hat{\theta}^{(i)}$ from $\mathbf{v}_1^{(i)}, \dots, \mathbf{v}_{N_i}^{(i)}$, $i = 1, 2$
2. Generate random samples $\mathbf{v}_1^{*(i)}, \dots, \mathbf{v}_{N_i}^{*(i)}$ from $f_i(\mathbf{v}|\hat{\theta}^{(i)})$, $i = 1, 2$
3. Generate random sample \mathbf{v}^* from $f_1(\mathbf{v}|\hat{\theta}^{(1)})$
4. Compute $\lambda_{\text{class}}^* = \lambda_{\text{class}}(\mathbf{v}^*, \mathbf{v}_1^{*(1)}, \dots, \mathbf{v}_{N_1}^{*(1)}, \mathbf{v}_1^{*(2)}, \dots, \mathbf{v}_{N_2}^{*(2)})$ (14)
5. Repeat steps 2 - 4 M times
6. Set λ_α from empirical (bootstrap) distribution of λ_{class}^*
 $\lambda_\alpha^* = \text{int}(\alpha(M+1))$ smallest value of $\{\lambda_{\text{class},j}^*\}_{j=1}^M$.

Using the bootstrap estimate of the critical value provides an approximate α significance level test (MacLachlan, 1987).

The generalized likelihood ratio in Equation (13) can also be used to establish a classification rule based on minimizing the expected cost of misclassification. In this case the cut-off value can be determined by generating bootstrap samples of test events, as in step 3 of the algorithm expressed in (14), from both classes instead of just the first, to estimate both types of probabilities of misclassification $P(2|1) = P[\lambda_{\text{class}} \leq \lambda_{\text{min}} | H_0]$ and $P(1|2) = P[\lambda_{\text{class}} > \lambda_{\text{min}} | H_1]$. Here, instead of setting the critical value so that $P(2|1) = \alpha$ is fixed, the cut-off λ_{min} is set such that the sum $q_1 c(2|1)P(2|1) + q_2 c(1|2)P(1|2)$ is minimized, where $c(2|1)$ and $c(1|2)$ are the costs of the two types of misclassification and q_i denotes the probability ($q_1 + q_2 = 1$) that an event is observed from π_i , $i = 1, 2$.

In practice q_i is estimated by the relative proportion of a particular type of event and the costs of misclassification are specified by the relative importance of each type of error. (Note that only the ratio of the costs is important.) Specifying the relative costs is

somewhat subjective, although qualitatively the cost of misclassifying a nuclear explosion is much higher than the other costs, i.e., $c(EQ|QB) = c(QB|EQ)$ while $c(EQ|EX) \gg c(EX|EQ)$. If the costs are taken to be equal, this procedure is equivalent to minimizing the total error rate. If the probabilities of observing events from the two populations are also approximately equal, this procedure is equivalent to minimizing the total probability of misclassification (see Anderson, 1984, Chapter 6).

For a mixture of correlated discrete and continuous variables, we follow the treatment of Olkin and Tate (1961). If the j th discrete variable Z_j has k_j categories (i.e., possible values), for $j = 1, \dots, r$, then the vector of discrete variables \mathbf{Z} may be expressed as a multinomial random variable $\mathbf{Y} = (Y_1, \dots, Y_k)'$, where $Y_m = 0$ or 1 , $m = 1, \dots, k$, $\sum_{m=1}^k Y_m = 1$, $k = \prod_{j=1}^r k_j$. Each distinct pattern of \mathbf{Z} defines a cell of \mathbf{Y} uniquely. It is assumed that the probability of obtaining an observation in cell m for population π_i is $p_m^{(i)}$, ($0 \leq p_m^{(i)} \leq 1$, $\sum_{m=1}^k p_m^{(i)} = 1$), $i = 1, 2$.

The probability density function of the continuous variables \mathbf{X} is conditional on the categorical variable \mathbf{Y} . That is, the distribution of P_n/S_n is treated as conditional on whether the event was onshore or offshore, for example. Then for the case of a multivariate normal distribution for the continuous variables, the joint distribution for population π_i is $f_i(\mathbf{V}|\theta^{(i)}) = f_{i,Y}(\mathbf{Y}|\theta_Y^{(i)})f_{i,XY}(\mathbf{X}|\theta_{XY}^{(i)}, \mathbf{Y})$, where $\theta_Y^{(i)} = (p_1^{(i)}, \dots, p_{k-1}^{(i)})$ and $\theta_{XY}^{(i)} = (\mu_m^{(i)}, \Sigma^{(i)})$ if \mathbf{Y} belongs to cell m .

2.3.3. Specific Cases

Baek et al. (1993) provide explicit expressions for the likelihood ratio for several cases involving combinations of multinomial and multivariate normal distributions for equal and unequal covariance matrices. Here, to illustrate the method, we present expressions for the portion of the likelihood ratio corresponding to multivariate normal distributions with equal and unequal covariance matrices. Inclusion of an independent Bernoulli variable is analogous to the derivation of Equation (10) and generalizes straightforwardly to a correlated multinomial variable. (See Baek et al., 1993, for explicit expressions for the latter case.)

The likelihood of two training samples from multivariate normal distributions is given by

$$L(\theta^{(1)}, \theta^{(2)} | \mathbf{x}_1^{(1)}, \dots, \mathbf{x}_{N_1}^{(1)}, \mathbf{x}_1^{(2)}, \dots, \mathbf{x}_{N_2}^{(2)}) = \prod_{i=1}^2 \left\{ (2\pi)^p |\Sigma^{(i)}| \right\}^{-N_i/2} \exp \left\{ -\frac{1}{2} \sum_{j=1}^{N_i} (\mathbf{x}_j^{(i)} - \mu^{(i)})' (\Sigma^{(i)})^{-1} (\mathbf{x}_j^{(i)} - \mu^{(i)}) \right\}. \quad (15)$$

The contribution to the likelihood coming from the new event is

$$L_1(\theta^{(i)} | \mathbf{x}) = \left\{ (2\pi)^p |\Sigma^{(i)}| \right\}^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mu^{(i)})' (\Sigma^{(i)})^{-1} (\mathbf{x} - \mu^{(i)}) \right\}. \quad (16)$$

For the case in which the covariance matrices are equal ($\Sigma^{(1)} = \Sigma^{(2)} = \Sigma$), it can be shown that the MLEs under H_0 are given by

$$\begin{aligned} \hat{\mu}_0^{(1)} &= \frac{N_1 \bar{\mathbf{x}}^{(1)} + \mathbf{x}}{N_1 + 1} \\ \hat{\mu}_0^{(2)} &= \bar{\mathbf{x}}^{(2)} \\ \hat{\Sigma}_0 &= \frac{1}{N_1 + N_2 + 1} \left[\mathbf{A} + \frac{N_1}{N_1 + 1} (\mathbf{x} - \bar{\mathbf{x}}^{(1)}) (\mathbf{x} - \bar{\mathbf{x}}^{(1)})' \right], \end{aligned} \quad (17)$$

where $\bar{\mathbf{x}}^{(i)} = \sum_{j=1}^{N_i} \mathbf{x}_j^{(i)} / N_i$ and \mathbf{A} is the pooled expression

$$\mathbf{A} = \sum_{i=1}^2 \sum_{j=1}^{N_i} (\mathbf{x}_j^{(i)} - \bar{\mathbf{x}}^{(i)}) (\mathbf{x}_j^{(i)} - \bar{\mathbf{x}}^{(i)})', \quad (18)$$

and the MLEs under the alternative hypothesis H_1 are given by

$$\begin{aligned} \hat{\mu}_1^{(1)} &= \bar{\mathbf{x}}^{(1)} \\ \hat{\mu}_1^{(2)} &= \frac{N_2 \bar{\mathbf{x}}^{(2)} + \mathbf{x}}{N_2 + 1} \\ \hat{\Sigma}_1 &= \frac{1}{N_1 + N_2 + 1} \left[\mathbf{A} + \frac{N_2}{N_2 + 1} (\mathbf{x} - \bar{\mathbf{x}}^{(2)}) (\mathbf{x} - \bar{\mathbf{x}}^{(2)})' \right]. \end{aligned} \quad (19)$$

Using Equations (15)–(19) and performing some simplifying algebra, the likelihood ratio defined in Equation (13) is given by

$$\lambda_1 = \frac{\left| \hat{\Sigma}_0 \right|^{-(N_1+N_2+1)/2}}{\left| \hat{\Sigma}_1 \right|^{-(N_1+N_2+1)/2}} = \left[\frac{1 + \frac{N_2}{N_1+1} (\mathbf{x} - \bar{\mathbf{x}}^{(2)})' \mathbf{A}^{-1} (\mathbf{x} - \bar{\mathbf{x}}^{(2)})}{1 + \frac{N_1}{N_1+1} (\mathbf{x} - \bar{\mathbf{x}}^{(1)})' \mathbf{A}^{-1} (\mathbf{x} - \bar{\mathbf{x}}^{(1)})} \right]^{(N_1+N_2+1)/2} \quad (20)$$

Note that if $N_2 = 0$, λ_1 is the same as the factor in the outlier likelihood ratio in Equation (10) corresponding to the multivariate normal variables.

For the case in which the covariance matrices are unequal ($\Sigma^{(1)} \neq \Sigma^{(2)}$), it can be shown that the MLEs under H_0 are given by

$$\begin{aligned} \hat{\mu}_0^{(1)} &= \frac{N_1 \bar{\mathbf{x}}^{(1)} + \mathbf{x}}{N_1 + 1} \\ \hat{\mu}_0^{(2)} &= \bar{\mathbf{x}}^{(2)} \\ \hat{\Sigma}_0^{(1)} &= \frac{1}{N_1 + 1} \left[\mathbf{A}^{(1)} + \frac{N_1}{N_1 + 1} (\mathbf{x} - \bar{\mathbf{x}}^{(1)}) (\mathbf{x} - \bar{\mathbf{x}}^{(1)})' \right] \\ \hat{\Sigma}_0^{(2)} &= \frac{1}{N_2} \mathbf{A}^{(2)}, \end{aligned} \quad (21)$$

where

$$\mathbf{A}^{(i)} = \sum_{j=1}^{N_i} (\mathbf{x}_j^{(i)} - \bar{\mathbf{x}}^{(i)}) (\mathbf{x}_j^{(i)} - \bar{\mathbf{x}}^{(i)})' \quad (22)$$

and the MLEs under the alternative hypothesis H_1 are given by

$$\begin{aligned} \hat{\mu}_1^{(1)} &= \bar{\mathbf{x}}^{(1)} \\ \hat{\mu}_1^{(2)} &= \frac{N_2 \bar{\mathbf{x}}^{(2)} + \mathbf{x}}{N_2 + 1} \\ \hat{\Sigma}_1^{(1)} &= \frac{1}{N_1} \mathbf{A}^{(1)} \\ \hat{\Sigma}_1^{(2)} &= \frac{1}{N_2 + 1} \left[\mathbf{A}^{(2)} + \frac{N_2}{N_2 + 1} (\mathbf{x} - \bar{\mathbf{x}}^{(2)}) (\mathbf{x} - \bar{\mathbf{x}}^{(2)})' \right]. \end{aligned} \quad (23)$$

Using Equations (15), (16), (21)-(23) and performing some simplifying algebra, the likelihood ratio defined in Equation (13) in this case is given by

$$\lambda_2 = \frac{|\hat{\Sigma}_0^{(1)}|^{-(N_1+1)/2} |\hat{\Sigma}_0^{(2)}|^{-N_2/2}}{|\hat{\Sigma}_1^{(1)}|^{-N_1/2} |\hat{\Sigma}_1^{(2)}|^{-(N_2+1)/2}} \quad (24)$$

$$= \left[\frac{(N_1+1)^{(N_1+1)/2} N_2^{N_2/2}}{N_1^{N_1/2} (N_2+1)^{(N_2+1)/2}} \right]^p \cdot \frac{\left[1 + \frac{N_2}{N_2+1} (\mathbf{x} - \bar{\mathbf{x}}^{(2)})' (\mathbf{A}^{(2)})^{-1} (\mathbf{x} - \bar{\mathbf{x}}^{(2)}) \right]^{(N_2+1)/2}}{\left[1 + \frac{N_1}{N_1+1} (\mathbf{x} - \bar{\mathbf{x}}^{(1)})' (\mathbf{A}^{(1)})^{-1} (\mathbf{x} - \bar{\mathbf{x}}^{(1)}) \right]^{(N_1+1)/2}}$$

Applications of λ_1 and λ_2 to seismic data are presented in Section 3.

Baek et al. (1993) investigated the performance of the classification test in three cases: normal distributions with equal covariance matrix, normal distributions with unequal covariance matrices, and a mixture of Bernoulli (discrete) and multivariate normal distributions with unequal covariance matrices. Their simulations showed that the bootstrap likelihood ratio statistic competes well with the W (Anderson, 1951) and Z (John, 1963) statistics whose asymptotic distributions are known, for moderate sample sizes $N_1 = N_2 = 50$ of two dimensional data, when two multivariate normal distributions with the equal covariance matrix are considered. They also showed that it performs quite well for both the multivariate normal case with unequal covariance matrices and the case of a mixture of binary and normal variates, where classical classification rules cannot control one of the probabilities of misclassification.

Moreover, the methodology considered here can be applied to any mixture of non-normal discrete and continuous variables, whenever the MLEs exist. It should be noted that the precision of the test depends on the sample sizes N_1 and N_2 , and the bootstrap replication size M . Small sample sizes may result in MLEs for the parametric bootstrap which are not close to the true parameter values. Adequate sample sizes for different dimensions of the classification variable should be studied more fully.

Last, to illustrate the usefulness of the bootstrap generalized likelihood ratio method for classifying seismic events, they applied it to observations at the ARCESS array in Norway which consist of mining blasts from two separate mines (HB6 and HD9) located in the Kola Peninsula of the former Soviet Union. Fifteen blasts were observed from mine HB6 and sixteen blasts were observed from mine HD9. They showed that the discrete variable day-of-week can contribute noticeably to the power of the test.

2.4. Preliminary Training Set Analyses

2.4.1. Tests for Normality

Since the statistics of most discriminants are in general quite complicated and the validity of discriminant analysis results depend on their proper treatment, we apply several tests to training sets to determine whether necessary assumptions hold. Although the Bootstrap GLR methods do not require that the continuous features are normally distributed, implementation of the tests simplify greatly if this assumption is made. In most cases, however, the continuous discriminants do not obey normal statistics.

To remedy this situation we first apply two hypothesis tests to determine whether the data are normal. Dyer (1974) listed seven tests for normality of which the Wilk-Shapiro (Shapiro and Wilk, 1965) and Anderson-Darling (Anderson and Darling, 1954) tests are generally the most powerful for a wide range of alternative distributions. The Wilk-Shapiro test is powerful for detecting skewed distributions and the Anderson-Darling test is powerful for detecting long-tailed distributions (Pettitt, 1977). In our analysis both tests are applied to each feature for all classes of events separately.

The Wilk-Shapiro W statistic is given by

$$W = \left(\sum_{i=1}^n a_i^{(n)} x_{(i)} \right)^2 / \sum_{i=1}^n (x_i - \bar{x})^2, \quad (25)$$

where $x_{(1)} \leq \dots \leq x_{(n)}$ are the ordered observations. The coefficients $a_i^{(n)}$ and percentage points of W are computed numerically and have been tabulated by Shapiro and Wilk (1965) and Seber (1984). Small values of W indicate departures from normality.

The Anderson-Darling A_n^2 statistic is given by

$$A_n^2 = -\frac{1}{n} \left[\sum_{i=1}^n (2i-1) \{ \log z_i + \log(1 - z_{n+1-i}) \} - n \right], \quad (26)$$

where

$$z_i = \Phi \left(\frac{x_{(i)} - \bar{x}}{s} \right), \quad (27)$$

$s^2 = \sum_i (x_i - \bar{x})^2 / (n-1)$ and Φ is the distribution function for $N(0,1)$. Normality is rejected for large values of A_n^2 . For $n \geq 4$ the critical value of the test is given approximately by $a_n = a_\infty (1 + c_1 n^{-1} + c_2 n^{-2})$, where a_∞ , c_1 and c_2 have been tabulated by Pettitt (1977) for various significance levels.

2.4.2. Box-Cox Transformations

If the normality is rejected by either test at fixed significance level, the feature is then transformed using the Box-Cox transformation (Box and Cox, 1964). Special cases of the Box-Cox transformation are the square root and the logarithm. The procedure automatically finds a power-law or logarithm transformation that maximizes the normal likelihood function of the transformed data. In general the Box-Cox transformation takes the form

$$x^{(\gamma)} = \begin{cases} \frac{x^\gamma - 1}{\gamma}, & \gamma \neq 0 \\ \log(x), & \gamma = 0 \text{ and } x > 0, \end{cases} \quad (28)$$

where x is the original discriminant variable and $x^{(\gamma)}$ is the transformed variable. (Note that this set of transformations is continuous in the limit that $\gamma \rightarrow 0$. Also the restriction of $x > 0$ can be circumvented by uniformly translating all values of a given feature for all classes so that they are all positive. The classification and outlier tests are invariant under uniform shifts.) The transformed values are inserted into the likelihood function of a normal distribution with the unknown mean and variance replaced by their maximum likelihood estimates. This maximizes the likelihood function with respect to the unknown parameters of the normal distribution given the transformed data, yielding

$$L_{\max}(\gamma) = -\frac{n}{2} \log \left[\sum_{j=1}^n (x_j^{(\gamma)} - \bar{x}^{(\gamma)})^2 \right] + (\gamma - 1) \sum_{j=1}^n \log x_j, \quad (29)$$

where n is the number of samples. Given the data, $L_{\max}(\gamma)$ is a function of the single variable γ . The value of γ that maximizes $L_{\max}(\gamma)$ is used as the transformation coefficient.

Transforming a feature for more than one class of events creates a complication. If a discriminant is transformed for each class separately, it is not at all obvious how a

new event of unknown type should be transformed, particularly for classical discriminant analysis approaches. Under a hypothesis testing approach to discriminant analysis, there are two options: (i) the event being tested could be transformed according to the class that corresponds to either the null or alternative hypotheses corresponding to the numerator or denominator of the likelihood ratio or (ii) a compromise Box-Cox transformation could be sought that makes the discriminant approximately normal for all classes. Since in some cases we may wish to consider a classical discriminant analysis approach in which the cost of misclassification is minimized, we will adopt the latter procedure. In this procedure we maximize the product of normal likelihood functions for all classes with respect to a single transformation coefficient.

This transformation is applied to each discriminant separately. Once transformed, each discriminant is tested again for normality. Although these transformations do not ensure normality in every case, we found that the overwhelming majority of features are accepted as normal after applying the transformations.

2.4.3. Tests for Equal Covariance Matrices

Robust application of classification tests also depend on whether covariance matrices for the features are equal for the various event types. If sample sizes are small or covariance matrices are not significantly different, using a classification method based on a pooled estimate of a single covariance matrix is more robust. If the covariance matrices are significantly different, estimating separate covariance matrices leads to a classification test with greater power (e.g., Seber, 1984).

We have implemented and applied a hypothesis test based on the F-distribution to determine whether the individual variances are equal. If any of the variances are unequal the corresponding covariance matrices are also unequal. However, since equal variances do not necessarily imply equal covariance matrices, we have also implemented a procedure to test for equal covariance matrices. A Monte Carlo comparison of several tests for equality of covariance matrices was given by Pervaiz and Skinner (1990). Based on their results, we selected a standard-error test provided by Layard (1974) which has comparable power to other tests under the assumption of normality, while being more robust to departures from normality.

Note that unless the sample sizes are relatively large, tests for equal covariance matrices typically have less power to reject cases in which they are unequal than testing

the variances separately and are usually more computationally intensive. Thus we use the standard error test for equal covariance matrices only if all variances are first accepted as equal by the F-test or if the sample sizes are relatively large (~100).

2.4.4. Treatment of Missing Data Values

Another practical statistical issue that arises is that of missing data values. Often particular seismic phases are censored due to poor signal-to-noise or interference with other events, which prevents them from being used in discriminants such as Pn/Lg. It is nevertheless useful to include such events with partial information in the training sets, rather than discarding them altogether. One approach is mean replacement which simply replaces a missing value of a particular feature with the mean computed for that feature from the remaining events. This is a simple procedure but it does not utilize available information regarding feature correlations. An approach that does so is called the EM (Estimation-Maximization) algorithm (Dempster et al., 1977). In the EM algorithm the sufficient statistics (first and second moments) of the missing values are estimated by regression analysis, based on MLEs computed from available data. The MLEs are then updated using the available data and the estimates of the sufficient statistics for the missing values. The procedure is iterated until the MLEs converge to within a specified tolerance. This approach has also been used by Jih et al. (1990) to improve estimates of device yields using regression analysis of partially censored data.

2.4.5. Feature Selection

In many cases, the number of available discriminants can be large (cf. Section 3.1) and it is not clear from the outset which are the best to use in a given situation. Thus, we have implemented and applied statistical procedures to select the optimal subset of discriminants based on training data. Features are first ranked based on error rates estimated using training data and the GLR classification algorithm for each univariate case. The leave-one-out technique is used to estimate the error rates. A stepforward selection procedure (Seber, 1984) is then used on the best univariate features to determine which set provides the best multivariate discrimination. The stepforward procedure starts with the best univariate feature and adds successive features as long as there is a continued reduction in the overall error rate estimate.

Applications of these preliminary training set analyses to seismic data sets are described in Sections 3.3 and 4.3.

3. MONITORING APPLICATIONS

3.1. Data Sets

Feature data used here were obtained from seismic waveform analysis performed by Baumgardt (1992a) using the Intelligent Seismic Event Identification System (ISEIS) (Baumgardt et al., 1991). We used two data sets in this study. The first consists of 10 earthquakes and 13 quarry blasts in the Vogtland region of western Bohemia (straddling Germany and the Czech Republic), which were recorded at GERESS. The second data set consists of 53 quarry blasts on the Kola Peninsula and 25 earthquakes near Steigen, Norway, recorded at ARCESS. The Vogtland and Steigen events are included in the CSS Ground-Truth Database and are discussed in more detail by Grant et al. (1993).

Event locations are shown in Figure 3; squares represent quarry blasts and circles represent earthquakes. Locations of the ARCESS, GERESS and NORESS arrays are also shown. The Vogtland region is roughly 180 km northwest of the GERESS array. Only one Vogtland earthquake was also recorded at NORESS. The Steigen events occurred roughly 440 km southwest of the ARCESS array. Baumgardt et al. (1992) argue that although the Kola mining blasts were not in the same region as the Steigen earthquakes, they are located at comparable distances from ARCESS and the paths are very similar in the same shield-type geology. Thus we consider the Kola Peninsula and Steigen events as a single Kola/Steigen data set. Some, but not all, of these events were also recorded at NORESS and FINESA. Hence we concentrated on the ARCESS data for these events.

ISEIS feature data include amplitude ratios, Pg/Lg , Pg/Sn , Pn/Lg , Pn/Pg , Pn/Sn , Rg/Lg , Rg/Sn , and Sn/Lg of either maximum or rms amplitude in 9 frequency bands (0.5-2.5, 2-4, 2.5-4.5, 3-5, 4-6, 5-7, 6-8, 8-10, and 8-16 Hz), which we denote $Pn/Sn(\max; 8-16 \text{ Hz})$, for example, spectral ratios of maximum and rms Lg , Pg , Pn , Rg and Sn in 3 different combinations of frequency bands (2-4 Hz/4-6 Hz, 2-5 Hz/5-8 Hz, 2-6 Hz/6-10 Hz), and the variance, skewness and kurtosis of detrended spectra and cepstra. (For further details on these features and how they were computed, see Baumgardt, 1992b; Baumgardt et al., 1992.) Not all of the features were available for every event and the total number was reduced from 204 to 5 at most for applications described below.

Scatter plots of Pn/Lg and Pn/Sn ratios of maximum amplitude are shown in Figure 4. Plots of Pn/Lg are shown on the left and plots of Pn/Sn are shown on the right.

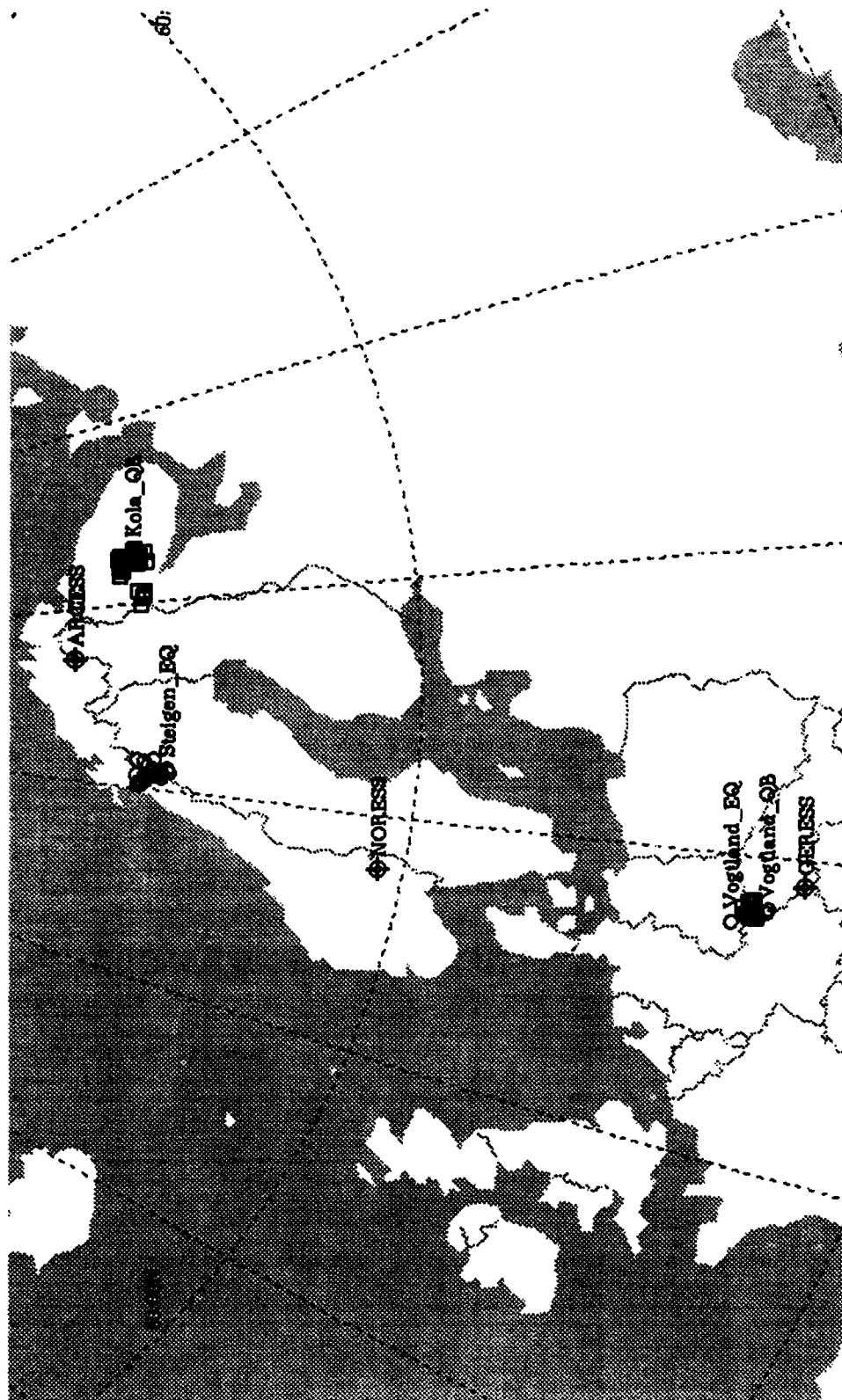


Figure 3. Locations of 10 earthquakes and 13 quarry blasts in the Vogland region, recorded at GERESS, and 25 earthquakes in the Steigen region and 53 quarry blasts in the Kola Peninsula, recorded at ARCESS.

The upper plots correspond to the Vogtland events recorded by GEC2 and the lower plots correspond to the Kola/Steigen events recorded by ARA0. The ratios are plotted for 9 frequency bands listed along the bottom frames. It is clear that high frequency Pn/Lg and Pn/Sn ratios are effective discriminants for the Vogtland region, particularly in the 8-16 Hz band. Note that the values of the amplitude ratios are very similar for all 9 bands for the earthquakes while there is considerable variation for the quarry blasts. The variance within each band is also larger for the quarry blasts.

For the Kola/Steigen events none of the frequency bands exhibit clear separation of the two event types. As for the Vogtland events, the amplitude ratios for the Kola quarry blasts have greater variation than the Steigen earthquakes. Baumgardt et al. (1992) noted that the overlap between the ratios for the two event types indicates that some of the blasts excite considerable shear wave energy. Unlike amplitude ratios for the Vogtland blasts, ratios for the Kola blasts do not continue to increase monotonically with frequency. Rather they tend to peak between 4 to 8 Hz. Thus, we found that Pn/Sn(max; 4-6 Hz), Pn/Sn(max; 5-7 Hz) and Pn/Lg(max; 5-7 Hz) discriminate the best in this region.

Spectral ratios of maximum and rms Lg in 3 frequency band combinations were also provided by Baumgardt (1992a) for the Vogtland events and are shown in Figure 5. Except for one quarry blast the two event types separate completely. Baumgardt et al. (1992) showed that this event was ripple-fired, causing large modulations in the spectrum which enhance the high frequencies. In contrast to the Vogtland events, Baumgardt et al. (1992) found that there was almost complete overlap of the Lg spectral ratios for the Kola/Steigen earthquakes and quarry blasts. We also found that there is nearly total overlap of the two populations using variance, skewness and kurtosis of either the spectra or cepstra. Thus we do not consider them in our analysis.

3.2. Multivariate Graphical Displays: Andrews' Plots

In addition to scatter plots we have also generated Andrews' Fourier plots as a means to visualize how multivariate characteristics of events compare as a whole and to compare many events simultaneously. Andrews (1972) suggested a useful technique to represent multivariate data as a 2-dimensional plot using a Fourier series expansion. For each event the coefficients of the series are the discriminant values that have been standardized so that all discriminants have the same overall mean and standard deviation computed over all of the events; this prevents a single discriminant from completely dominating the plot.

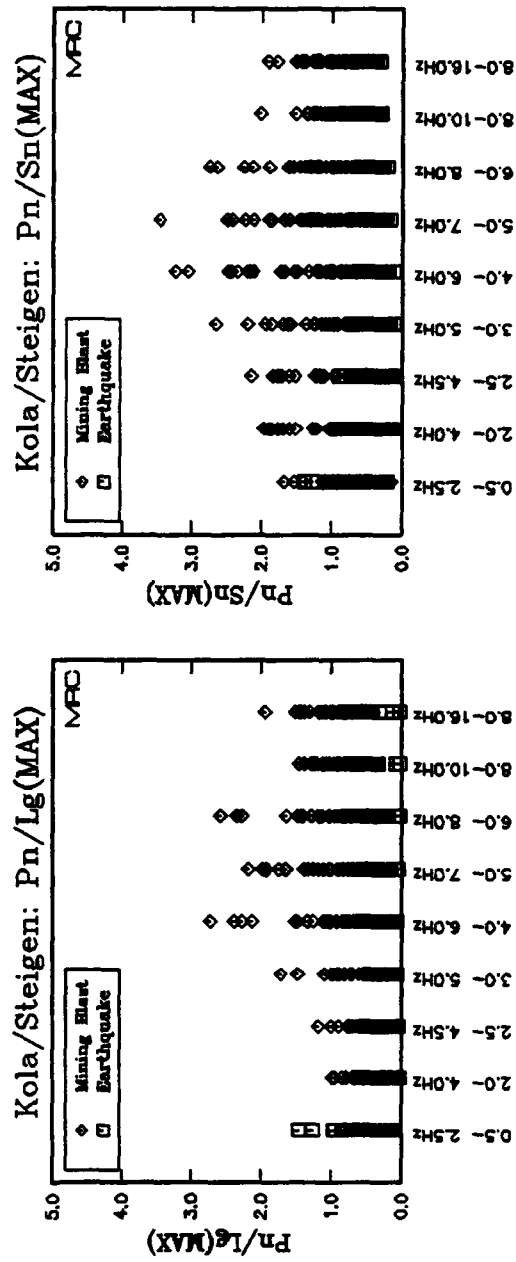
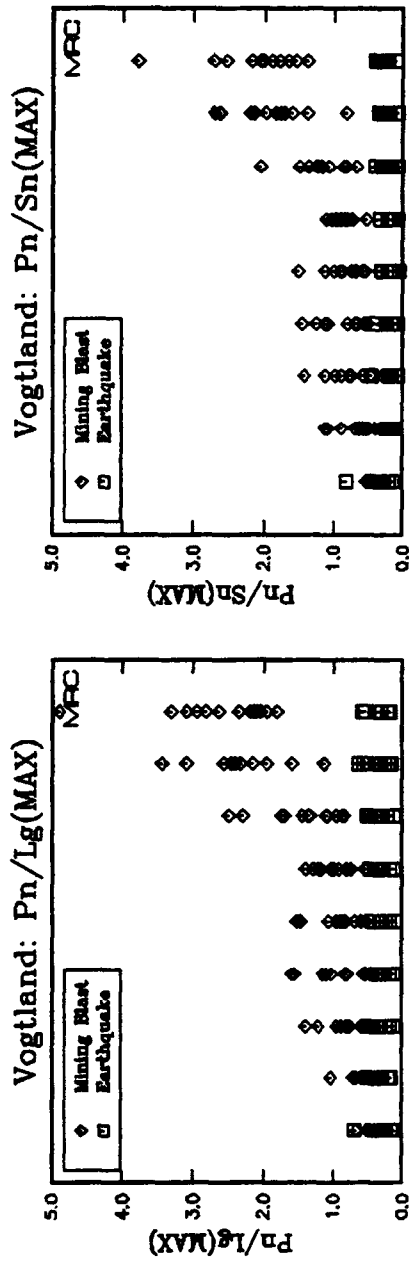


Figure 4. Scatter plots of Pn/Lg (left) and Pn/Sn (right) ratios of maximum amplitude in 9 frequency bands for the Vogtland events (upper) recorded by GEC2 and the Kola/Steigen events (lower) recorded by ARA0.

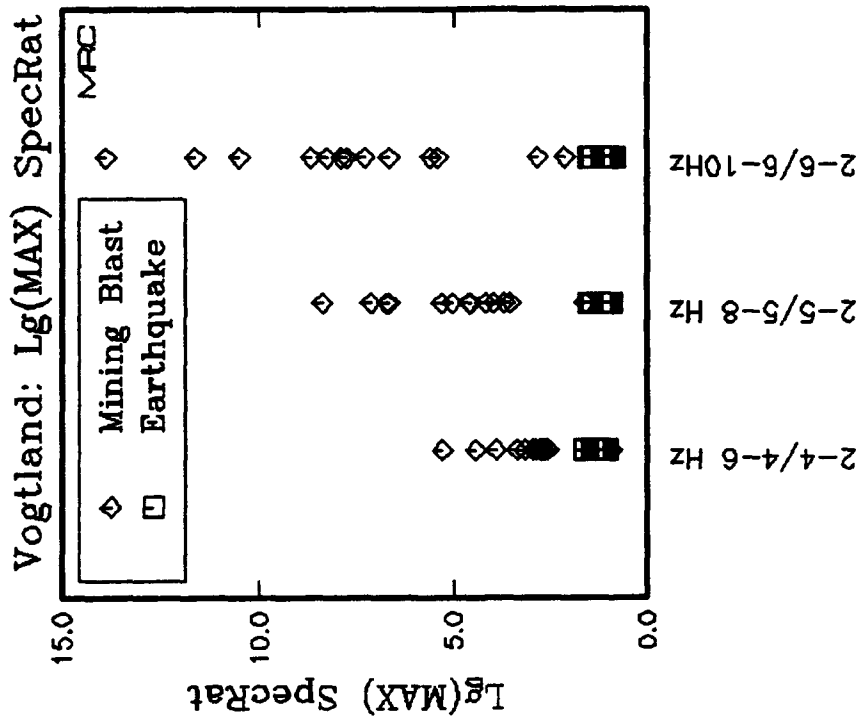
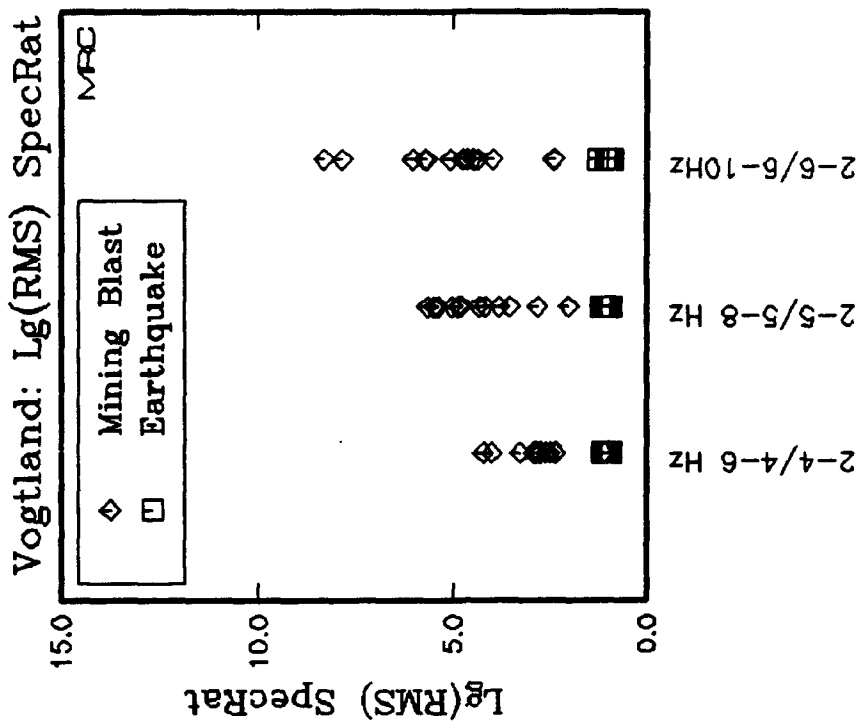


Figure 5. Scatter plots of spectral ratios of maximum (left plot) and rms (right plot) Lg for 3 frequency band combinations for the Vogtland data set.

Andrews' plots for the Vogtland events are shown in Figure 6. The discriminants used as coefficients of the Fourier series are listed in the upper lefthand legend. The relative shift upward of the curves for the blasts is due to greater values of $Pn/Lg(\max; 8-16\text{Hz})$. Larger values of $Pn/Sn(\max; 8-16\text{Hz})$ and Lg spectral ratios for the blasts also lead to greater amplitudes for the first and second harmonics. From these plots it is easy to discern quickly that the earthquakes and blasts are considerably different based on these discriminants.

Andrews' plots for the Kola and Steigen events are shown in Figure 7. Again the discriminants are listed in the upper lefthand legend. These plots show that the majority of blasts are considerably different than the earthquakes, based on the amplitude ratios used, although there is some overlap between the two populations for all three discriminants. It also shows that there is much greater variation in the discriminants for quarry blasts than for earthquakes.

3.3. Results of Preliminary Training Set Analyses

An example of the preliminary training set normality analysis is shown in Figure 8. The histogram plots are of $Pn/Sn(\max; 4-6\text{ Hz})$ values for the Steigen earthquakes (left) and Kola blasts (right). The lower frames were produced before the feature values were transformed and the upper frames correspond to the values after applying the Box-Cox transformation. The legends in each frame give the results of the tests for normality, performed at 0.05 significance level. Before transforming the data, the distributions for both earthquakes and blasts are skewed and, hence, normality was rejected. The transformed variable is not skewed and normality is now accepted. Quantile-Quantile (Q-Q) plots for the same case are shown in Figure 9. Q-Q plots are useful for visualizing whether data obey normal statistics. If the data are approximately normal, the values of the sample quantiles computed from the data (square markers) should correspond to the quantiles of a normal distribution represented by the straight line. Figure 9 shows that the transformed data are normal to a good approximation.

Similar results for Kola and Steigen events are shown in Figures 10 and 11 for $Pn/Sn(\max; 5-7\text{ Hz})$ and in Figures 12 and 13 for $Pn/Lg(\max; 5-7\text{ Hz})$. Normality of the Box-Cox transformed $Pn/Sn(\max; 5-7\text{ Hz})$ values is accepted for both classes, but it is rejected for the transformed $Pn/Lg(\max; 5-7\text{ Hz})$ values. Note that normality of the untransformed $Pn/Lg(\max; 5-7\text{ Hz})$ values is accepted for the Steigen earthquakes but not for the Kola blasts. In attempting to find a compromise transformation for both classes,

Vogtland DataSet (GEC2)

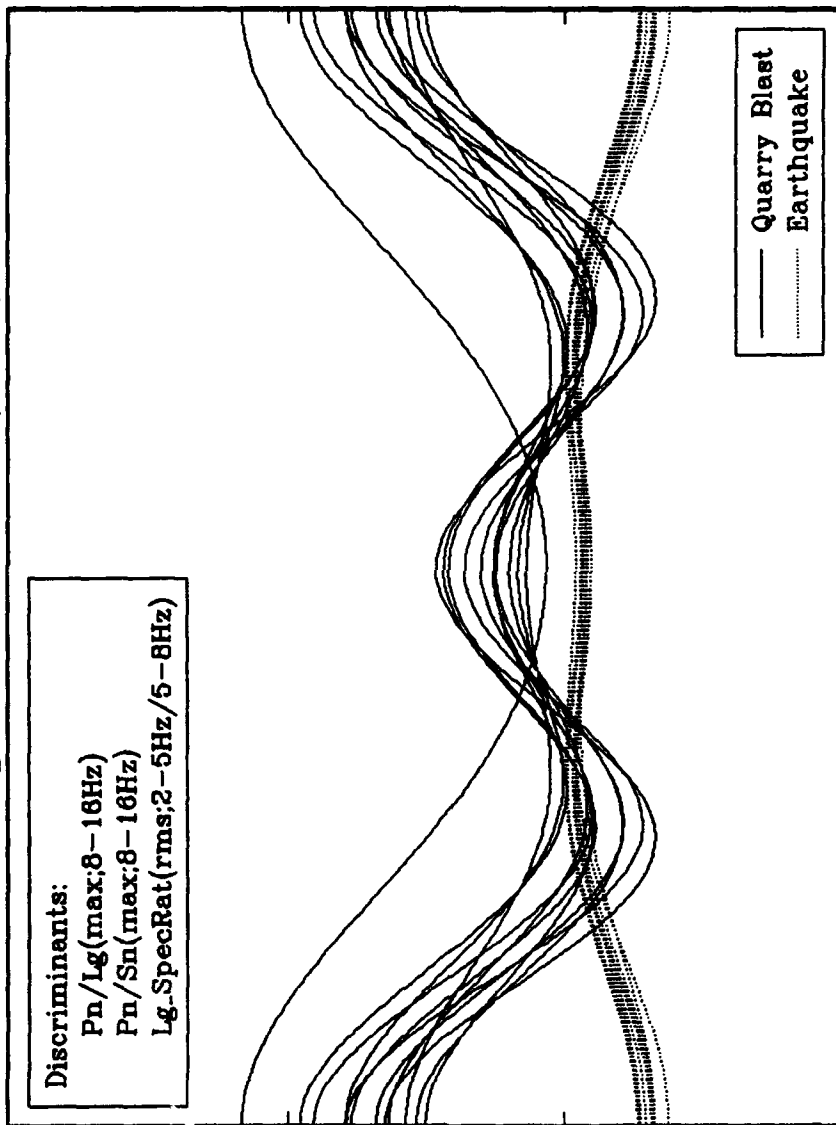


Figure 6. Andrew's plots for 10 earthquakes and 13 quarry blasts in the Vogtland region, using the discriminants listed in the upper legend as coefficients of the Fourier series. All three discriminants are larger for the blasts than for the earthquakes.

Kola/Steigen DataSet (ARAO)

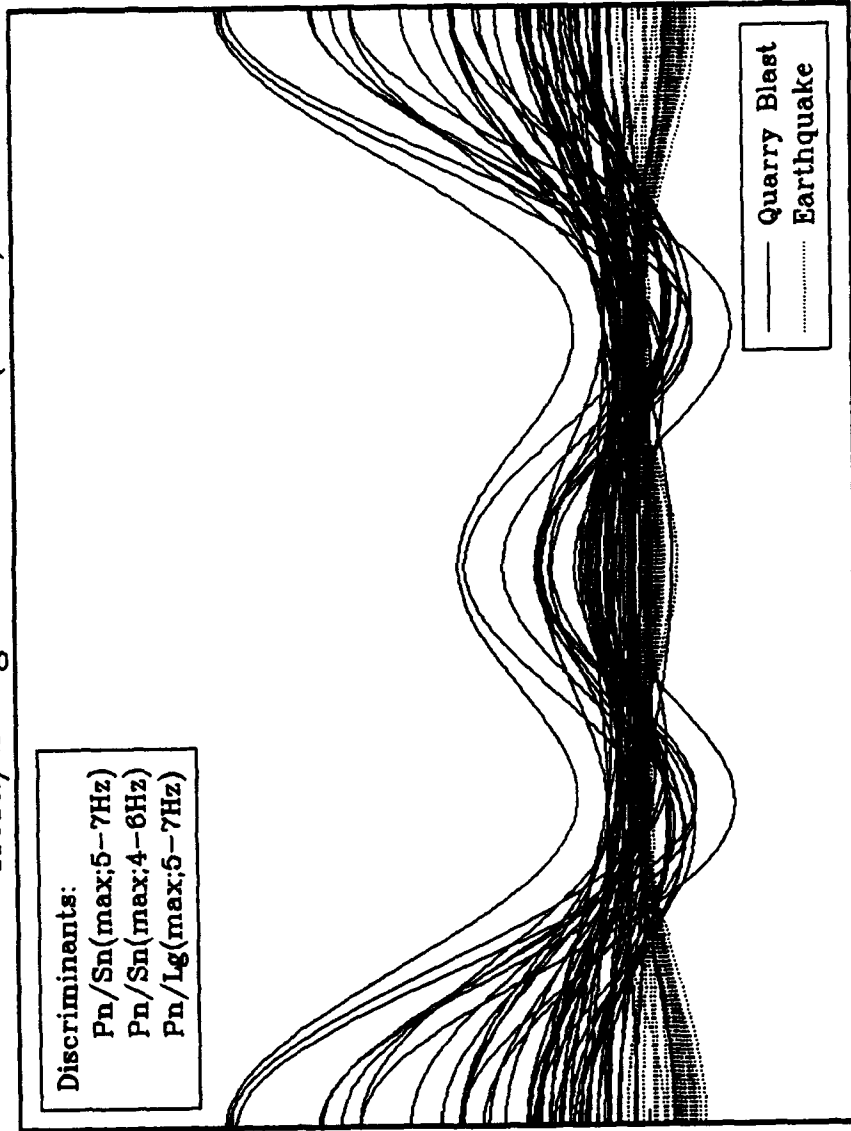


Figure 7. Andrew's plots for 25 Steigen earthquakes and 53 Kola blasts, using the discriminants listed in the upper legend as coefficients of the Fourier series. All three discriminants are typically much larger for the blasts than for the earthquakes.

Histograms: Kola/Steigen DataSet: Pn/Sn(max;4-6Hz)

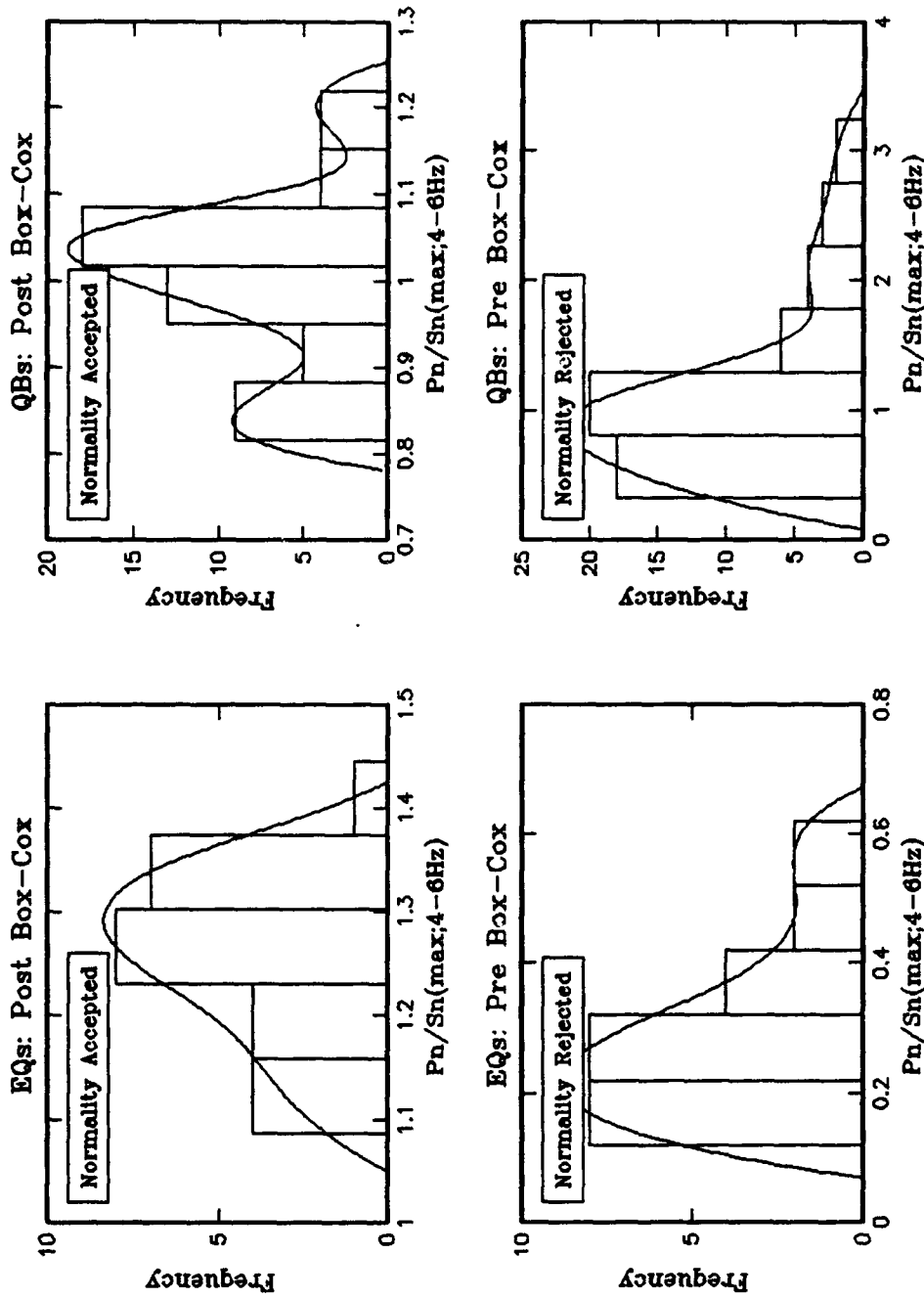


Figure 8. Histogram plots of Pn/Sn(max; 4-6 Hz) for Steigen earthquakes (left) and Kola quarry blasts (right), before (lower) and after (upper) applying the Box-Cox transformation. The legends in each frame contain the results of the tests for normality.

Q-Q Plots: Kola/Steigen Dataset: Pn/Sn(max;4-6Hz)

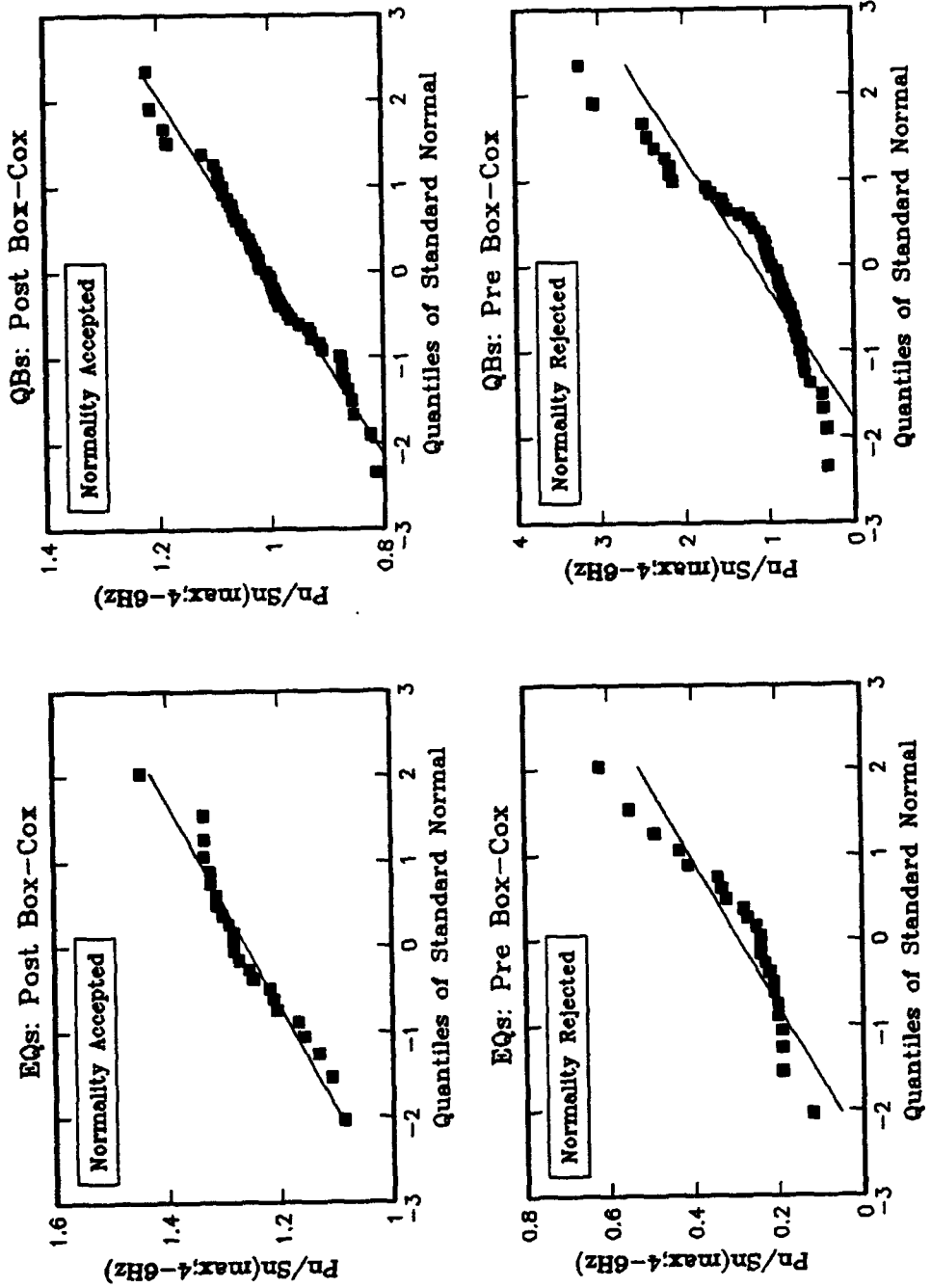


Figure 9. Quantile-Quantile plots of Pn/Sn(max; 4-6 Hz) for Steigen earthquakes (left) and Kola quarry blasts (right), before (lower) and after (upper) applying the Box-Cox transformation. The legends in each frame contain the results of the tests for normality. Data sets that are approximately normal lie closely along the straight line.

Histograms: Kola/Steigen DataSet: Pn/Sn(max;5-7Hz)

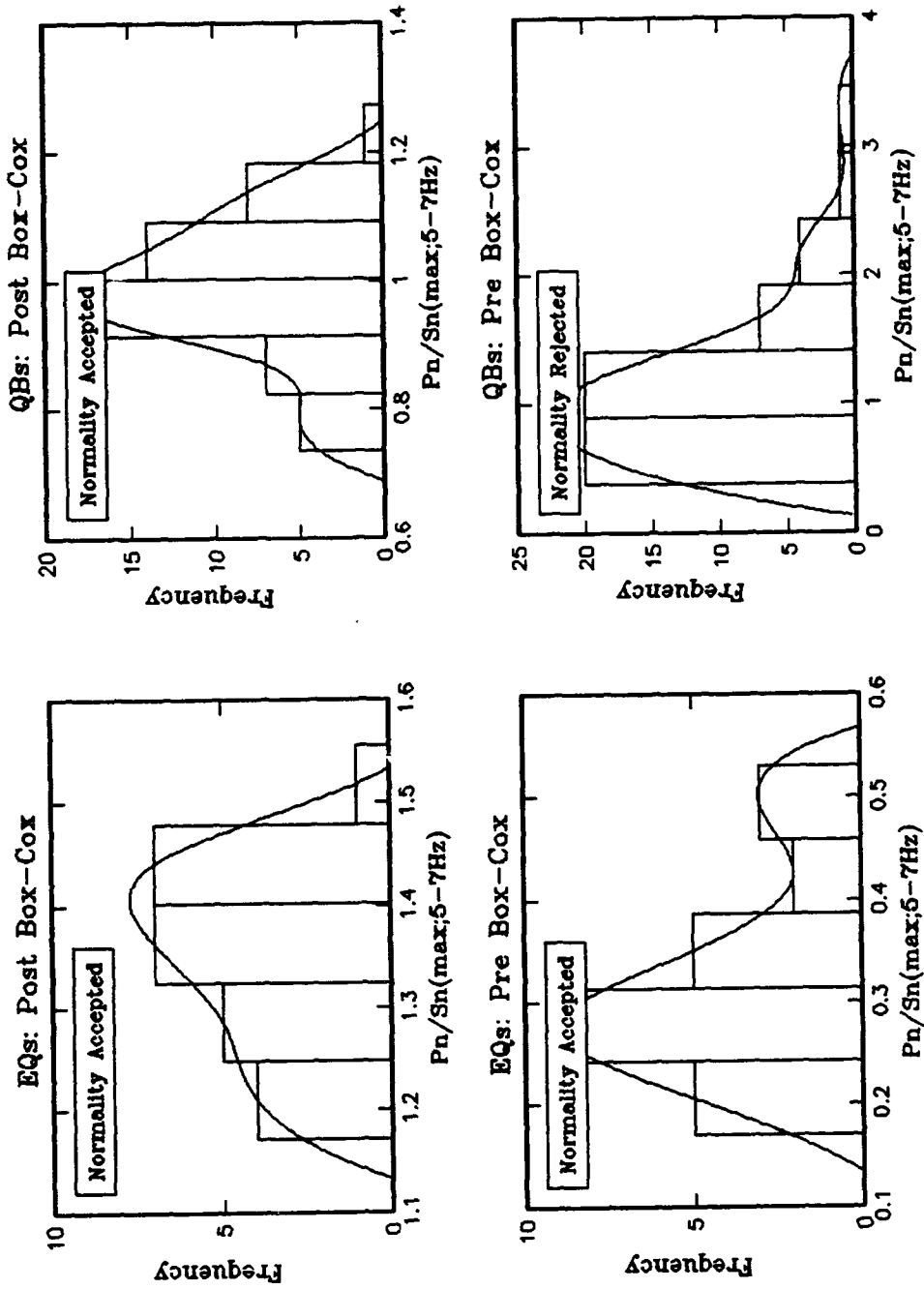


Figure 10. Same as Figure 8 but for Pn/Sn(max; 5-7 Hz).

Q-Q Plots: Kola/Steigen DataSet: Pn/Sn(max;5-7Hz)

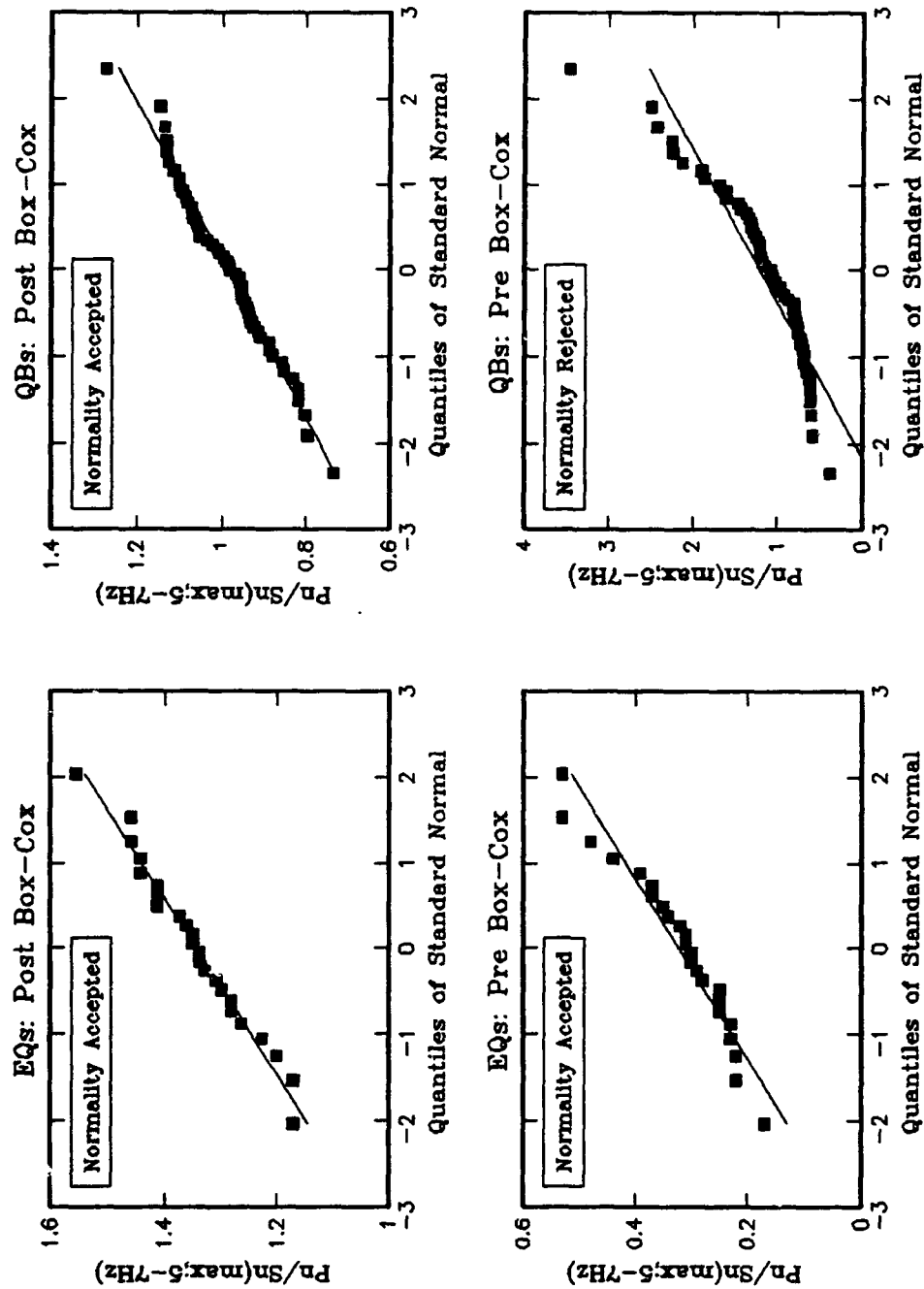


Figure 11. Same as Figure 9 but for Pn/Sn(max; 5-7 Hz).

Histograms: Kola/Steigen DataSet: Pn/Lg(max;5-7Hz)

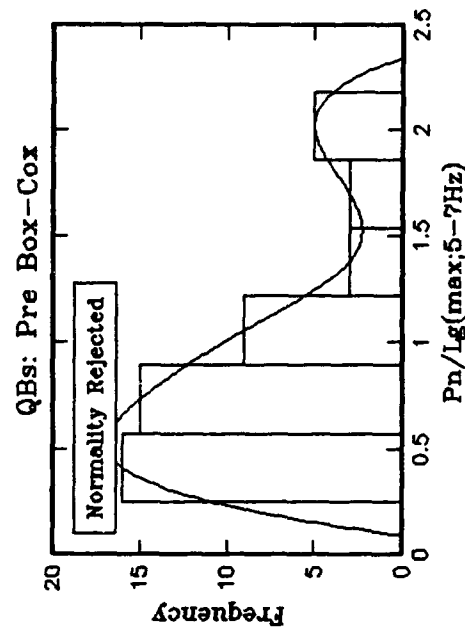
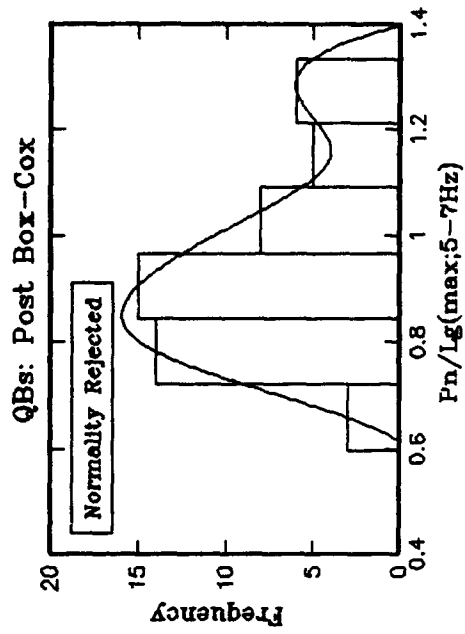
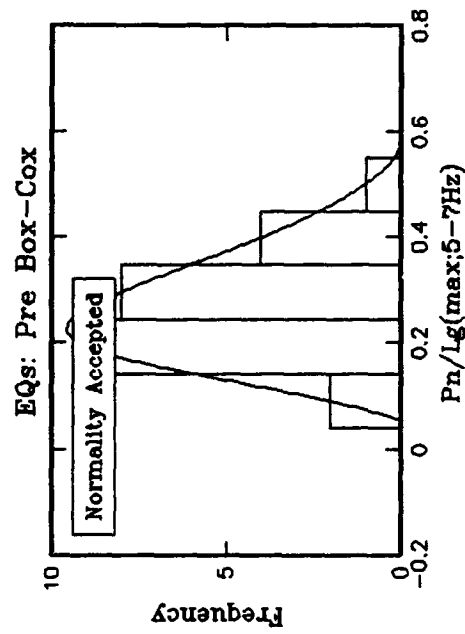
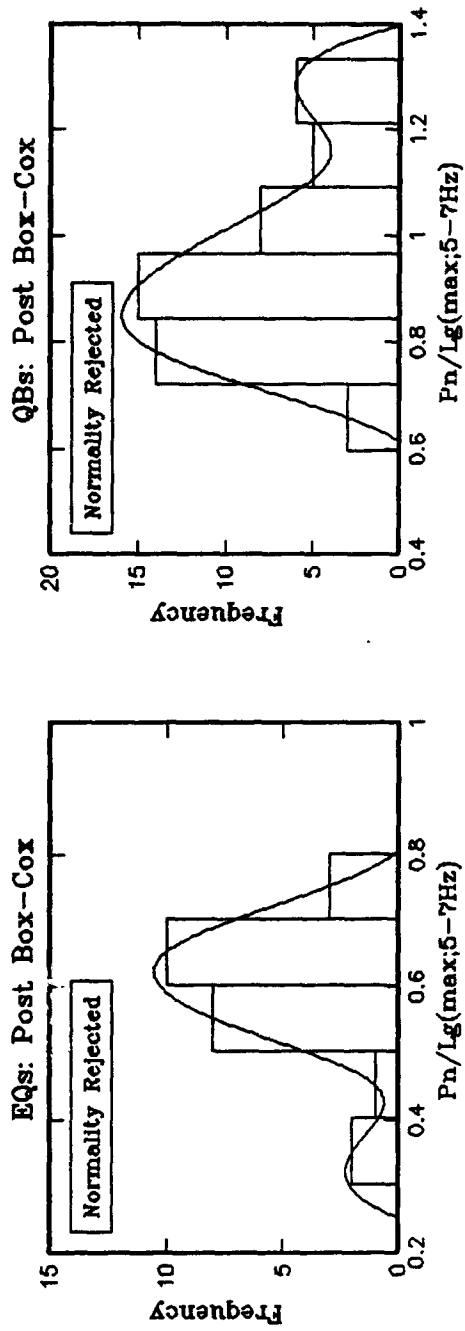


Figure 12. Same as Figure 8 but for Pn/Lg(max; 5-7 Hz).

Q-Q Plots: Kola/Steigen Dataset: Pn/Lg(max;5-7Hz)

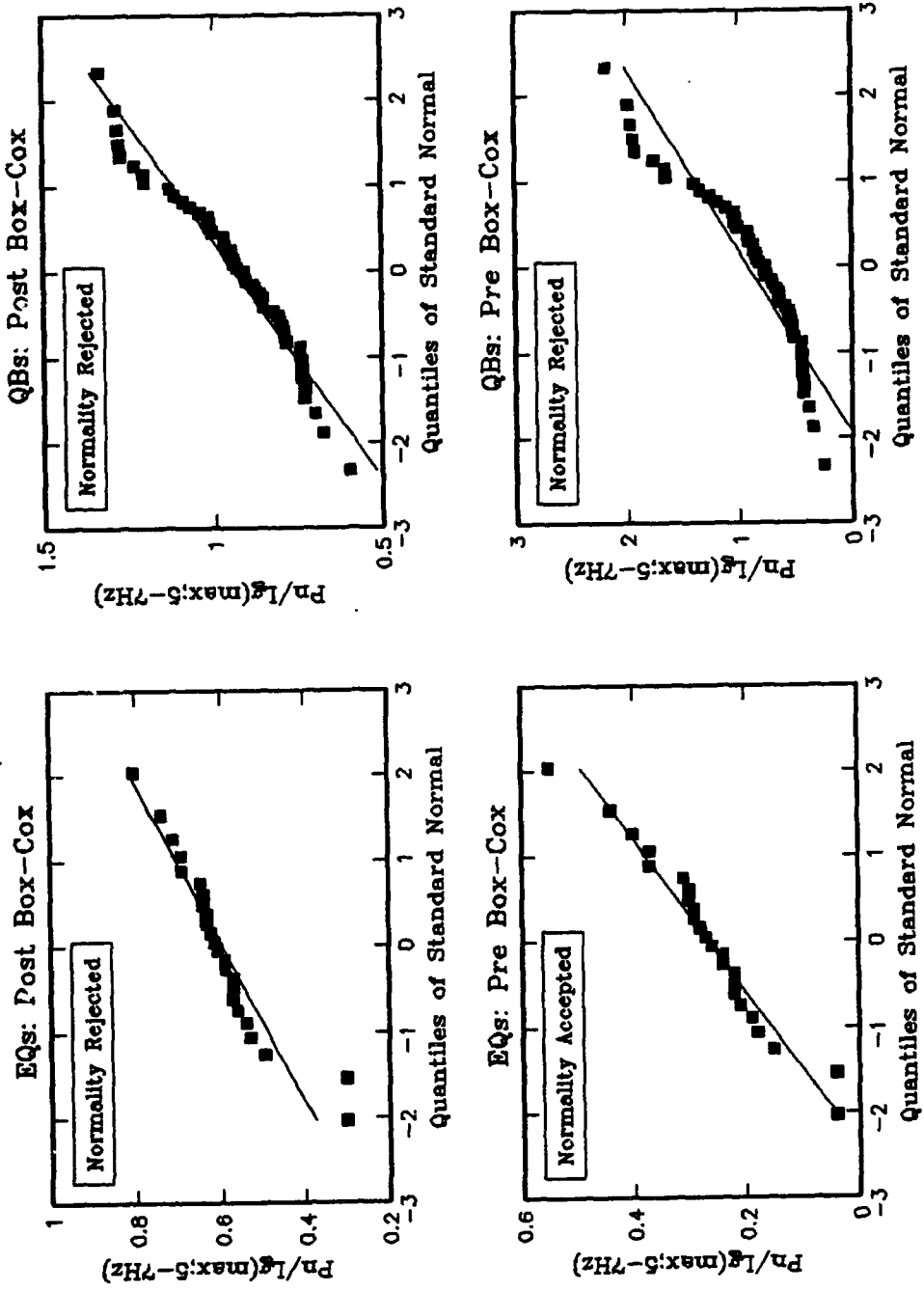


Figure 13. Same as Figure 9 but for Pn/Lg(max; 5-7 Hz).

the transformed Pn/Lg(max: 5-7 Hz) values for the Kola blasts are more nearly normal than before but the values for the Steigen earthquakes are less normal (Figure 13). This illustrates the fact that the joint Box-Cox transformation does not guarantee normality. (Note that the assumption of normality is accepted for this discriminant if the significance level is set at 0.025.) However, 72 of the 108 ISEIS features we received for the Kola and Steigen events and 116 of 144 for the Vogtland events were found to be approximately normal for both classes after applying the Box-Cox transformation.

Figures 14 and 15 show similar Q-Q plots of Pn/Lg(max: 8-16 Hz) and Pn/Sn(max: 8-16 Hz), respectively, for the Vogtland events. Normality is accepted for both discriminants in both classes after applying the Box-Cox transformation. Normality was accepted by both tests for the rms Lg spectral ratio values for the ratio of the 2-5 Hz to 5-8 Hz bands prior to applying the Box-Cox transformation. Hence, no transformation was made and a similar plot for this discriminant was not generated.

Note that in many cases the earthquake features are approximately normal even before transforming the data, while this is rarely the case for mining blast features which have much more skewed distributions with longer tails at the high end of the amplitude and spectral ratios.

The F-test was also applied to the features and the hypothesis of equal variances was rejected in most cases. Based on our ranking procedure we found that Pn/Sn(max; 4-6 Hz), Pn/Sn(max; 5-7 Hz) and Pn/Lg(max; 5-7 Hz) discriminate the best in the combined Kola/Steigen region, while Pn/Lg(max; 8-16 Hz), Pn/Sn(max; 8-16 Hz) and the spectral ratio Lg(rms; 2-5 Hz/5-8 Hz) discriminate the best in the Vogtland region. Equality of the covariance matrices for these sets of features was rejected in both cases and, hence, were treated as unequal in the GLR classification tests discussed in the following section.

3.4. Monitoring a New Region

Suppose that under a CTBT we need to monitor a new region for which a seismic array has been installed recently and to date has observed only a few of events, e.g., earthquakes. As new events are observed, how well does the outlier test perform for detecting explosions? If explosion training data also become available, how much more accurately can we identify events in a particular region using the classification test?

Q-Q Plots: Vogtland DataSet: Pn/Lg(max;8-16Hz)

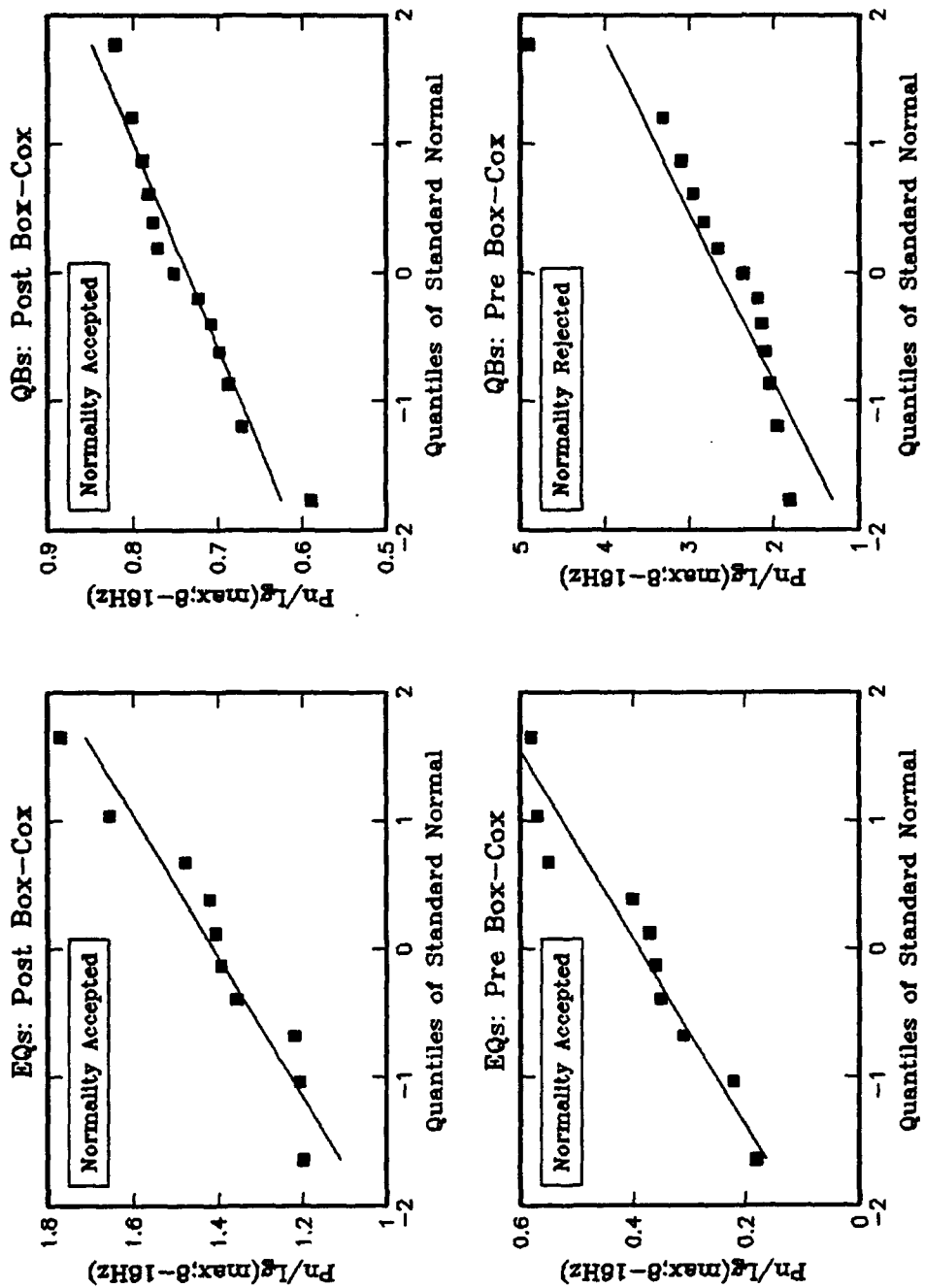


Figure 14. Quantile-Quantile plots of Pn/Lg(max; 8-16 Hz) for Vogtland earthquakes (left) and quarry blasts (right), before (lower) and after (upper) applying the Box-Cox transformation. The legends in each frame contain the results of the tests for normality. Data sets that are approximately normal lie closely along the straight line.

Q-Q Plots: Vogtland DataSet: Pn/Sn(max;8-16Hz)

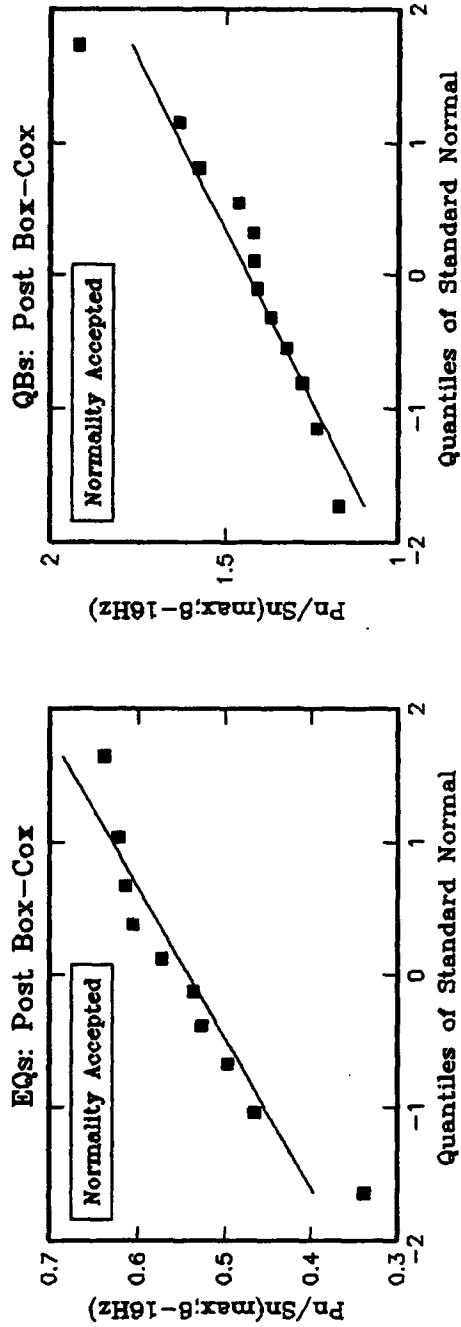


Figure 15. Same as Figure 14 but for Pn/Sn(max; 8-16 Hz).

For the purpose of illustration, consider the hypothetical case in which we are monitoring the regions surrounding the GERESS and ARCESS arrays for mining blasts. Assuming that we only have several earthquakes to start with in each region and treating each blast as a new observation, the GLR outlier test was applied to determine whether we can detect these events as being different. Each earthquake was also tested as a new event using the leave-one-out procedure. This allows us to estimate the false alarm rate and, in practice, to determine if there are any suspicious events included in the training data. In the following the significance level was set at 0.025.

Using the 10 Vogtland earthquakes and Pn/Sn(max; 8-16 Hz), Pn/Lg(max; 8-16 Hz) and Lg(rms; 2-5 Hz/5-8 Hz) as discriminants, we found that all 13 Vogtland quarry blasts were detected as outliers while, appropriately, none of the earthquakes treated as new events were. Using the 25 Steigen earthquakes and Pn/Sn(max; 4-6 Hz), Pn/Sn(max; 5-7 Hz) and Pn/Lg(max; 5-7 Hz) as discriminants, 51 of 53 (96.2%) Kola blasts were detected as outliers and 21 of 24 (87.5%) Steigen earthquakes were identified correctly.

The fact that 3 of 24 earthquakes were called outliers when they should not have been is not particularly detrimental since these events could be investigated further using other, possibly non-seismic, techniques. A much more significant and positive result is that a very high percentage (100% and 96.2%) of the events we were trying to monitor were in fact identified as suspicious without requiring training data for this type of event.

In the second part of this analysis, the GLR classification test was trained with both earthquake and quarry blast data to assess how error rates can be reduced if training data are available for all relevant event types. Each event was classified using the leave-one-out procedure. As for the outlier test, all Vogtland events were classified correctly. For the Kola/Steigen case, 52 of 53 (98.1%) Kola blasts and 23 of 24 (95.8%) Steigen earthquakes were classified correctly. This illustrates how the classification test provides greater accuracy over the outlier test for situations where more data are available.

3.5. Visualization of Likelihood Ratio Test Results

A useful way to intuitively convey the results of the GLR classification and outlier tests is graphically. Since the likelihood ratio statistic mathematically combines the multivariate discriminants for the all of the training events and the new event in question into a single univariate variable, the bootstrap distribution of the likelihood ratio and its value for the event being tested can be plotted easily in 2-dimensions.

Figure 16 shows graphical representations of the GLR outlier test results for the Vogtland (upper) and Kola/Steigen (lower) data sets. The plots on the left (right) correspond to cases for which the earthquake (quarry blast) data sets were used as the training set, while testing the alternate set. The training and test event sets are listed in the upper legends and the discriminants used are listed in the legends just below. The values of the likelihood ratios computed from the training data and each test event are depicted by the triangles. The distributions shown are smoothed histograms of the likelihood ratio obtained by bootstrapping 500 samples of the training data and new event under the null hypothesis H_0 , i.e., where the test event is from the same population as the training events. For example, to produce the distribution in the upper lefthand frame, 500 samples of 10 earthquakes for the training set and 1 earthquake to be tested were generated randomly from a normal distribution with the same mean and covariance matrix that were estimated from the actual Vogtland earthquake data. For each random bootstrap sample, the likelihood ratio is computed, binned and smoothed to generate the distribution shown. The vertical lines represent the critical values of the tests for various significance levels listed in the lowest legend in each frame. The critical value is set such there are only $100\alpha\%$ of the bootstrap test events that are rejected falsely, where events whose likelihood ratios fall to the left of the critical value are rejected as belonging to the same population as the training sample at the corresponding significance level.

The upper plots show that all Vogtland blasts are outliers of the earthquake population and vice versa, even at 0.01 significance level or less. The lower lefthand plot shows that all but two Kola blasts are outliers of the Steigen earthquake population at 0.01 significance level, another is rejected at 0.05 significance level, and the last is not rejected at any of the significance levels considered. The lower righthand plot shows that none of the Steigen earthquakes are outliers of the Kola blast population at any of the significance levels considered.

There are two reasons why the GLR outlier test performed so well for the cases considered in Section 3.4. First, the likelihood ratio methodology is very powerful which is why it has been used so extensively in previous statistical analyses. Second, the variances of the earthquake features are small and there is clean separation between the Vogtland classes and relatively good separation between the Kola blast and Steigen earthquake classes. The contrasting results in the lower plots of Figure 16 illustrate a significant point; namely, even though the separation of the classes are the same for the two cases, if feature variances are relatively large for the training set, it is much more

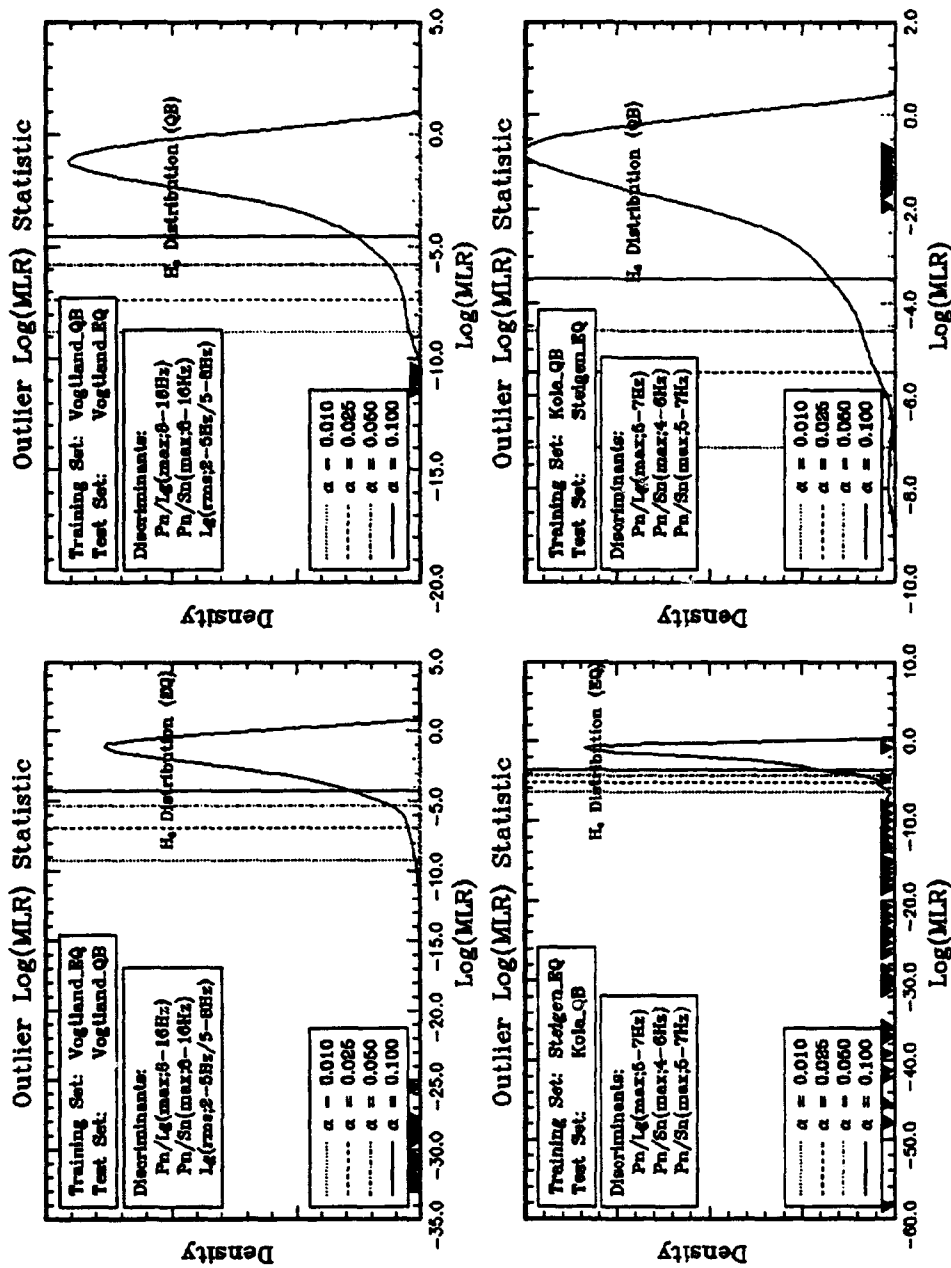


Figure 16. Graphical representation of the GLR outlier tests for the Vogtland and Kola/Steigen data sets. The distributions shown are smoothed histograms of the likelihood ratio (LR) obtained by bootstrapping. The LRs for the events that were tested are depicted by the triangles. The training and test event sets are listed in the upper legends in each plot. The vertical lines represent critical values of the tests for various significance levels listed in the lower legend. Events whose LRs fall to the left of the critical value are rejected as belonging to the training set population at that significance level. Because the variances of the Kola quarry blast discriminants are so large, none of the Steigen earthquakes are outliers of the Kola quarry blast population.

difficult to detect outliers. Sufficient separation between the Vogtland classes allows the outlier test to overcome the large feature variances of the Vogtland quarry blast data set (Figure 16, upper righthand frame).

Figure 17 shows similar plots for GLR classification test results for the Vogtland (upper) and Kola/Steigen (lower) data sets. The two training sets and the test event set for each case are listed in the upper legends and the discriminants are listed just below as before. The values of the likelihood ratios computed from the training data and each test event are depicted by the triangles. Here two bootstrap distributions are plotted for each case. The H_0 distributions shown are smoothed histograms of the likelihood ratio obtained by generating 500 bootstrap samples of the two training sets and a new event under the null hypothesis H_0 , i.e., where the test event is from the same population as the first training set. The H_1 distributions are generated similarly except that the bootstrap test events are sampled from the other population under the alternative hypothesis H_1 . The vertical lines represent the critical values of the tests for various significance levels listed in the lowest legend in each frame. Events whose likelihood ratios fall to the left of the critical value are rejected as belonging to the same population as the first training set at the given significance level. The thick vertical line shows the cut-off value of the classical classification procedure based on minimizing the total misclassification probability. Events falling to one side or the other of this line are classified with the corresponding event type.

Figure 17 illustrates a couple of important points. First, the classical cut-off value is different than, and may be to the right or left of the critical values of the hypothesis test depending on the training data and the significance level. The hypothesis test approach controls the percentage of events of the first type that are falsely rejected, while for the classical approach this false alarm rate varies from case to case but the overall misclassification probability is minimized. Both approaches are useful in different settings, the hypothesis testing approach being most useful when misclassification of a particular type of event is far more serious than other types of misclassification.

Second, using training data for both classes in the classification procedure offers significant improvement over the outlier test, particularly when the feature variances of the first training set are large. For example, comparing the lower righthand frames in Figures 16 and 17, we see that by including the earthquake training data the Steigen

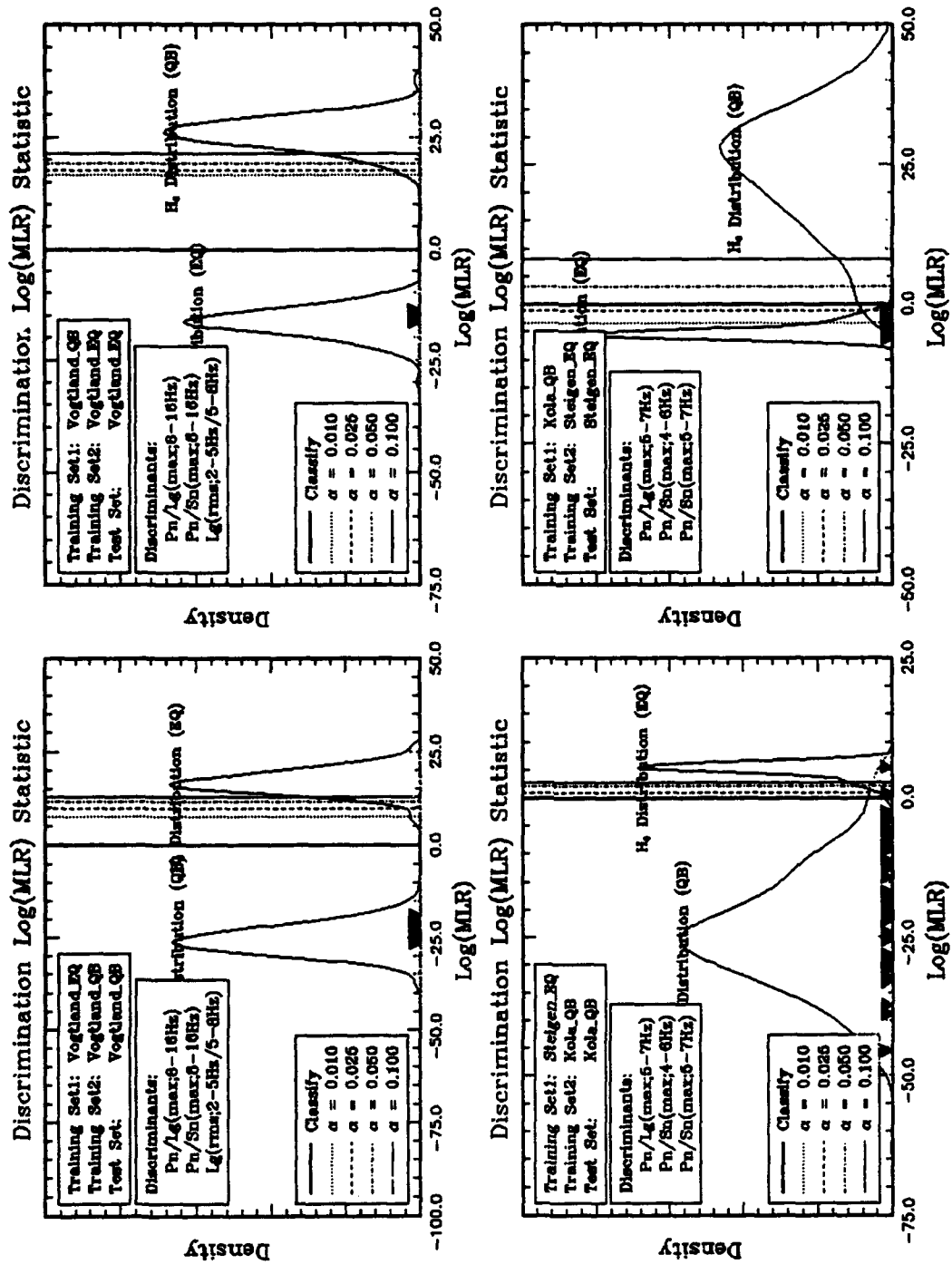


Figure 17. Graphical representation of the GLR classification tests for the Vogtland and Kola/Steigen data sets. The distributions shown are smoothed histograms of the likelihood ratio (LR) obtained by bootstrapping. The vertical lines represent critical values of the test for various significance levels listed in the lowest legend. The thick vertical line corresponds to the critical value which minimizes the total misclassification rate. The LRs for the events that were tested are depicted by the triangles.

earthquakes are now rejected in most cases as being Kola blasts by the classification test while they were not rejected by the outlier test.

3.6. Contaminated Training Set Study

The purpose of this study was to determine whether the GLR outlier test can be used to detect contaminating events in training sets which lack ground-truth and to assess the impact that contaminated data sets can have on error rates if they are used for future monitoring. To carry out this study, we randomly selected 2 blasts from each data set and inserted them in the corresponding earthquake sets. We then ran the GLR outlier test on each event in the contaminated earthquake set using the leave-one-out procedure and the same discriminants as above. Note that even as either blast is selected as the new event there is still another blast in the training set, thus making them more difficult to detect. If an outlier is detected, however, it is removed from the set and the remainder are re-tested.

To make this problem somewhat interesting for the Vogtland data set, we selected the 2 quarry blasts that are most similar to the earthquakes, rather than selecting 2 at random. Even so, both blasts were detected on the first pass. If the blasts are not removed from the earthquake training set, 2 other quarry blasts of the remaining 11 are undetected when subsequently tested. This analysis was repeated by randomly selecting 2 Kola blasts and inserting them in the Steigen earthquake set. Only one of the blasts was detected on the first pass, but once removed, the other one was also detected. If the 2 Kola blasts are not removed from the Steigen earthquake training set, 7 of the remaining 51 Kola blasts are undetected when subsequently tested. So the probability of identifying an event as an earthquake given it was actually a quarry blast, denoted $P(EQ|QB)$, goes from 0 for the uncontaminated case to 18.2% for the Vogtland region. Similarly for the Kola/Steigen region, $P(EQ|QB)$ goes from 3.8% for the uncontaminated case to 13.7%.

Thus, in addition to missing events that are included in training sets erroneously, identification error rates for future events can be affected substantially. This study shows that the GLR outlier test can be a useful tool for detecting peculiar events in training sets which lack ground-truth provided the number of contaminating events is of the order of 20% or less, depending on the discriminants. In practice, peculiar events can be flagged for further expert analyst review or corroboration with seismic event bulletins.

4. IDENTIFICATION ANALYSIS OF THE 31 DECEMBER 1992

NOVAYA ZEMLYA EVENT

4.1. Introduction

Based on origin analysis by the Intelligent Monitoring System (Bache et al., 1990), a small regional event (M_L 2.26) occurred at origin time 12/31/92 09:29:24 (GMT), latitude 73.58 and longitude 55.21 on Novaya Zemlya. (We will refer to this event by its origin identification number, ORID=361575.) Considerable interest in this event is motivated by its relevance to future CTBT or NPT monitoring scenarios since it occurred in an area of very low seismicity, had magnitude corresponding to a fully decoupled 1 kt nuclear test, and was detected by only a few arrays at regional distances (Ryall, 1993b). Of further interest is the fact that it occurred in a region where previous underground nuclear tests had been conducted.

Using amplitude ratios of seismic waveforms recorded at ARCESS, we performed statistical tests of hypothesis in order to assess the identification of the 31 December 1992 Novaya Zemlya (NZ) event. The outlier detection and classification methods described in Section 2 were applied to event 361575 using training data for previous nuclear explosions, earthquakes and quarry blasts with similar (but certainly not identical) paths to ARCESS. The data sets were provided to us by Baumgardt (1993a), who also used them to classify this event using an alternative classification approach. The preliminary training set analyses described in Section 2.4 were also applied to the data sets used here.

In the following, we describe the data sets and discriminants used in the analysis, results of the preliminary training set analyses, results of the outlier and classification tests, and our conclusions and recommendations based on this analysis.

4.2. Training Sets and Discriminants

The feature data (discriminant values) used here were obtained from seismic waveform analysis performed by Baumgardt (1993a) using the Intelligent Seismic Event Identification System (ISEIS) (Baumgardt et al., 1991). They consist of Pn/Sn ratios of maximum amplitude in five frequency bands, 4-6, 5-7, 6-8, 8-10, and 8-16 Hz, recorded by ARA0. Data for four events at NZ were provided, three of which are known nuclear explosions (EXs) and the fourth is the event in question. For comparison we also used Pn/Sn(max) in the same five frequency bands for the 53 quarry blasts (QBs) which

occurred in the Kola Peninsula region, 24 earthquakes (EQs) which occurred in the Steigen region, and 5 EQs which occurred in the direction of the Mid-Atlantic Ridge relative to ARCESS, near Spitsbergen.

A map of the event and seismic array locations is shown in Figure 18. The circles represent the EQs in the Steigen region of Norway (lower left) and in the direction of the Mid-Atlantic Ridge (upper center). The squares show the locations of the Kola Peninsula QBs. The triangles represent the 3 previous EXs at NZ, and the asterisk depicts the location of the event at NZ on 31 December 1992. The location of the ARCESS array at which these events were observed is also shown.

A scatter plot of the amplitude ratios for all of these events is shown in Figure 19. The legend in the upper lefthand corner associates the markers with the event types; event 361575 is listed as RU for regional unknown. The five frequency bands in which the ratios were computed are listed along the bottom. It is evident that the ratios for event 361575 are significantly different than those for either the NZ EXs or the Steigen EQs in all frequency bands. There are some bands for which 361575 has Pn/Sn values similar to the Mid-Atlantic Ridge EQs (e.g., 4-6, 5-7 and 8-10 Hz), while it has noticeably different values in the 6-8 and 8-16 Hz bands. The Pn/Sn values for event 361575 are consistent with those of the Kola QBs in all frequency bands.

Andrews' plots for all of the events considered here are shown in Figure 20. The four amplitude ratios listed in the upper lefthand legend were used as coefficients of the Fourier series. The line types corresponding to the various event types are given in the righthand legend. Note that the curve for event 361575 is significantly different than those for the previous NZ EXs, as well as the Steigen and Mid-Atlantic Ridge EQs, while it is quite similar to a number of the curves for the Kola QBs. To clarify this Figure 21 shows a comparison for all but the Kola QBs and Figure 22 shows a comparison for only event 361575 and the Kola QBs. Note that the curve for event 361575 does look similar to the curve for one of the Mid-Atlantic Ridge EQs in the middle of the plot. This is due to the fact that the values of Pn/Sn are nearly identical in the 5-7 Hz band (cf. Figure 19). Overall, however, the curves do look different. From Figure 22 it is clear that event 361575 has very similar amplitude ratios to the Kola QBs in all of the frequency bands.

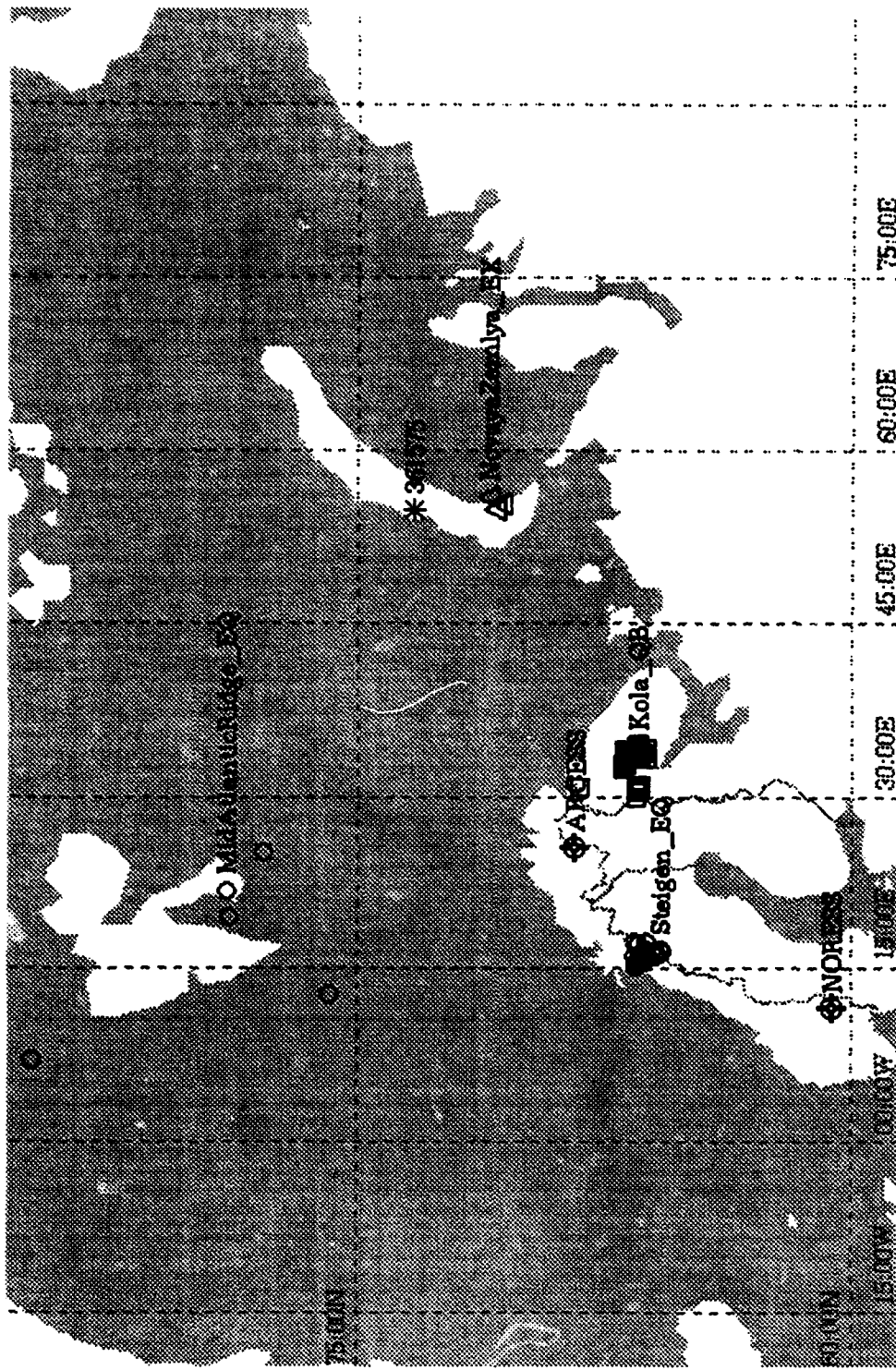


Figure 18. Locations of the seismic events and arrays used in the identification analysis of the Novaya Zemlya event on 31 December 1992 (ORID=361575). Only data from ARA0 were used in the analysis.

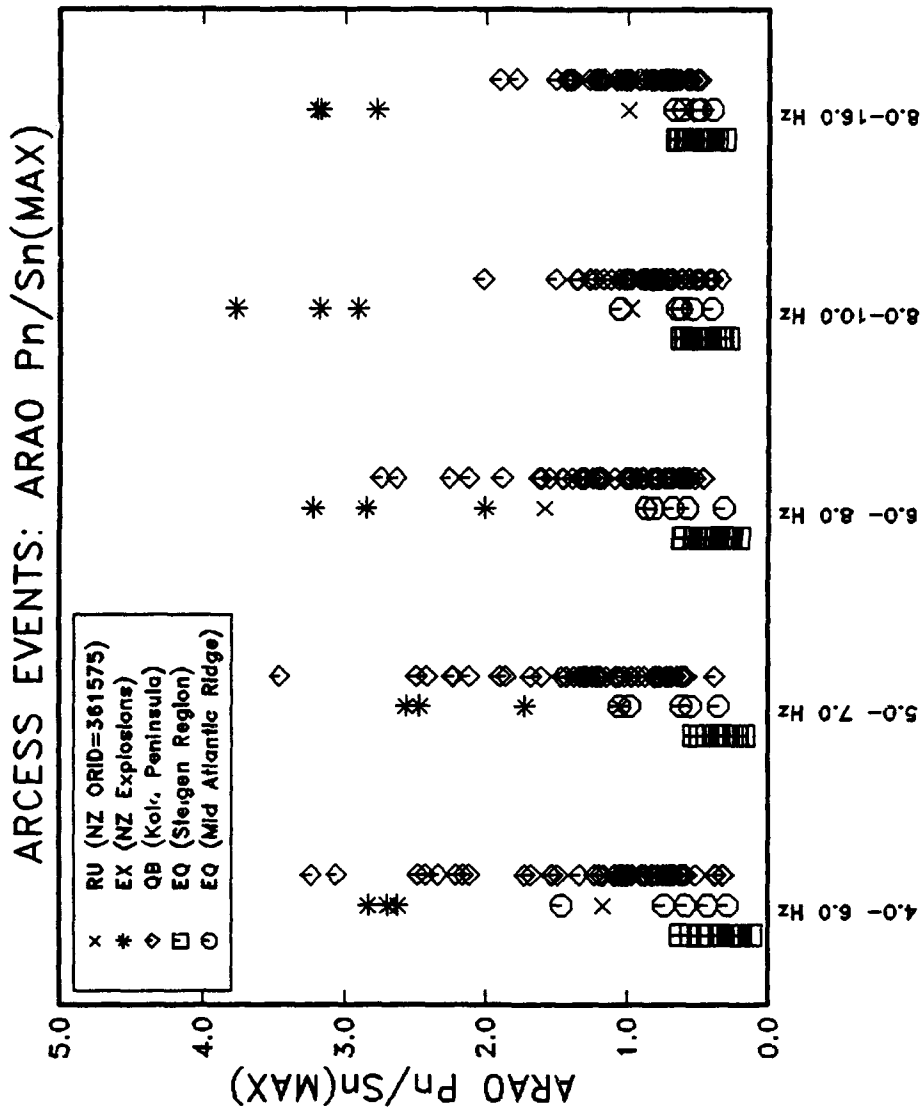


Figure 19. Scatter plot of Pn/Sn ratios of maximum amplitude in five frequency bands for all events used in the analysis of the Novaya Zemlya event on 31 December 1992. Event 361575 is listed as RU for regional unknown.

FOURIER PLOTS OF ARCESS EVENTS (ARAO)

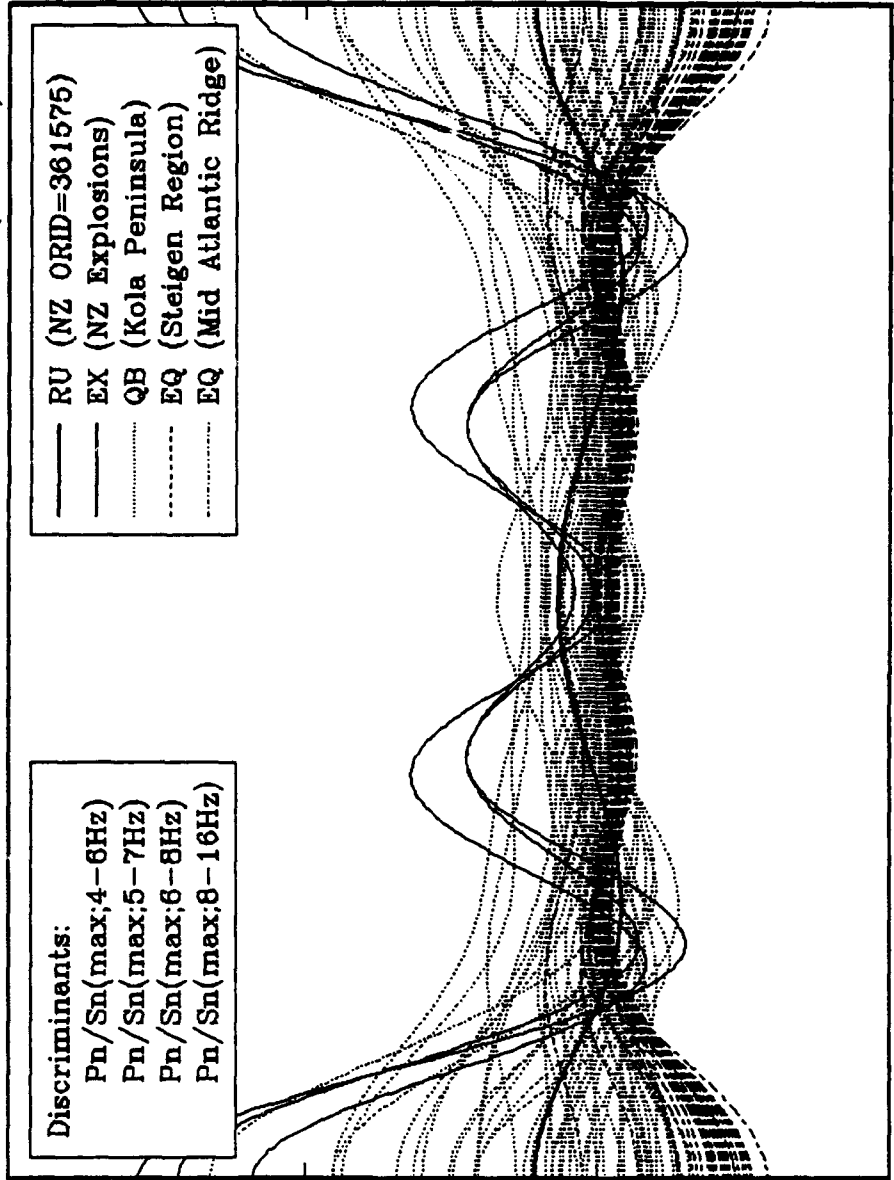


Figure 20. Andrews' Fourier plots for all events used in the analysis of event 361575. The discriminants used as coefficients of the Fourier series are given in the left-hand legend and the line types associated with the various event types are given in the right-hand legend.

FOURIER PLOTS OF ARCESS EVENTS (ARAO)

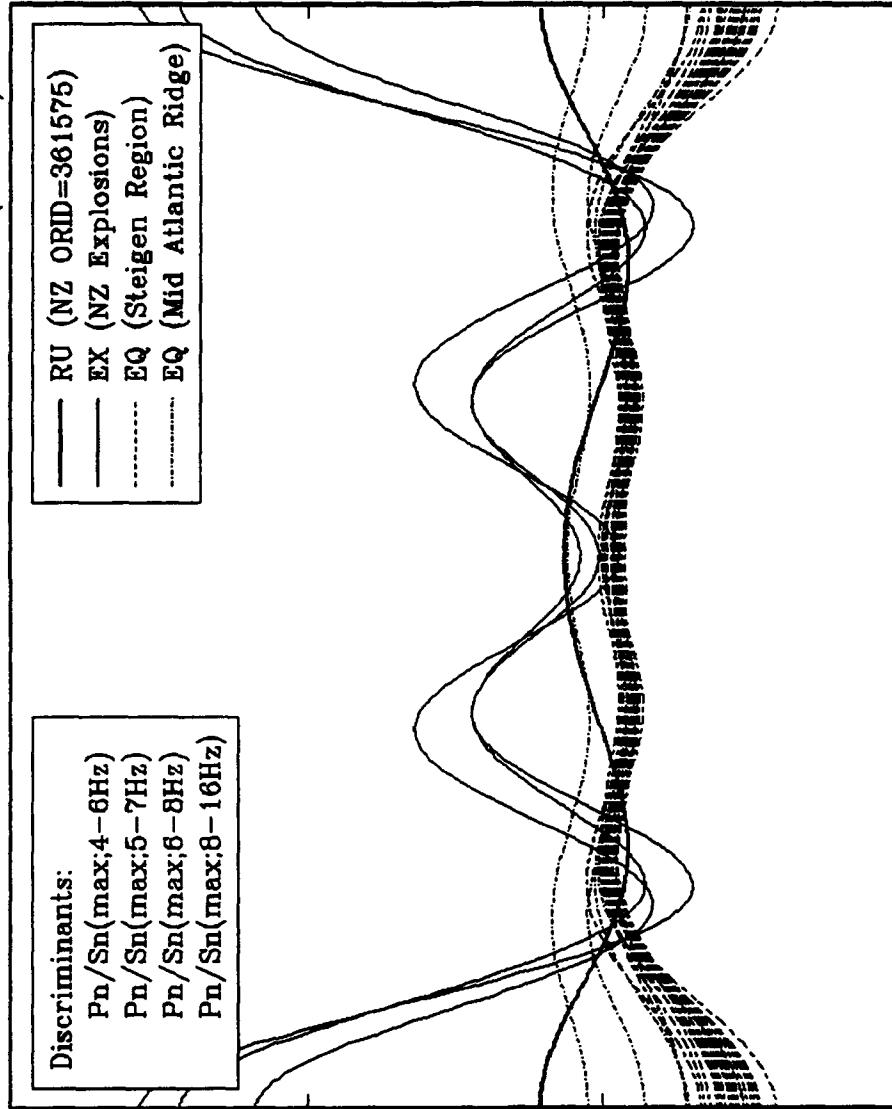


Figure 21. Andrews' Fourier plots for all events used in the analysis of event 361575 except the Kola quarry blasts.

FOURIER PLOTS OF ARCESS EVENTS (ARAO)

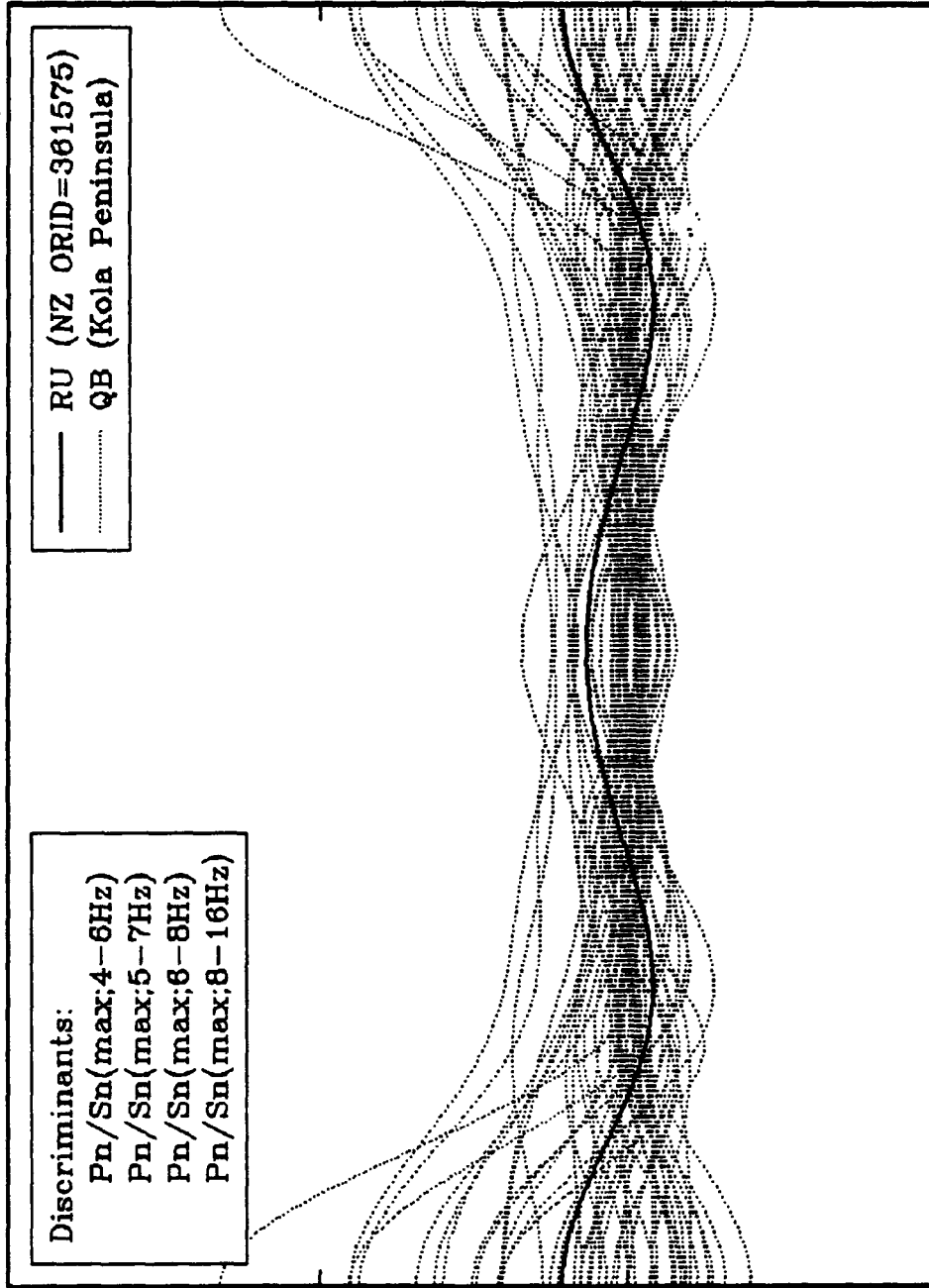


Figure 22. Andrews' Fourier plots for event 361575 and the Kola quarry blasts.

4.3. Results of Preliminary Training Set Analyses

We applied the Wilk-Shapiro and Anderson-Darling tests for normality to each discriminant at 0.05 significance level. Recall that we require that normality be accepted by both tests and for all classes of events. There was insufficient evidence to reject normality of the Pn/Sn ratios in any frequency band for the NZ EXs due to the small sample size. Prior to applying the Box-Cox transformations, normality of the Pn/Sn ratios was rejected in all frequency bands for the Kola QBs and in all but two frequency bands (0.5-2.5 Hz and 8-16 Hz) for the combined Steigen/Mid-Atlantic Ridge EQs. After applying the Box-Cox transformations all of the features used in this study are accepted as normal for all classes of events. Figures 8-11 in Section 3.3 illustrate the results of the normality analyses on the training sets for Pn/Sn(max; 4-6 Hz) and Pn/Sn(max; 5-7 Hz).

We also applied a hypothesis test at 0.05 significance level based on the F-distribution to determine whether the variances of the discriminants for the various event types are equal. Because there were only 3 events in the NZ EX training set, the covariance matrices were treated as equal for all tests involving this set. This is a more *robust procedure for classification involving small data sets*. For classification into Kola QB or Steigen/Mid-Atlantic Ridge EQ classes, the variances and, hence, the covariance matrices were found to be unequal and were treated accordingly in the classification test.

We also ranked the Pn/Sn ratios by their discriminatory power after first applying the normality and equal variance tests and Box-Cox transformations. Results for the five best Pn/Sn discriminants are summarized in Table 1.

Table 1. Results of normality and equal variance tests for ranked features.

<u>Rank</u>	<u>Feature</u>	<u>Normal</u>	<u>Equal Var.</u>
1	Pn/Sn(max;5-7 Hz)	Yes	No
2	Pn/Sn(max;4-6 Hz)	Yes	No
3	Pn/Sn(max;8-16 Hz)	Yes	No
4	Pn/Sn(max;6-8 Hz)	Yes	Yes
5	Pn/Sn(max;8-10 Hz)	Yes	Yes

4.4. Discriminant Analysis Results

After first performing the preliminary training set analyses just described, we performed the following tests of hypothesis on event 361575:

Outlier Tests:

- Outlier Test #1: tested as an outlier of the NZ EX population
- Outlier Test #2: tested as an outlier of the Kola QB population
- Outlier Test #3: tested as an outlier of the Steigen EQ population
- Outlier Test #4: tested as an outlier of the Mid-Atlantic Ridge population

Classification Tests:

- Classification Test #1: classified in NZ EX or Kola QB populations
- Classification Test #2: classified in Kola QB or the combined EQ populations

For the classification tests we also considered the Steigen and Mid-Atlantic Ridge EQs separately and found that the results remain qualitatively the same. We also examined the range of significance levels for which we reject a particular hypothesis. For some cases the limiting significance levels at which particular hypotheses were rejected were much higher than commonly acceptable. Recall that if an event is rejected at a relatively high significance level, the probability that an error was made is accordingly high. Although the standard testing approach is to fix the significance level, this analysis provides useful information regarding the range of significance levels for which a particular hypothesis is rejected or accepted. Last, we also provide classification results using the same GLR method, but setting the classification rule such that the overall misclassification rate is minimized.

4.4.1. Outlier Test Results

Outlier Test #1. Since the training sample size of NZ EXs is so small, the number of features that can be used is two or less; otherwise, the sample covariance matrix used in the likelihood ratio of the outlier test is singular. In fact, Seber (1984) recommends that only one discriminant be used for such small samples since it is very difficult to obtain a reasonable estimate of a correlation coefficient. Thus using only ARA0 Pn/Sn(max;8-16) for the 3 NZ EXs, event 361575 was tested and rejected as an EX at the 0.025 significance level. In fact it is also rejected at the 0.015 significance level. This is significant statistical evidence that event 361575 is different than previous NZ EXs.

Outlier Test #2. Using the outlier test trained on 53 Kola QBs with 5 features, ARA0 Pn/Sn(max) at 4-6, 5-7, 6-8, 8-10 and 8-16 Hz, event 361575 is accepted as a Kola QB at the 0.025 significance level. It is rejected only if the significance level is greater than or equal to 0.34. This means that we would reject event 361575 as belonging to the same population as Kola QBs only if we are willing to falsely reject 1 out of 3 actual Kola QBs as not being a Kola QB. So for a wide range of reasonable significance levels the evidence is insufficient to reject the hypothesis that this event is from the same population as Kola QBs.

Outlier Test #3. Using the outlier test trained on 24 Steigen EQs with the same 5 features as in the previous test, event 361575 is rejected as a Steigen EQ at the 0.025 significance level. This is strong statistical evidence that 361575 is not from the same population as Steigen EQs.

Outlier Test #4. As for the first test, the training sample size of 5 Mid-Atlantic Ridge EQs is small. Using only ARA0 Pn/Sn(max;8-16) for these 5 EQs, event 361575 is not rejected as a Mid-Atlantic Ridge EQ at the 0.025 significance level, but is at the 0.035 significance level and above. This suggests that event 361575 is different than previous Mid-Atlantic Ridge EQs, with a 0.035 probability that an actual Mid-Atlantic Ridge EQ would be falsely rejected.

Results are represented graphically in Figure 23 for the 4 outlier tests. The distributions shown are smoothed histograms of the likelihood ratio (LR) obtained by bootstrapping the corresponding training set. The vertical lines shown correspond to critical values of the tests for various significance levels given in the lowest legend. The LR for event 361575 is denoted by the triangle marker. The training set and discriminants are listed in the upper legends. Figure 23 clearly shows that the LR for event 361575 is in the tails of the EX and EQ LR distributions, while it lies roughly in the middle of the LR distribution for the Kola QBs.

4.4.2. Classification Test Results

Classification Test #1. Using the GLR classification test trained on the 3 NZ EXs and 53 Kola QBs with ARA0 Pn/Sn(max) at 4-6, 5-7, 6-8 and 8-16 Hz, event 361575 is rejected as an NZ EX at the 0.025 significance level. It is also rejected at the 0.01 significance level. This is strong statistical evidence that event 361575 is not from the same population as NZ EXs, with preferred classification as a Kola QB. (Note that the

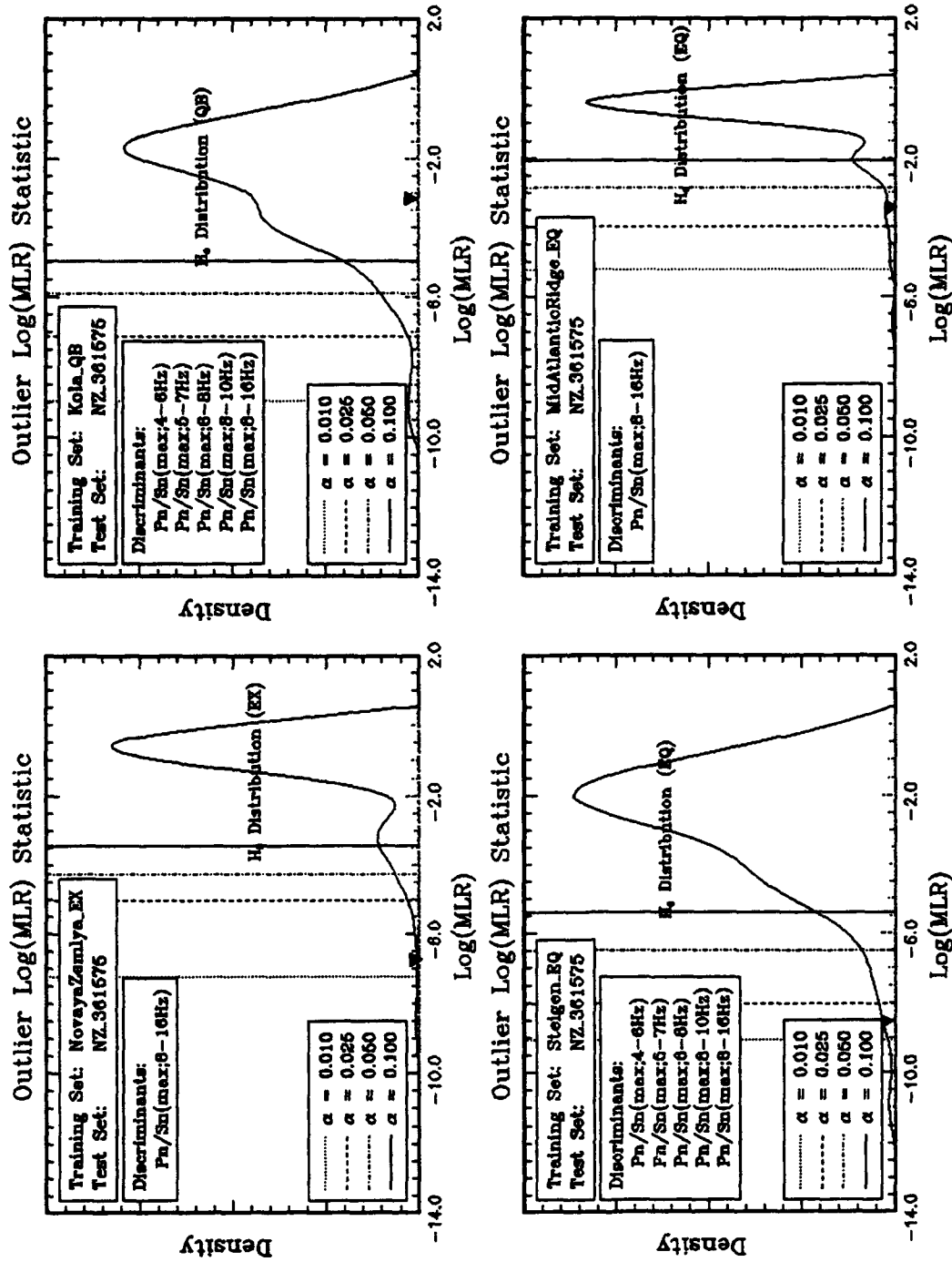


Figure 23. Graphical representation of the four GLR outlier tests applied to NZ event 361575. The distributions shown are smoothed histograms of the likelihood ratio (LR) obtained by bootstrapping. The LR for event 361575 is depicted by the triangle. The vertical lines represent critical values of the test for various significance levels listed in the lowest legend. The training sets and discriminants used for each test are listed in the upper and middle legends, respectively.

assumption of equal covariance matrices was used for this test. This allowed a pooled estimate of the covariance matrix to be computed for the four features using 56 samples.)

Classification Test #2. Since event 361575 was classified as a Kola QB in the first of these tests, it was then tested as belonging to either the Kola QB or the combined Steigen/Mid-Atlantic Ridge EQ sets. Using the GLR classification test, based on these training sets and the same four features, event 361575 is rejected as an EQ at the 0.025 significance level, set with respect to the EQs. It is also rejected at the 0.01 significance level of falsely rejecting an EQ. Again, this is strong statistical evidence that 361575 is not from the same population as the Steigen or Mid-Atlantic Ridge EQs; classification as a Kola QB is preferred. It does not imply that event 361575 could not also belong to some other class of events that are not included in the training sets.

Results of these tests are represented graphically in Figure 24. The distributions shown are smoothed histograms of the LR, bootstrapped from the training data, where the H_0 (null) distribution corresponds to LRs if the new event is from the first population and the H_1 (alternative) distribution corresponds to LRs if the new event is from the second population. As for the graphical representations of the outlier tests, the critical values of the tests for various significance levels are depicted by the vertical lines. The thick vertical line is the value of the LR which minimizes the total misclassification rate. Classification based on minimizing the total misclassification rate also allocate 361575 to the QB population.

4.5. Conclusions and Recommendations Regarding Event 361575

With the exception of the outlier test trained on $Pn/Sn(\max;8-16)$ for the 5 Mid-Atlantic Ridge EQs, the results show that event 361575 is rejected statistically by all of the tests at 0.025 significance as belonging to the EX or EQ populations. In most cases, these results also held for smaller significance levels, i.e., 0.01 for both of the GLR classification tests. Event 361575 is also rejected as a Mid-Atlantic Ridge EQ if we are willing to accept an earthquake false alarm rate of 3.5% or larger.

The outlier and classification results also show that there is insufficient evidence to reject event 361575 as belonging to the Kola QB population for all of the relevant tests at 0.025 significance. Even if the significance level of the outlier test is 0.33, event 361575 is not rejected. Classification based on minimizing the total misclassification rate clearly allocates event 361575 to the QB population. Thus, based on these tests and the

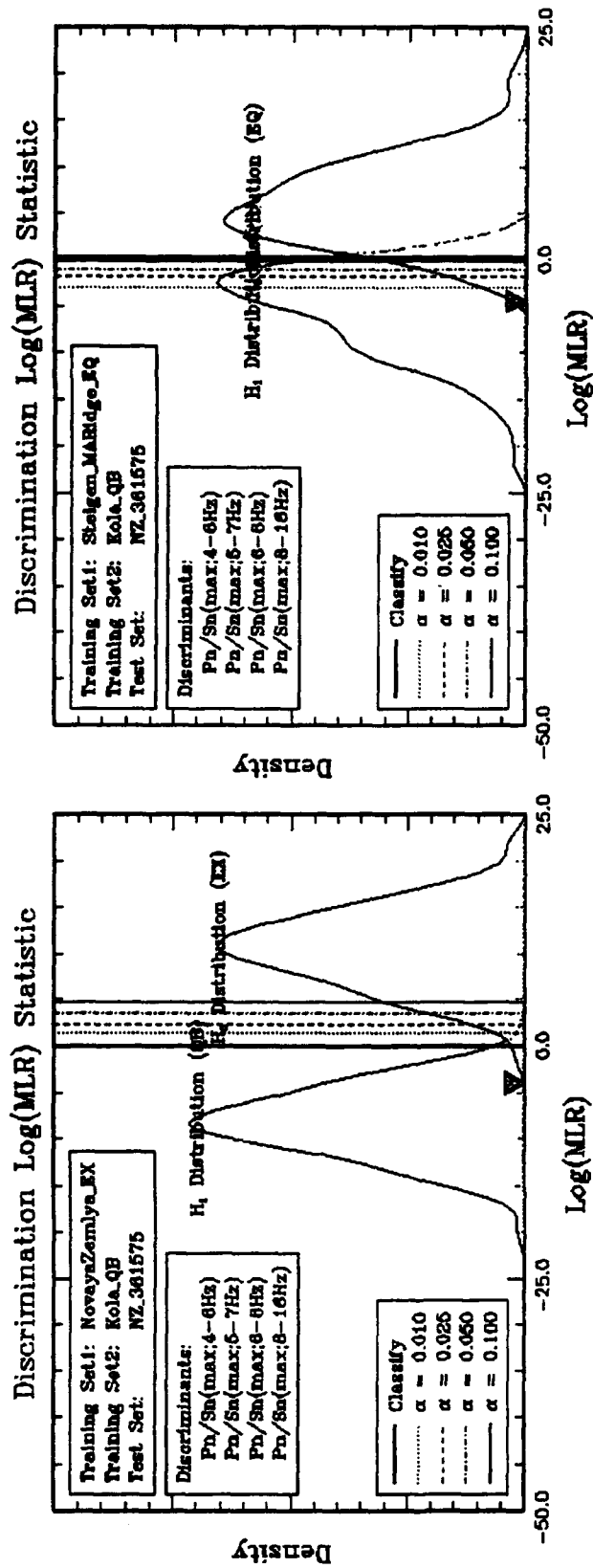


Figure 24. Graphical representation of the two GLR classification tests applied to NZ event 361575. The distributions shown are smoothed histograms of the LR obtained by bootstrapping. The LR for event 361575 is depicted by the triangle. The vertical lines represent critical values of the test for various significance levels listed in the lowest legend. The thick vertical line corresponds to the critical value which minimizes the total misclassification rate. The training sets and discriminants used for each test are listed in the upper and middle legends, respectively.

training data for the three classes considered, event 361575 is classified as belonging to the Kola QB population.

Note that these results do not imply that this event was necessarily a quarry blast, only that this classification is strongly preferred over the alternatives and there is insufficient evidence to reject it as a Kola QB, based on the training data. Since we do not know the actual distribution from which this event was produced, we cannot estimate the probability that the event was classified as a Kola QB when it might, in fact, be something else. Thus, the strongest conclusion that we can draw, *based on the data*, is that event 361575 was not a nuclear explosion similar to others at NZ nor an earthquake similar to those in the Steigen region or in the direction of the Mid-Atlantic Ridge.

Although our quantitative measures are substantially different, our qualitative results are the same as those of Baumgardt (1993b), who also concluded, based on a neural net approach and a similar set of discriminants, that event 361575 looks most like a Kola QB. This is not surprising since the training sets are effectively the same and no matter how the amplitude ratios for event 361575 are examined, plotted or statistically tested, they look similar to the Kola QB amplitude ratios.

Pulli and Dysart (1993) also applied a neural network technique trained on data for events in Scandinavia from the NORESS array. Using broadband and three narrowband measurements of Pn/Sn and Pn and Sn cepstral variances extracted from signals recorded at the ARCESS array, event 361575 was again classified as a chemical explosion with very high confidence (see Pulli and Dysart, 1993, Figure 7). In seeming contrast to the neural network result, they argue that examination of the individual features shows that the Pn and Sn cepstral variances for the event in question fall easily within the ranges for earthquakes. They also note that the value of the Pn/Sn spectral ratio from 10-20 Hz is well below the peak in the distribution for chemical explosions and fits into the distribution of earthquakes.

In response to a request of the Seismological Service of the Ministry of Defense, Russian Federation, regarding the origin of event 361575, Ryall (1993b) notes that the response was "that on the day in question there were no blasting activities, either for military or construction purposes, on the territory of the Novaya Zemlya test range."

Although our outlier and classification test results are consistent with all other examinations of the ISEIS feature data, a key question remaining to be answered is whether the discriminants based on the observed seismic signals are truly characteristic of the source. Our results should be tempered by the caveat that there may be significant path differences to ARCESS from the various regions which warrant investigation. For example, $P_n/S_n(\max)$ ratios for explosions differ dramatically for the NZ to ARCESS path as compared to the NZ to NORESS path. The path distance from NZ to NORESS is roughly double that from NZ to ARCESS. Figure 25 shows $P_n/S_n(\max)$ in 5 frequency bands for the 3 known nuclear explosions at NZ recorded by the ARCESS (ARA0) and NORESS (NRA0) arrays. The NRA0 P_n/S_n values are significantly larger than the ARA0 values for the same events. This suggests that S_n attenuates more rapidly with distance than P_n .

Even though all discriminants were obtained from signals recorded by a common instrument at ARCESS for our study, path differences could yield different results for the tests comparing event 361575 to Kola or Steigen events. Distances from ARCESS range from 200-300 km for the Kola QBs, 400-500 km for the Steigen EQs, and 900-1100 km for the Mid-Atlantic Ridge EQs and the NZ events (Baumgardt, 1993b). More rapid attenuation of S_n with distance relative to P_n would cause event 361575 to have higher P_n/S_n ratios than a similar event occurring in the Steigen or Kola Peninsula regions due to the longer propagation distance from NZ. Applying distance corrections may lead to ratios that are far more consistent with earthquakes than with quarry blasts.

This illustrates a key issue that must be addressed in order to monitor a CTBT effectively. Namely, even state-of-the-art classification methods can identify events accurately only if the features used as input are representative of the source type. For situations in which data are sparse for a specific region, such as the case here and as we can anticipate for future monitoring, path corrections are crucial if training sets from distinct regions and at different distances are to be used successfully.

Note, however, that our outlier test based on the 3 NZ EXs is not subject to this caveat as are the tests using data from other regions performed by us and others. Thus until path differences can be investigated suitably, this test may provide the only reliable statistical evidence that event 361575 was not a nuclear test. Even though the outlier test accounts for sample size, a larger nuclear explosion training set for NZ events is desired to strengthen our conclusions. Furthermore, as noted by Ryall (1993c) and Sereno

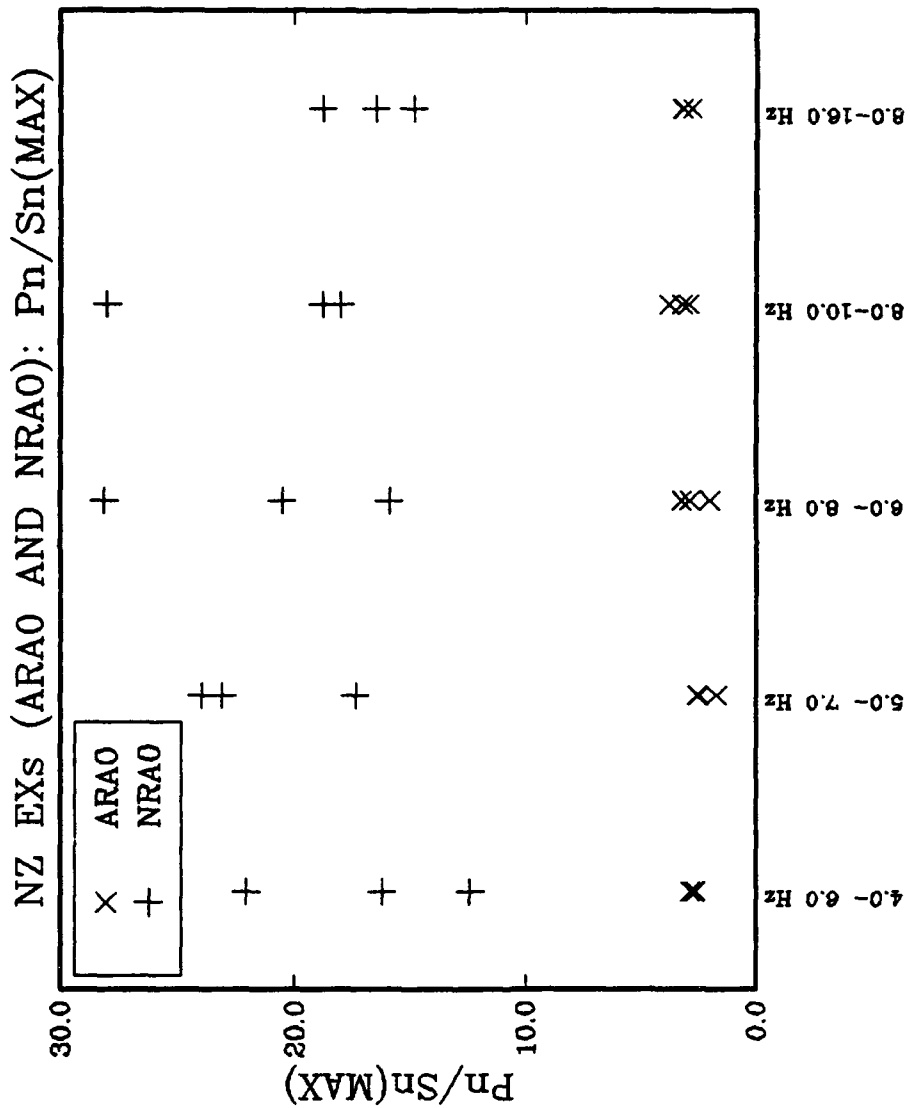


Figure 25. Scatter plot of Pn/Sn ratios of maximum amplitude in five frequency bands for 3 nuclear explosions at the Novaya Zemlya test site that were recorded in common by the ARCESS (ARAO) and NORESS (NRAO) arrays. The Pn/Sn values for the events recorded by NRAO are considerably larger than those recorded by ARAO, apparently due to greater Sn attenuation relative to Pn over roughly double the propagation distance from NZ to NRAO as compared to the NZ to ARAO path.

(1993), this result should also be tempered by the caveat that, even if event 361575 was a nuclear explosion, there are likely to be significant differences in the signal generated by that event as compared to those generated by the three known nuclear explosions at NZ, based simply on source size. This issue also warrants future investigation.

As a final comment, the relatively limited set of discriminants used here does not take full advantage of the generalized likelihood ratio approach which can also treat interesting discrete discriminants (contextual features and others) such as presence of cepstral peaks, offshore/onshore, high/low seismicity, high/low magnitude, etc. Based on analysis by Sereno and Wahl (1993), it is not clear whether these features provide additional useful information in this case but they could certainly be included in our outlier and classification procedures for more complete analysis.

5. CONCLUSIONS AND RECOMMENDATIONS

The results presented in this report illustrate the usefulness of the GLR outlier test for monitoring a new region in which few data are available and to test data sets which lack ground-truth in order to detect suspicious events that warrant further study before including them in training sets. We found that the outlier test detected 100% and 96.2% of the "suspicious" events in the Vogtland and Kola/Steigen regions, with relatively small false alarm rates, using only small samples (10 and 25) of earthquake data to train the algorithm. We showed in Section 3.6 that if contaminated training sets are used, the error rates can be unnecessarily high. The outlier test, however, was able to detect the contaminating events in the earthquake training sets for proportions of 10-20% contamination and perhaps higher.

We have also applied the outlier test to the problem of script-matching to associate seismic events to particular mines. In this case one of the script variables used is day-of-the-week which is a discrete variable. Use of this discriminant is based on the observation that some mining operations perform blasting more heavily on particular days of the week than others. We will report on these results in the future.

We have provided a version of the GLR outlier test to Dr. Thomas Sereno of SAIC to be used as a script-matching module of EVID, an event identification subsystem of the Intelligent Monitoring System at the Center for Seismic Studies.

The results also showed that the GLR classification test provides an improved procedure for event identification for cases in which training data are available for all relevant classes of events. The test identified all of the Vogtland events, 98.1% (52 of 53) of the Kola blasts and 95.8% (23 of 24) of the Steigen earthquakes correctly. We also showed how the bootstrap procedure allows one of the misclassification rates to be controlled or, alternatively, the total misclassification rate associated with all event types to be minimized.

In addition, we presented preliminary training set analyses to test appropriate assumptions and to transform the data, if necessary, to simplify implementation of the outlier and classification tests and to ensure that the necessary assumptions hold. We showed that the techniques presented to test and transform the data to normality are very effective for this purpose. We also discussed methods that are useful for testing whether covariance matrices are equal, for treating missing data values and for selecting the best multivariate discriminants.

We also presented two types of graphical techniques to help visualize comparisons of multivariate data (Andrews' plots) and the results of the outlier and classification tests.

The preliminary analyses and the outlier and classification tests were applied to the 31 December 1992 Novaya Zemlya event. This case represents a archetypical example of the type of monitoring that is likely to be performed under a CTBT or NPT. While the results obtained were consistent with the data and with other event identification results obtained by other methods, it highlights the need to carefully treat data for path differences. Application of our outlier test based on previous nuclear explosions at Novaya Zemlya was not subject to this concern, which illustrates its usefulness in this setting. Scaling relations of discriminants for sources of widely varying magnitude may, however, be needed.

The collection of statistical methods presented here, as well as others to perform path corrections and test transportability of discriminants, should be (and in fact are being) included in a single software system to provide a robust, rigorous, cohesive and complete framework for routine seismic event identification. This system is being developed with appropriate logic and a Motif-style X Window graphical user interface to

allow semi-automated application, or to allow statisticians and seismologists alike to use these methods interactively with greater ease.

We plan to start applying these methods to monitoring particular regions of current interest.

The case study of the 31 December 1992 Novaya Zemlya event illustrates that further efforts are needed to understand where and why discriminants work and to develop techniques to transport discriminants from one region to another. We are pursuing this currently using the Soviet data set discussed in the Executive Summary and a Western U.S. data set established and discussed by Taylor et al. (1989). Our emphasis is on statistically testing the equality of discriminant distributions along various paths to first determine where they are equivalent and work on a consistent basis. This will be the subject of a future report.

REFERENCES

- Adushkin, V., I. Kitov, V. Kovalenko, A. Peshkov, G. Shedlovskii, D. Sultanov (1992). Private Communication.
- Anderson, T.W. (1984). *An Introduction to Multivariate Analysis*, Second Edition, John Wiley and Sons, Inc., New York.
- Anderson, T.W. (1973). An asymptotic expansion of the distribution of the standardized classification statistic W , *Ann. Stat.*, **1**, 964-972.
- Anderson, T.W. (1958). *An Introduction to Multivariate Analysis*, John Wiley and Sons, Inc., New York.
- Anderson, T.W. (1951). Classification by multivariate analysis, *Psychometrika*, **16**, 631-650.
- Anderson, T.W. and D.A. Darling (1954). A test of goodness-of-fit, *J. Am. Stat. Assoc.*, **49**, 765-769.
- Andrews, D.F. (1972). Plots of high dimensional data, *Biometrics*, **28**, 125-136.
- Bache, T.C., S.R. Bratt, J. Wang, R.M. Fung, C. Kobryn and J.W. Given (1990). The Intelligent Monitoring System, *Bull. Seism. Soc. Am.*, **80**, 1833-1851.
- Baek, J., H. L. Gray, and W. A. Woodward (1992). A generalized likelihood ratio test in outlier detection or script matching, submitted to *J. Am. Stat. Assoc.*, Southern Methodist University.
- Baek, J., H. L. Gray, W. A. Woodward and M.D. Fisk (1993). A bootstrap generalized likelihood ratio test in discriminant analysis, submitted to *J. Am. Stat. Assoc.*, Southern Methodist University.
- Baumgardt, D.R. (1993a). Private Communication.
- Baumgardt, D.R. (1993b). Seismic waveform feature analysis and discrimination of the December 31, 1992 Novaya Zemlya event, Paper in Ryall (1993b), PL-TR-93-2160, ADA271458.
- Baumgardt, D.R. (1992a). Private Communication.
- Baumgardt, D.R. (1992b). Investigation of Seismic Discriminants in Eurasia, Final Technical Report SAS-TR-92-81, ENSCO, Inc., Springfield, VA.

- Baumgardt, D.R. (1987). Case-Based Reasoning Applied to Regional Seismic Event Characterization, *Proceedings of the 9th Annual DARPA/AFGL Seismic Research Symposium*, GL-TR-90-0300, Phillips Laboratory, Hanscom AFB, MA, 173-178, ADA229025.
- Baumgardt, D.R., J. Carney, M. Maxson and S. Carter (1992). Evaluation of Regional Discriminants using the Intelligent Seismic Event Identification System, Semi-Annual Technical Report SAS-TR-93-38, ENSCO, Inc., Springfield, VA.
- Baumgardt, D.R., G. Young and K. Ziegler (1991). Design and Development of the Intelligent Seismic Event Identification System: Design Considerations and Processing for Regional Event Identification, Technical Report PL-TR-91-2211, Phillips Laboratory, Hanscom AFB, MA, ADA246793.
- Box, G.E.P. and D.R. Cox (1964). An analysis of transformations, *J. R. Stat. Soc.*, **26**, 211-252.
- Burger, R.W., L.J. Burdick and T. Lay (1986). Estimating the relative yields of Novaya Zemlya tests by waveform intercorrelation, *Geophys. J. R. Astr. Soc.*, **87**, 523-537.
- Caroni, C. and P. Prescott (1992). Sequential application of Wilks' multivariate outlier test, *Appl. Stat.*, **41**, 355-364.
- Dempster, A.P., N.M. Laird and D.B. Rubin (1977). Maximum likelihood estimation from incomplete data via the EM algorithm, *J. Royal Stat. Soc.*, **39**, 1-38.
- Dyer, A.R. (1974). Comparison of tests for normality with a cautionary note, *Biometrika*, **61**, 185-189.
- Efron, B. (1982). The jackknife, the bootstrap and other resampling plans, *SIAM*, Philadelphia.
- Efron, B. (1979). Bootstrap methods: another look at the jackknife, *Ann. Stat.*, **7**, 1-26.
- Fisk, M. D., R. W. Alewine, H. L. Gray, G. D. McCartor (1992). Multivariate Seismic Calibration for the Novaya Zemlya Test Site, PL-TR-92-2251, Phillips Laboratory, Hanscom AFB, MA, ADA261725.
- Fisk, M. D. and H. L. Gray (1993). Event Identification Analysis of the Novaya Zemlya Event on 31 December 1992 using Outlier and Classification Likelihood Ratio Tests, (Paper in Ryall, 1993b, PL-TR-93-2160, ADA271458), MRC-R-1449, Mission Research Corp., Santa Barbara, CA.

- Fisk, M. D., H. L. Gray, G. D. McCartor and G. L. Wilson (1991a). A Constrained Bayesian Approach for Testing TTBT Compliance, PL-TR-91-2170, Phillips Laboratory, Hanscom AFB, MA, ADA253288.
- Fisk, M. D., H. L. Gray, G. D. McCartor and G. L. Wilson (1991b). Estimates of the Covariance Parameters for Novaya Zemlya and Robustness of Point Estimates for Unequal Yields, MRC-N-940, Mission Research Corp., Santa Barbara, CA.
- Grant, L., J. Coyne and F. Ryall (1993). CSS Ground-Truth Database: Version 1 Handbook, Technical Report C93-05, Science Applications International Corp., Center for Seismic Studies, Arlington, VA.
- Gray, H. L., J. Baek, G. D. McCartor and W. A. Woodward (1992). A Bayesian Method for Testing TTBT Compliance with Unknown Intercept and Slope, SMU Technical Report.
- Gray, H. L., W. A. Woodward, G. D. McCartor and M. D. Fisk (1993). A bootstrap generalized likelihood ratio test in discriminant analysis, in *Proceedings of the 15th Annual Seismic Research Symposium, 8-10 September 1993*, Report PL-TR-93-2160 (Eds. J. F. Lewkowitz and J. M. McPhetres), Phillips Laboratory, Hanscom AFB, MA, ADA271458.
- GSE/SW/62 (1988). IA system - Interactive program for association and locations of seismic events. Conference of Disarmament, Working Paper, 17 March 1988.
- Israelsson, H. (1992). RMS Lg as a Yield Estimator in Eurasia, PL-TR-92-2117(I), Phillips Laboratory, Hanscom AFB, MA, ADA256692.
- Jih, R.-S. and C. S. Lynnes (1992). Studies of Regional Phase Propagation in Eurasia, PL-TR-93-2003, Phillips Laboratory, Hanscom AFB, MA, ADA262801.
- Jih, R.-S., R.R. Shumway, D.W. Rivers, R.A. Wagner and T.W. McElfresh (1990). Magnitude:Yield Relationship at Various Nuclear Test Sites - A Maximum Likelihood Approach using Heavily Censored Explosive Yields, GL-TR-90-0107, Phillips Laboratory, Hanscom AFB, MA, ADA223490.
- John, S. (1963). On classification by the statistics R and Z , *Ann. Inst. Stat. Math.*, **14**, 237-246.
- Johnson, R.A. and D.W. Wichern (1988). *Applied Multivariate Statistical Analysis*, Prentice Hall, New Jersey.

- Kanazawa, M. (1979). The asymptotic cut-off point and comparison of error probabilities in covariate discriminant analysis, *J. Japan Stat. Soc.*, **9**, 7-17.
- Knoke, J.D. (1982). Discriminant analysis with discrete and continuous variables, *Biometrics*, **38**, 191-200.
- Krzanowski, W.J. (1975). Discrimination and classification using both binary and continuous variables, *J. Am. Stat. Assoc.*, **70**, 782-790.
- Krzanowski, W.J. (1979). Some linear transformations for mixtures of binary and continuous variables, with particular reference to linear discriminant analysis, *Biometrika*, **66**, 33-39.
- Krzanowski, W.J. (1980). Mixtures of continuous and categorical variables in discriminant analysis, *Biometrics*, **36**, 493-499.
- Krzanowski, W.J. (1982). Mixtures of continuous and categorical variables in discriminant analysis: A hypothesis-testing approach, *Biometrics*, **38**, 991-1002.
- Layard, M.W.J. (1974). A Monte Carlo comparison of tests for equality of covariance matrices, *Biometrika*, **16**, 461-465.
- McLachlan, G. J. (1987). On bootstrapping the likelihood ratio test statistic for the number of components in a normal mixture, *App. Stat.*, **36**, 318-324.
- Nicholson, W.L., R.W. Mensing and H.L. Gray (1991). Private Communication.
- Nutli, O.W. (1988). Lg magnitudes and yield estimates for underground Novaya Zemlya nuclear explosions, *Bull. Seis. Soc. Am.*, **78**, 873-884.
- Nutli, O.W. (1986). Lg magnitudes of selected East Kazakhstan underground explosions, *Bull. Seis. Soc. Am.*, **76**, 1241-1251.
- Olkin, I. and R.F. Tate, (1961). Multivariate Correlation Models with Mixed Discrete and Continuous Variables, *Ann. Math. Stat.*, **22**, 92-96.
- Pervaiz, M.K. and C.J. Skinner (1990). A Monte Carlo comparison of Elliptical-Theory and other tests for equality of covariance matrices, *Australian J. Stat.*, **32**, 71-86.
- Pettitt, A.N. (1977). Testing the normality of several independent samples using the Anderson-Darling statistic, *Appl. Stat.*, **26**, 156-161.
- Pulli and Dysart (1993). Identification analysis of the Dec. 31, 1992 Novaya Zemlya event, Paper in Ryall (1993b), PL-TR-93-2160, ADA271458.
- Richards, P. G. (1992). Private Communication.

- Ringdal, F. and J. Fyen (1991). RMS Lg analysis of Novaya Zemlya explosion recordings, Semiannual Technical Summary, 1 Oct 1990 - 31 Mar 1991, *NORSAR Scientific Report No. 2-90/91*, NTNF/NORSAR, Kjeller, Norway.
- Ryall, A. (1993a). Seismic Identification Issues and the Novaya Zemlya Event of 31 December 1992, *Presentation at the 15th Annual Seismic Research Symposium*, 8-10 September 1993, Vail, CO, PL-TR-93-2160, ADA271458.
- Ryall, A. (1993b). The Novaya Zemlya Event of 31 December 1992 and Seismic Identification Issues, *Supplemental volume of papers to Proceedings of the 15th Annual Seismic Research Symposium*, 8-10 September 1993, Vail, CO, PL-TR-93-2160, ADA271458.
- Ryall, A. (1993c). Private Communication.
- Schwartz, S. (1992). Private Communication.
- Seber, G.A.F. (1984). *Multivariate Observations*, John Wiley, New York.
- Sereno, T.J. (1993). Private Communication.
- Sereno, T.J. and D. Wahl (1993). A fuzzy-logic approach to regional seismic event identification: application to the Novaya Zemlya event on 31 December 1992, Paper in Ryall (1993b), PL-TR-93-2160, ADA271458.
- Shapiro, S.S. and M.B. Wilk (1965). An analysis of variance test for normality (complete samples), *Biometrika*, 52, 591-611.
- Shumway, R.H. and Z.A. Der (1990). Multivariate calibration and yield estimation for nuclear tests, Preprint, University of California, Davis.
- Sykes, L.R. and S. Ruggi (1988). Soviet nuclear testing, in *Nuclear Weapon Databook* (Volume IV, Chapter 10), Natural Resources Defense Council, Washington, D.C.
- Taylor, S.R., M.D. Denny, E.S. Vergino and R.E. Glaser (1989). Regional discrimination between NTS explosions and western U.S. earthquakes, *Bull. Seism. Soc. Am.*, 79, 1142-1176.
- Tu, C. and C. Han (1982). Discriminant analysis based on binary and continuous variables, *J. Am. Stat. Ass.*, 77, 447-454.
- Welch, B.L. (1939). Note on discriminant functions, *Biometrika*, 31, 218-220.
- Wilks, S.S. (1963). Multivariate statistical outliers," *Sankhya*, 25, 407-426.

Prof. Thomas Ahrens
Seismological Lab, 252-21
Division of Geological & Planetary Sciences
California Institute of Technology
Pasadena, CA 91125

Prof. Keiiti Aki
Center for Earth Sciences
University of Southern California
University Park
Los Angeles, CA 90089-0741

Prof. Shelton Alexander
Geosciences Department
403 Deike Building
The Pennsylvania State University
University Park, PA 16802

Prof. Charles B. Archambeau
CIRES
University of Colorado
Boulder, CO 80309

Dr. Thomas C. Bache, Jr.
Science Applications Int'l Corp.
10260 Campus Point Drive
San Diego, CA 92121 (2 copies)

Prof. Muawia Barazangi
Institute for the Study of the Continent
Cornell University
Ithaca, NY 14853

Dr. Jeff Barker
Department of Geological Sciences
State University of New York
at Binghamton
Vestal, NY 13901

Dr. Douglas R. Baumgardt
ENSCO, Inc
5400 Port Royal Road
Springfield, VA 22151-2388

Dr. Susan Beck
Department of Geosciences
Building #77
University of Arizona
Tucson, AZ 85721

Dr. T.J. Bennett
S-CUBED
A Division of Maxwell Laboratories
11800 Sunrise Valley Drive, Suite 1212
Reston, VA 22091

Dr. Robert Blandford
AFTAC/TT, Center for Seismic Studies
1300 North 17th Street
Suite 1450
Arlington, VA 22209-2308

Dr. Stephen Bratt
ARPA/NMRO
3701 North Fairfax Drive
Arlington, VA 22203-1714

Dr. Lawrence Burdick
IGPP, A-025
Scripps Institute of Oceanography
University of California, San Diego
La Jolla, CA 92093

Dr. Robert Burrige
Schlumberger-Doll Research Center
Old Quarry Road
Ridgefield, CT 06877

Dr. Jerry Carter
Center for Seismic Studies
1300 North 17th Street
Suite 1450
Arlington, VA 22209-2308

Dr. Eric Chael
Division 9241
Sandia Laboratory
Albuquerque, NM 87185

Dr. Martin Chapman
Department of Geological Sciences
Virginia Polytechnical Institute
21044 Derring Hall
Blacksburg, VA 24061

Mr Robert Cockerham
Arms Control & Disarmament Agency
320 21st Street North West
Room 5741
Washington, DC 20451,

Prof. Vernon F. Cormier
Department of Geology & Geophysics
U-45, Room 207
University of Connecticut
Storrs, CT 06268

Prof. Steven Day
Department of Geological Sciences
San Diego State University
San Diego, CA 92182

Marvin Denny
U.S. Department of Energy
Office of Arms Control
Washington, DC 20585

Dr. Cliff Frolich
Institute of Geophysics
8701 North Mopac
Austin, TX 78759

Dr. Zoltan Der
ENSCO, Inc.
5400 Port Royal Road
Springfield, VA 22151-2388

Dr. Holly Given
IGPP, A-025
Scripps Institute of Oceanography
University of California, San Diego
La Jolla, CA 92093

Prof. Adam Dziewonski
Hoffman Laboratory, Harvard University
Dept. of Earth Atmos. & Planetary Sciences
20 Oxford Street
Cambridge, MA 02138

Dr. Jeffrey W. Given
SAIC
10260 Campus Point Drive
San Diego, CA 92121

Prof. John Ebel
Department of Geology & Geophysics
Boston College
Chestnut Hill, MA 02167

Dr. Dale Glover
Defense Intelligence Agency
ATTN: ODT-1B
Washington, DC 20301

Eric Fielding
SNEE Hall
INSTOC
Cornell University
Ithaca, NY 14853

Dan N. Hagedorn
Pacific Northwest Laboratories
Battelle Boulevard
Richland, WA 99352

Dr. Petr Firbas
Institute of Physics of the Earth
Masaryk University Brno
Jecna 29a
612 46 Brno, Czech Republic

Dr. James Hannon
Lawrence Livermore National Laboratory
P.O. Box 808
L-205
Livermore, CA 94550

Dr. Mark D. Fisk
Mission Research Corporation
735 State Street
P.O. Drawer 719
Santa Barbara, CA 93102

Prof. David G. Harkrider
Seismological Laboratory
Division of Geological & Planetary Sciences
California Institute of Technology
Pasadena, CA 91125

Prof Stanley Flatte
Applied Sciences Building
University of California, Santa Cruz
Santa Cruz, CA 95064

Prof. Danny Harvey
CIRES
University of Colorado
Boulder, CO 80309

Prof. Donald Forsyth
Department of Geological Sciences
Brown University
Providence, RI 02912

Prof. Donald V. Helmberger
Seismological Laboratory
Division of Geological & Planetary Sciences
California Institute of Technology
Pasadena, CA 91125

Dr. Art Frankel
U.S. Geological Survey
922 National Center
Reston, VA 22092

Prof. Eugene Herrin
Institute for the Study of Earth and Man
Geophysical Laboratory
Southern Methodist University
Dallas, TX 75275

Prof. Robert B. Herrmann
Department of Earth & Atmospheric Sciences
St. Louis University
St. Louis, MO 63156

Prof. Lane R. Johnson
Seismographic Station
University of California
Berkeley, CA 94720

Prof. Thomas H. Jordan
Department of Earth, Atmospheric &
Planetary Sciences
Massachusetts Institute of Technology
Cambridge, MA 02139

Prof. Alan Kafka
Department of Geology & Geophysics
Boston College
Chestnut Hill, MA 02167

Robert C. Kemerait
ENSCO, Inc.
445 Pineda Court
Melbourne, FL 32940

Dr. Karl Koch
Institute for the Study of Earth and Man
Geophysical Laboratory
Southern Methodist University
Dallas, Tx 75275

Dr. Max Koontz
U.S. Dept. of Energy/DP 5
Forrestal Building
1000 Independence Avenue
Washington, DC 20585

Dr. Richard LaCoss
MIT Lincoln Laboratory, M-200B
P.O. Box 73
Lexington, MA 02173-0073

Dr. Fred K. Lamb
University of Illinois at Urbana-Champaign
Department of Physics
1110 West Green Street
Urbana, IL 61801

Prof. Charles A. Langston
Geosciences Department
403 Deike Building
The Pennsylvania State University
University Park, PA 16802

Jim Lawson, Chief Geophysicist
Oklahoma Geological Survey
Oklahoma Geophysical Observatory
P.O. Box 8
Leonard, OK 74043-0008

Prof. Thorne Lay
Institute of Tectonics
Earth Science Board
University of California, Santa Cruz
Santa Cruz, CA 95064

Dr. William Leith
U.S. Geological Survey
Mail Stop 928
Reston, VA 22092

Mr. James F. Lewkowicz
Phillips Laboratory/GPEH
29 Randolph Road
Hanscom AFB, MA 01731-3010(2 copies)

Mr. Alfred Lieberman
ACDA/VI-OA State Department Building
Room 5726
320-21st Street, NW
Washington, DC 20451

Prof. L. Timothy Long
School of Geophysical Sciences
Georgia Institute of Technology
Atlanta, GA 30332

Dr. Randolph Martin, III
New England Research, Inc.
76 Olcott Drive
White River Junction, VT 05001

Dr. Robert Masse
Denver Federal Building
Box 25046, Mail Stop 967
Denver, CO 80225

Dr. Gary McCartor
Department of Physics
Southern Methodist University
Dallas, TX 75275

Prof. Thomas V. McEvelly
Seismographic Station
University of California
Berkeley, CA 94720

Dr. Art McGarr
U.S. Geological Survey
Mail Stop 977
U.S. Geological Survey
Menlo Park, CA 94025

Dr. Keith L. McLaughlin
S-CUBED
A Division of Maxwell Laboratory
P.O. Box 1620
La Jolla, CA 92038-1620

Stephen Miller & Dr. Alexander Florence
SRI International
333 Ravenswood Avenue
Box AF 116
Menlo Park, CA 94025-3493

Prof. Bernard Minster
IGPP, A-025
Scripps Institute of Oceanography
University of California, San Diego
La Jolla, CA 92093

Prof. Brian J. Mitchell
Department of Earth & Atmospheric Sciences
St. Louis University
St. Louis, MO 63156

Mr. Jack Murphy
S-CUBED
A Division of Maxwell Laboratory
11800 Sunrise Valley Drive, Suite 1212
Reston, VA 22091 (2 Copies)

Dr. Keith K. Nakanishi
Lawrence Livermore National Laboratory
L-025
P.O. Box 808
Livermore, CA 94550

Prof. John A. Orcutt
IGPP, A-025
Scripps Institute of Oceanography
University of California, San Diego
La Jolla, CA 92093

Prof. Jeffrey Park
Kline Geology Laboratory
P.O. Box 6666
New Haven, CT 06511-8130

Dr. Howard Patton
Lawrence Livermore National Laboratory
L-025
P.O. Box 808
Livermore, CA 94550

Dr. Frank Pilotte
HQ AFTAC/TT
1030 South Highway A1A
Patrick AFB, FL 32925-3002

Dr. Jay J. Pulli
Radix Systems, Inc.
201 Perry Parkway
Gaithersburg, MD 20877

Dr. Robert Reinke
ATTN: FCTVTD
Field Command
Defense Nuclear Agency
Kirtland AFB, NM 87115

Prof. Paul G. Richards
Lamont-Doherty Geological Observatory
of Columbia University
Palisades, NY 10964

Mr. Wilmer Rivers
Teledyne Geotech
314 Montgomery Street
Alexandria, VA 22314

Dr. Alan S. Ryall, Jr.
ARPA/NMRO
3701 North Fairfax Drive
Arlington, VA 22203-1714

Dr. Richard Sailor
TASC, Inc.
55 Walkers Brook Drive
Reading, MA 01867

Prof. Charles G. Sammis
Center for Earth Sciences
University of Southern California
University Park
Los Angeles, CA 90089-0741

Prof. Christopher H. Scholz
Lamont-Doherty Geological Observatory
of Columbia University
Palisades, NY 10964

Dr. Susan Schwartz
Institute of Tectonics
1156 High Street
Santa Cruz, CA 95064

Secretary of the Air Force
(SAFRD)
Washington, DC 20330

Office of the Secretary of Defense
DDR&E
Washington, DC 20330

Thomas J. Sereno, Jr.
Science Application Int'l Corp.
10260 Campus Point Drive
San Diego, CA 92121

Dr. Michael Shore
Defense Nuclear Agency/SPSS
6801 Telegraph Road
Alexandria, VA 22310

Dr. Robert Shumway
University of California Davis
Division of Statistics
Davis, CA 95616

Dr. Matthew Sibol
Virginia Tech
Seismological Observatory
4044 Derring Hall
Blacksburg, VA 24061-0420

Prof. David G. Simpson
IRIS, Inc.
1616 North Fort Myer Drive
Suite 1050
Arlington, VA 22209

Donald L. Springer
Lawrence Livermore National Laboratory
L-025
P.O. Box 808
Livermore, CA 94550

Dr. Jeffrey Stevens
S-CUBED
A Division of Maxwell Laboratory
P.O. Box 1620
La Jolla, CA 92038-1620

Lt. Col. Jim Stobie
ATTN: AFOSR/NL
110 Duncan Avenue
Bolling AFB
Washington, DC 20332-0001

Prof. Brian Stump
Institute for the Study of Earth & Man
Geophysical Laboratory
Southern Methodist University
Dallas, TX 75275

Prof. Jeremiah Sullivan
University of Illinois at Urbana-Champaign
Department of Physics
1110 West Green Street
Urbana, IL 61801

Prof. L. Sykes
Lamont-Doherty Geological Observatory
of Columbia University
Palisades, NY 10964

Dr. David Taylor
ENSCO, Inc.
445 Pineda Court
Melbourne, FL 32940

Dr. Steven R. Taylor
Los Alamos National Laboratory
P.O. Box 1663
Mail Stop C335
Los Alamos, NM 87545

Prof. Clifford Thurber
University of Wisconsin-Madison
Department of Geology & Geophysics
1215 West Dayton Street
Madison, WS 53706

Prof. M. Nafi Toksoz
Earth Resources Lab
Massachusetts Institute of Technology
42 Carleton Street
Cambridge, MA 02142

Dr. Larry Turnbull
CIA-OSWR/NED
Washington, DC 20505

Dr. Gregory van der Vink
IRIS, Inc.
1616 North Fort Myer Drive
Suite 1050
Arlington, VA 22209

Dr. Karl Veith
EG&G
5211 Auth Road
Suite 240
Suitland, MD 20746

Prof. Terry C. Wallace
Department of Geosciences
Building #77
University of Arizona
Tuscon, AZ 85721

Phillips Laboratory
ATTN: XPG
29 Randolph Road
Hanscom AFB, MA 01731-3010

Dr. Thomas Weaver
Los Alamos National Laboratory
P.O. Box 1663
Mail Stop C335
Los Alamos, NM 87545

Phillips Laboratory
ATTN: GPE
29 Randolph Road
Hanscom AFB, MA 01731-3010

Dr. William Wortman
Mission Research Corporation
8560 Cinderbed Road
Suite 700
Newington, VA 22122

Phillips Laboratory
ATTN: TSML
5 Wright Street
Hanscom AFB, MA 01731-3004

Prof. Francis T. Wu
Department of Geological Sciences
State University of New York
at Binghamton
Vestal, NY 13901

Phillips Laboratory
ATTN: PL/SUL
3550 Aberdeen Ave SE
Kirtland, NM 87117-5776 (2 copies)

Prof Ru-Shan Wu
University of California, Santa Cruz
Earth Sciences Department
Santa Cruz
, CA 95064

Dr. Michel Bouchon
I.R.I.G.M.-B.P. 68
38402 St. Martin D'Herès
Cedex, FRANCE

ARPA, OASB/Library
3701 North Fairfax Drive
Arlington, VA 22203-1714

Dr. Michel Campillo
Observatoire de Grenoble
I.R.I.G.M.-B.P. 53
38041 Grenoble, FRANCE

HQ DNA
ATTN: Technical Library
Washington, DC 20305

Dr. Kin Yip Chun
Geophysics Division
Physics Department
University of Toronto
Ontario, CANADA

Defense Intelligence Agency
Directorate for Scientific & Technical Intelligence
ATTN: DTIB
Washington, DC 20340-6158

Prof. Hans-Peter Harjes
Institute for Geophysics
Ruhr University/Bochum
P.O. Box 102148
4630 Bochum 1, GERMANY

Defense Technical Information Center
Cameron Station
Alexandria, VA 22314 (2 Copies)

Prof. Eystein Husebye
NTNF/NORSAR
P.O. Box 51
N-2007 Kjeller, NORWAY

TACTEC
Battelle Memorial Institute
505 King Avenue
Columbus, OH 43201 (Final Report)

David Jepsen
Acting Head, Nuclear Monitoring Section
Bureau of Mineral Resources
Geology and Geophysics
G.P.O. Box 378, Canberra, AUSTRALIA

Ms. Eva Johannisson
Senior Research Officer
FOA
S-172 90 Sundbyberg, SWEDEN

Dr. Peter Marshall
Procurement Executive
Ministry of Defense
Blacknest, Brimpton
Reading FG7-FRS, UNITED KINGDOM

Dr. Bernard Massinon, Dr. Pierre Mechler
Societe Radiomana
27 rue Claude Bernard
75005 Paris, FRANCE (2 Copies)

Dr. Svein Mykkeltveit
NTNT/NORSAR
P.O. Box 51
N-2007 Kjeller, NORWAY (3 Copies)

Prof. Keith Priestley
University of Cambridge
Bullard Labs, Dept. of Earth Sciences
Madingley Rise, Madingley Road
Cambridge CB3 0EZ, ENGLAND

Dr. Jorg Schlittenhardt
Federal Institute for Geosciences & Nat'l Res.
Postfach 510153
D-30631 Hannover, GERMANY

Dr. Johannes Schweitzer
Institute of Geophysics
Ruhr University/Bochum
P.O. Box 1102148
4360 Bochum 1, GERMANY

Trust & Verify
VERTIC
Carrara House
20 Embankment Place
London WC2N 6NN, ENGLAND