PNNL-11337

Statistical Algorithms for a Comprehensive Test Ban Treaty Discrimination Framework

D.N. Anderson
K.T. Higbee
N.E. Miller
T. Redgate
A.C. Rohay
D.N. Hagedorn
Technical Editor: N.D. Foote

October 1996

Prepared for the U.S. Department of Energy
under Contract DE-AC06-76RLO 1830

PNNL-11337

DISCLAIMER

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

Acknowledgments

The authors acknowledge the support of Leslie Casey and the Department of Energy's Office of National Security and Non-Proliferation Research and Development (NN-20) for funding this report. The authors assume full responsibility for any inaccuracies or omissions.

Summary

Seismic discrimination is the process of identifying a candidate seismic event as an earthquake or explosion using information from seismic waveform features (seismic discriminants). In the CTBT setting, low energy seismic activity must be detected and identified. A defensible CTBT discrimination decision requires an understanding of false-negative (declaring an event to be an earthquake given it is an explosion) and false-positive (declaring an event to be an explosion given it is an earthquake) rates. These rates are derived from a statistical discrimination framework.

A discrimination framework can be as simple as a single statistical algorithm or it can be a mathematical construct that integrates many different types of statistical algorithms and CTBT technologies (e.g., seismic, hydroacoustic, infrasound and radionuclide). In either case, the result is the identification of an event and the numerical assessment of the accuracy of an identification, that is, false-negative and false-positive rates.

In Anderson et al. (1996), eight statistical discrimination algorithms are evaluated relative to their ability to give results that effectively contribute to a decision process and to be interpretable with physical (seismic) theory. These algorithms can be discrimination frameworks individually or components of a larger framework. The eight algorithms are linear discrimination (LDA), quadratic discrimination (QDA), variably regularized discrimination (VRDA), flexible discrimination (FDA), logistic discrimination, K-th nearest neighbor (KNN), kernel discrimination, and classification and regression trees (CART).

In this report, the performance (accuracy in identifying the source of seismic activity) of these eight algorithms, as applied to regional seismic data, is documented. The discriminants were constructed with an automated approach—phases and energy measurements involved no human analyst interaction. A preliminary velocity model was used to identify phases. Seven seismic stations, at both quiet and noisy locations, were used in the study. The ground-truth data used in this study has some of the characteristics that might initially typify training data in future regions of interest under a CTBT. Based on the findings in Anderson et al. (1996) and this analysis:

- CART is an appropriate algorithm for an automated CTBT setting. CART has many attractive features, such as interpretability, sequential decision rules, the ability to integrate discrete and continuous measurements into a decision, and the ability to manage missing values.

Our analysis supports the assertion that with these eight algorithms, false-negative rates can be as high as 20-25% in poorly characterized regions. It is likely that ground-truth data, similar to the data used in this report, will not be available in future regions of interest. In this case, other approaches need to be initially adopted in order to characterize false-negative and false-positive rates. A thoughtful, technical effort, directed at characterizing the probability structure of discriminants for explosions, is critical to the CTBT ratification and monitoring effort.

Contents

List of Figures

List of Tables

1 Introduction

In this report, the performance (accuracy in identifying the source of seismic activity) of eight statistical discrimination algorithms, as applied to regional seismic data, is documented. The algorithms are linear discrimination (LDA), quadratic discrimination (QDA), variably regularized discrimination (VRDA), flexible discrimination (FDA), logistic discrimination, K-th nearest neighbor (KNN), kernel discrimination, and classification and regression trees (CART). A detailed description of these algorithms can be found in McLachlan (1992).

- LDA assumes that the discriminants from both natural and man-made sources are multivariate normal with equal covariance. The LDA rule assigns a candidate event to the source whose mean is closest to the candidate event, using a Mahalanobis distance measure.

- FDA is a generalization of LDA. FDA reformulates LDA as a least squares linear regression problem and then substitutes non-parametric regression techniques in place of the least squares approach.

- QDA assumes the discriminants from natural and man-made sources are both multivariate normal with possibly unequal covariance. The QDA decision rule is composed of the Mahalanobis distance between a candidate event and each source mean.

- VRDA generalizes LDA and QDA by forming a covariance, for each source, that is the weighted average between LDA and QDA type covariances. Here, the weights change from source to source and are determined from the training data.

- Logistic discrimination models the probability that an event is an earthquake or explosion as a function of seismic discriminants. Logistic discrimination can be viewed as tossing an earthquake/explosion coin where the probability of explosion depends on observed seismic discriminants.

- Kernel discrimination uses non-parametric models of the probability structure of source discriminants to form a likelihood ratio decision rule.

- KNN discrimination assigns a candidate event to the source with the largest number of points in the nearest k points around the candidate event.

- CART is a non-parametric method that seeks to partition a training sample of seismic discriminants into regions, each with a homogeneous event source. The end product of a classification tree is a collection of if–then questions (a decision tree) that can be applied to measured seismic discriminants.

A CTBT discrimination framework/algorithm is evaluated with three general criteria:

1. Does the framework/algorithm give results that effectively contribute to a decision process?

2. Can the framework/algorithm be interpreted with physical (seismic) theory?

3. Can the framework/algorithm accurately identify the source of seismic activity?

In Anderson et al. (1996), the eight statistical discrimination algorithms listed previously were evaluated relative to items 1 and 2. Germane to items 1 and 2, some of the more relevant criteria are:

- Discrete & continuous data (discriminants) allowed — Polarity of first motion is discrete and $m_b - M_s$ is continuous. The framework/algorithm should be applicable to both types of measurements.

- Missing data handled directly — In a CTBT setting, it is very probable that not all seismic measurements will be seen by each monitoring station. The framework/algorithm should be able to adapt to this scenario.

- Easily Understood Algorithm — The framework/algorithm should be sophisticated but lucid. It should be easy to integrate into the CTBT monitoring environment and should not have an excessively complicated structure.

- Easily interpretable results — The framework/algorithm should give an easily interpretable identification of an unknown event, that is, false-negative (declaring an event to be an earthquake given it is an explosion) and false-positive (declaring an event to be an explosion given it is an earthquake) rates, and a class membership score.

- Works well with small group sizes — In many regional CTBT settings few ground-truth measurements may be available for explosions. The framework/algorithm should accommodate solutions to this problem.

- Few assumptions — A framework/algorithm with a large number of assumptions is usually less applicable. Assumptions may compensate for the lack of ground-truth data. Assumptions can also be potential complications or points of disagreement in a ratification and monitoring setting.

- Sequential decision rules — The framework/algorithm should require the computation of measurements only when necessary for a decision.

Table 1: Discrimination Algorithm Comparison

Selected Issues/Criteria	LDA	QDA	VRDA	FDA	Logistic	Kernel	KNN	CART
Allows Both Discrete & Continuous Data	N	N	N		Y			Y
Missing Data Handled Directly	N	N	N	N	N	N	N	Y
Easily Understood Algorithm	Y			N	Y		Y	Y
Easily Interpretable Results	Y	Y	Y	Y	Y	Y	Y	Y
Works Well with Small Group Sizes	Y	N	Y		N	N	N	N
Few Assumptions	N	N	N	Y	Y	Y	Y	Y
Sequential Decision Rules	N	N	N	N	N	N	N	Y
Accommodates Multi-Modal Sources	N	N	N	Y		Y	Y	Y
Automatic Parameter Selection	Y	Y	Y	N		N	N	N

Y means yes and N means no.
A blank entry indicates one of the following:
1) A middle answer between Y and N is appropriate.
2) Either Y or N is appropriate, depending on setting or parameters.
3) A definitive answer is unavailable.

- Accommodates multi-modal sources — The framework/algorithm should have the ability to accurately differentiate between sources characterized by seismic measurement data with two or more modes (that is, data with two or more local distribution peaks or density maxima).

- Parameters are automatically selected — The framework/algorithm should have the ability to automatically estimate or select the values of the model parameters used by the framework.

A concise summary, in terms of these criteria, is presented in Table 1. All of these criteria are strengths. In the table, "Y" means yes, indicating that the algorithm possesses the strength. A blank indicates that either a definitive answer is unavailable, or an answer depends on the application. As seen in Table 1, no single algorithm is ideal for every criterion — There is often a trade-off between the criteria. One example of such a trade-off can be seen by comparing the entries in Table 1 for the criteria of *Few Assumptions* and *Works Well with Small Group Sizes*.

Among the many seismic and statistical concerns involved with monitoring a CTBT is the construction of a discrimination framework. The capability or accuracy of the framework must be understood and characterized with a concerted technical effort. In this report, the accuracy (item 3) of eight statistical discrimination algorithms is evaluated. A description of the seismic data used in this evaluation is presented in Section 2. Section 3 compares the different discrimination algorithms. Section 4 discusses the importance of characterizing the probability structure of discriminants from nuclear explosions. Summary comments are presented in Section 5.

2 Data Description

The ground-truth data used in this study have some of the characteristics that might initially typify training data in future regions of interest under a CTBT. The discriminants were calculated with an automated approach—phases and energy measurements involved no human analyst interaction. A preliminary velocity model was used to identify phases. Seven seismic stations, at both quiet and noisy locations, were used in the study. These stations had the most complete data and most of the time windows contained at least some discernible signal. The station locations are: Albuquerque, NM (ANMO); Columbia, CA (CMB); Goldstone, CA (GSC); Isadora, CA (ISA); Pasadena, CA (PAS); Pinyon Flats, CA (PFO); and Tucson, AZ (TUC). Station-to-event distance varied from 200 to 800 km. The data set was built from events occurring during the period 1990 to 1995 at the Nevada Test Site (NTS) located in southern Nevada. The seismic data were gathered from the IRIS Data Management Center located at the University of Washington. These data consisted 80 earthquakes, 20 explosions and 2 cavity collapses for a total of 102 events. Event magnitudes ranged from 2.8 to 5.6.

The Seismic Analysis Code (SAC) developed at Lawrence Livermore National Laboratory was used to make the phase and noise measurements for the events. Earthquake locations and origin times from the National Earthquake Information Center (NEIC) were used. A preliminary velocity model was uniformly applied to all events to pick the phase arrival times that were used to make the energy measurements of signal and noise. No refinements to the arrival times were made on an event-by-event basis. Velocities of 6.1 and 3.6 km/sec were used to predict the arrival times of the P_g and L_g phases, respectively. The locations used had depth and epicentral errors in the range from 10 to 20 km, relative to precise local array locations. A total of nine band-pass filters were used on each waveform: 0.5-1, 1-2, 2-4, 4-6, 6-8, 4-8, 1-3, 3-6, and 3-9 Hz. For each bandpass filter, several RMS energy measurements were made:

(a). a measurement of noise, $\pm$ 15 seconds around the origin time (used for noise removal processing and to create signal-to-noise ratios).

(b). measurements of the first 10 seconds after model estimated arrival times for the P_g and L_g waves (used to create signal-to-noise ratios).

(c). measurements of P_g and L_g phases.

P_n was not included in this work, because the time window for P_n is short and therefore dependent on precise location information, not typically available in a new region of interest.

Clean energy measurements were created by extracting the noise from measured signals in items (b) and (c) using the formula:

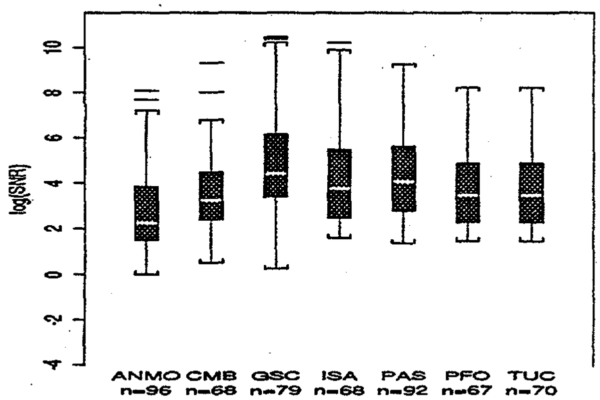
$$\text{cleaned energy signal} = \sqrt{(\text{measured signal})^2 - (\text{noise})^2}. \tag{1}$$

In order to create signal-to-noise ratios (SNR) specifically for each of the P_g and L_g measurements, the cleaned signals in item (b) were divided by the noise in item (a). The SNRs were inputs to the creation of weights applied at the time of averaging. Initially, four different candidate data sets were created. Each candidate data set contained station-averaged phase ratio discriminants. Two methods of averaging were investigated: arithmetic and geometric. Additionally, averages were constructed with and without weights. Optimal discrimination was obtained by using the logarithm of a weighted geometric average. Specifically, the phase ratio discriminants were formed with

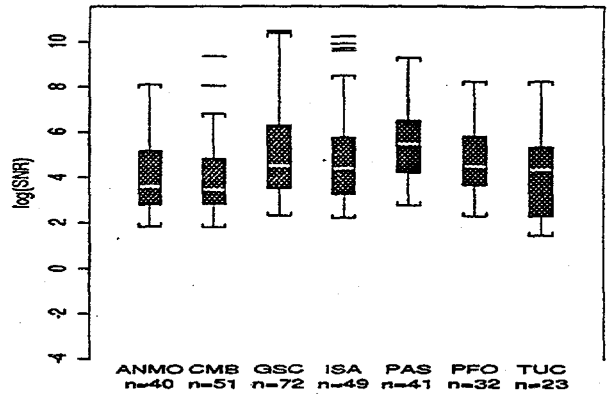
$$\text{phase ratio discriminant} = \frac{1}{W}\sum_{i=1}^{7} w_i \log\left(\frac{P_{g_i}}{L_{g_i}}\right) \tag{2}$$

where i indicates station, $w_i = \log(\text{SNR } P_{g_i})/\log(\text{SNR } L_{g_i})$, and $W = \sum_{i=1}^{7} w_i$. Here, SNR P_{g_i} is the SNR using the cleaned signal in item (b) for P_g and the noise in item (a); SNR L_{g_i} is the SNR using the cleaned signal in item (b) for L_g and the noise in item (a). Phase ratio discriminants were created for each of the frequency bands. Only events with P_g and L_g log(SNR) greater than 1.0 were included in the average. Box plots of log(SNR) and log(SNR) given log(SNR) > 1.0 for the P_g phase are given in Figure 1. The number of events (n) used to construct each box plot is given (note that none of the stations recorded all 102 events). This figure illustrates the impact that the log(SNR) > 1.0 constraint has on these data. The final data set consisted of phase ratio discriminants, at frequency bands 0.5-1, 1-2, 2-4, 4-6, 6-8, 4-8, 1-3, 3-6, and 3-9 Hz, for 71 earthquakes and 19 explosions.

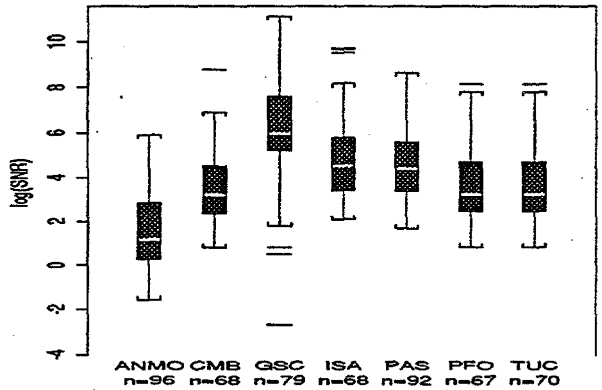
A subset of these phase ratio discriminants was identified and used in the analysis. In general, as the number of discriminants used in an algorithm increases, discriminative ability also increases. However, after a certain point, adding more discriminants can actually degrade performance. If several discriminants are strongly correlated, then including all of them in the construction of a discrimination algorithm will tend to fit random noise rather than seismic structure. This property is analogous to over fitting a linear regression model. The subset of phase ratio discriminants was identified in two different ways:



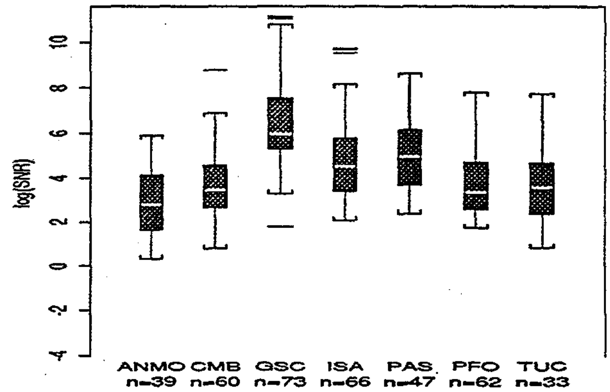
(a) Frequency Band = 0.5-1 Hz.



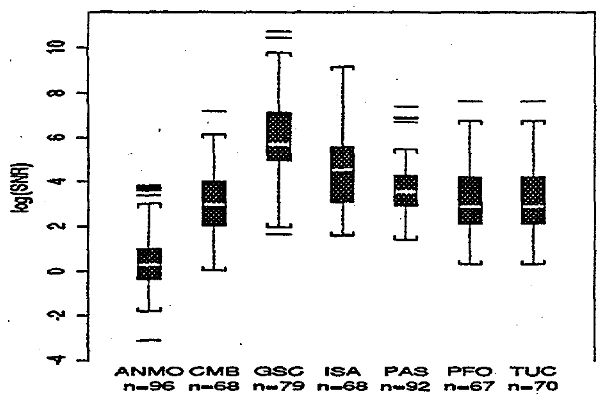
(b) Frequency Band = 0.5-1 Hz and log(SNR) > 1.0.



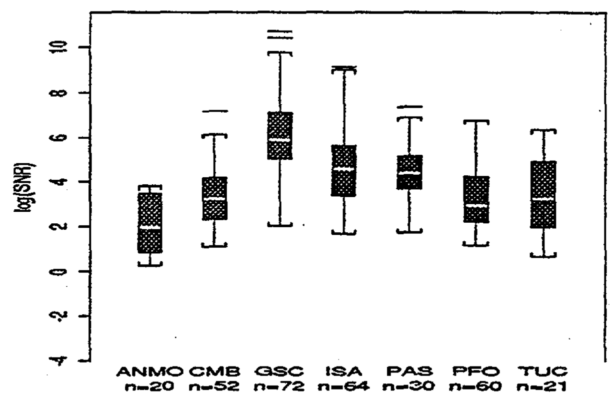
(c) Frequency Band = 2-4 Hz.



(d) Frequency Band = 2-4 Hz and log(SNR) > 1.0.



(e) Frequency Band = 4-6.



(f) Frequency Band = 4-6 and log(SNR) > 1.0.

Figure 1: Box plots of log(SNR) and log(SNR) given log(SNR) > 1.0 for the P_g phase. The value of n is the number of events used to construct the box plot.

- Stepwise regression for all frequency bands with the addition of dummy noise variables.
- All-possible-subsets method for all frequency bands for subsets of sizes 2, 3, 4, and 5.

The Stepwise procedure in SAS was used with the added dummy noise variables (SAS, 1989). The idea was to create noise variables and add them to the list of potential variables (see Miller (1990)). When the stepwise procedure selected a noise variable as a discriminant, the phase ratio discriminants identified in the previous step were chosen as the optimal discriminants. In the second variable selection method, all possible pairs, three-tuples four-tuples and five-tuples, were also created; the corresponding linear discrimination models were fit; and the model with the best cross-validated overall error rate (percent correct decision) was selected. The all-possible subsets-method, while computationally more challenging, is preferred over the stepwise procedure. Both methods were examined as in some cases the all-possible-subset method would not be feasible. These two variable selection techniques, however, yielded the same three-phase ratio discriminants at 0.5-1, 2-4, and 4-6 Hz frequencies as important variables for this data set. To summarize, the data set used to evaluate the discrimination algorithms consisted of 71 earthquakes and 19 explosions with phase ratio discriminants constructed at band widths of 0.5-1, 2-4, and 4-6 Hz.

3 Comparison of Discrimination Algorithms

The eight algorithms used in this analysis were chosen to represent a reasonable range of the numerous algorithms available in the statistics community. The data set was partitioned into 10 training data sets and 10 prediction data sets. Each training set comprised 90% of the original data set and each prediction data set made up the remaining 10%. Each of the eight algorithms were applied to each of the 10 training data sets. The rates used to assess performance were combined over the 10 prediction data sets. Figure 2 summarizes the performance of the eight algorithms. Table 2 gives estimates of the false-negative and false-positive rates for this analysis. While the false-negative rate for CART is 32%, we note that the full modeling features (non-orthogonal cuts, see Breiman et al. (1984)) of CART were not available for this analysis. Using all of the features available to the CART method would greatly improve the false-negative error rate. The decision boundaries for each algorithm were constructed with equal source prior probabilities. Error costs were not incorporated into the construction of decision boundaries. In terms of the overall error rate in Figure 2, all of the eight algorithms appear to perform about the same. In terms of the false-negative and false-positive rates in Table 2, some algorithms perform better than others. For example, CART and Logistic perform best in terms of a false-positive rate. However, the kernel algorithm performs best in terms of a false-negative rate. The impact of these

CART

True Source	Identified as Ex	Identified as Eq
Ex	13	6
Eq	4	67

Overall Error Rate = 11.1%

KNN

True Source	Identified as Ex	Identified as Eq
Ex	16	3
Eq	8	63

Overall Error Rate = 12.2%

Logistic

True Source	Identified as Ex	Identified as Eq
Ex	14	5
Eq	4	67

Overall Error Rate = 10%

LDA

True Source	Identified as Ex	Identified as Eq
Ex	17	2
Eq	8	63

Overall Error Rate = 11.1%

QDA

True Source	Identified as Ex	Identified as Eq
Ex	15	4
Eq	9	62

Overall Error Rate = 14.4%

FDA

True Source	Identified as Ex	Identified as Eq
Ex	15	4
Eq	7	64

Overall Error Rate = 12.2%

Kernel

True Source	Identified as Ex	Identified as Eq
Ex	18	1
Eq	11	60

Overall Error Rate = 13.3%

VRDA

True Source	Identified as Ex	Identified as Eq
Ex	15	4
Eq	8	63

Overall Error Rate = 13.3%

Figure 2: Summary Tables (Confusion Matrices) and Overall Error Rates.

Algorithm	Pr(false-negative)	Pr(false-positive)
CART	32%	6%
Logistic	26%	6%
FDA	21%	10%
VRDA	21%	11%
QDA	21%	13%
KNN	16%	11%
LDA	11%	11%
Kernel	5%	15%

Table 2: Estimated False-Negative and False-Positive Rates.

two errors on a decision can be integrated into an algorithm by adopting a decision theory approach to discrimination. This approach integrates prior probabilities and error costs into a discrimination decision.

4 Probability Structure of Discriminants from Nuclear Explosions

The CTBT ratification process, as well as implementation of a CTBT monitoring system, will generate vigorous discussion of the probabilities of

- declaring an event to be an earthquake, given it is an explosion (false-negative)
- declaring an event to be an explosion, given it is an earthquake (false-positive).

For example, evidence to support the accuracy of a monitoring system will be prerequisite to the imposition of an on-site inspection directed at a CTBT signatory. The capability to identify low energy seismic activity is assessed with these probabilities. Both of these probabilities are essential components in calculating the probability that

- an event is an explosion given it is declared to be an earthquake
- an event is an earthquake given it is declared to be an earthquake.

If an event is an explosion and is declared to be an earthquake, then a serious error has been made. If an event is an earthquake and is declared to be an earthquake, then a correct and cost effective decision has been made. Discussions and efforts to resolve these and other analogous probabilities are critical to the successful implementation of a CTBT.

In general, discrimination involves two basic activities. Seismic discriminants that evidentially identify the source of a seismic event are selected, and secondly, these discriminants are integrated into an appropriate statistical discrimination framework. These activities cannot be completed without some characterization of the multivariate probability structure of the discriminants for each seismic source, that is, earthquakes and explosions.

Because the availability of data from nuclear explosions is limited to a small number of regions in the world, the CTBT research community has discussed constructing discrimination strategies independent of any ground-truth explosion data. An outlier approach proposed by Gray et al. (1996) is based on the general idea of characterizing the regional seismic data from naturally occurring events and mining activities, and then declaring as suspicious those future events that are not statistically similar to this population. This approach precludes any estimate of the false-negative error rate. Adopting any algorithm or framework to monitor low energy seismic activity is tenuous without an understanding of the associated false-negative error rate. Without some type of characterization of the regional seismic properties of nuclear explosions, it is not possible to quantify the false-negative error. The outlier approach only provides an estimate of the probability of incorrectly calling an earthquake an outlier.

A simple graphical example illustrates the problem involved when no information is used to characterize nuclear explosions for a region. Figure 3 shows four scenarios. In the first, Figure 3(a), only earthquakes are characterized and a decision boundary is set that allows 5% of earthquakes to be called explosions. Because explosions are not characterized in any way in Figure 3(a), no indication is given of the error rate in labeling explosions as earthquakes. Figures 3(b), 3(c), and 3(d) indicate three (out of an infinite number) possible arrangements based on different explosion characteristics. In Figure 3(b) a 1% error rate of labeling explosions as earthquakes is obtained while retaining a 5% error rate in labeling earthquakes as explosions. For Figure 3(c), the error rate of labeling explosions as earthquakes is 5%, and for Figure 3(d), the rate is 25%. The false-negative rates in Table 2 are comparable to Figure 3(d). This simple illustration points out that an outlier approach is not able to quantify the chances of labeling an explosion as an earthquake, unless information concerning explosions is available. Figure 3 uses a simple one-sided decision boundary on only one normally distributed measurement or feature. However, the same idea holds for two-sided decision boundaries, higher dimensional problems, and non-normal (including mixture) distributions.

Several approaches can be adopted to characterize nuclear explosions in regions of interest. Some of these strategies are:

1. Develop seismic theory that will permit transportation of nuclear explosion data from existing weapon test sites to future regions of interest.

2. Use data from regional simulations of nuclear explosions.

3. Design and conduct chemical calibration explosions.

4. Use data from chemical explosions of opportunity, such as mining explosions.

5. Base regional nuclear explosion characterization on expert opinion.

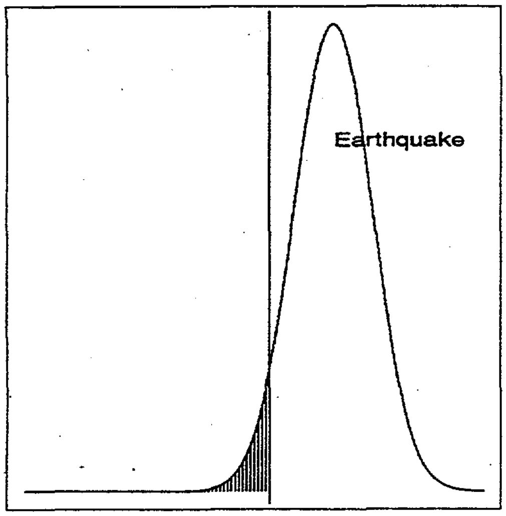
6. Use any combination of these strategies.

The characterization of the probability structure of explosions and earthquakes is critical research for CTBT discrimination. Without this research,

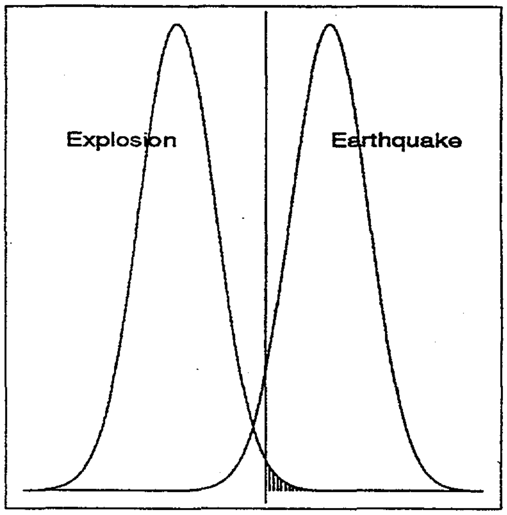
- optimal discriminants cannot be identified for regions of interest
- an optimal discrimination framework cannot be constructed for regions of interest
- performance capabilities of a regional discrimination framework cannot be assessed.

A thoughtful, technical effort to characterize nuclear explosions must be undertaken in support of the CTBT. The outlier approach to discrimination, while expedient, appears to cloud a discussion of the issues demonstrated in this section. The CTBT research community should seek the best possible resolution of the probabilities of

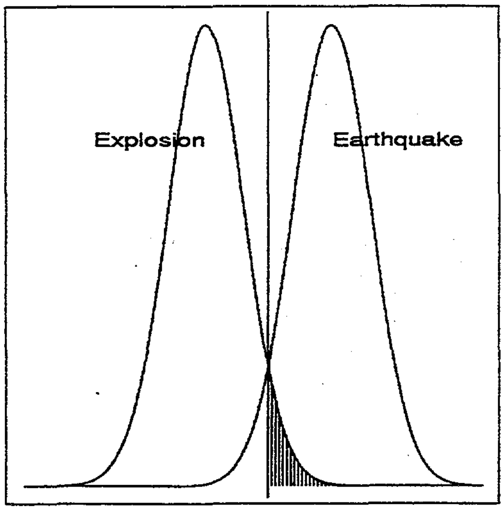
- declaring an event to be an earthquake given it is an explosion
- declaring an event to be an explosion given it is an earthquake.



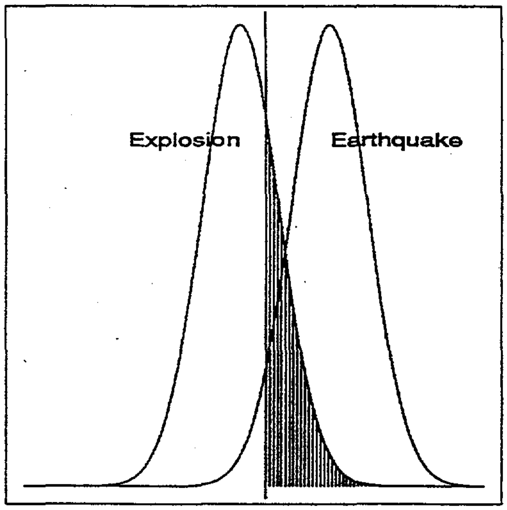
(a) Decision boundary giving 5% error rate of calling an earthquake an explosion. Only earthquake distribution information is used. The error rate of calling an explosion an earthquake is not available.



(b) Earthquake distribution and decision boundary of Figure 3(a) (giving 5% error rate of calling an earthquake an explosion) and explosion distribution giving a 1% error rate of calling an explosion an earthquake.



(c) Earthquake distribution and decision boundary of Figure 3(a) (giving 5% error rate of calling an earthquake an explosion) and explosion distribution giving a 5% error rate of calling an explosion an earthquake.



(d) Earthquake distribution and decision boundary of Figure 3(a) (giving 5% error rate of calling an earthquake an explosion) and explosion distribution giving a 25% error rate of calling an explosion an earthquake.

Figure 3: Error Rates for Earthquake and Various Explosion Populations.

5 Conclusions

The analysis in this report, in combination with the conclusions from Anderson et al. (1996), indicate:

- CART is an appropriate algorithm for an automated CTBT setting. CART has many attractive features, such as interpretability, sequential decision rules, the ability to integrate discrete and continuous measurements into a decision, and the ability to manage missing values.

- All of the algorithms in this report can provide corroborative evidence to support a decision. These algorithms can be integrated into a seismic analyst tool box or a discrimination framework.

- Many technologies will contribute to a decision on the disposition of a seismic event. The discrimination algorithms in this report can contribute, across all monitoring technologies, to the CTBT discrimination problem.

- In this analysis, overall error rates are near 10%. Further, this analysis supports the assertion that false-negative rates can be as high as 20-25% in poorly characterized regions.

- A thoughtful, technical effort directed at characterizing the probability structure of discriminants for explosions is critical to the CTBT ratification and monitoring effort.

References

Anderson, D. N., Anderson, K. K., Hagedorn, D. N., Higbee, K. T., Miller, N. E., Redgate, T., and Rohay, A. C. (1996). Statistical classification methods applied to seismic discrimination. Technical Report PNNL-11192, Pacific Northwest National Laboratory, Richland, WA.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). *Classification and Regression Trees.* Chapman & Hall, New York, NY.

Gray, H., Sain, S., Frawley, W., and Woodward, W. (1996). On detecting outliers in mixed populations. Technical Report PL-TR-96-2049, Phillips Laboratory, Hanscom Air Force Base.

McLachlan, G. J. (1992). *Discriminant Analysis and Statistical Pattern Recognition.* John Wiley & Sons, Inc., New York, NY.

Miller, A. (1990). *Subset Selection in Regression.* Chapman & Hall, New York, NY.

SAS (1989). *SAS/STAT User's Guide, Version 6, Fourth Edition, Volume 2.* SAS Institute, Inc., Cary, NC.

Distribution

No. Of Copies		Copies	
2	DCI/ACIS Arms Control Intelligence Staff Washington, DC 20505 J. Filson L.S. Turnbull	5	Phillips Laboratory Geophysics Branch Earth Sciences Div. PL/GPE 29 Randolph Road Hanscom AFB, MA 01731-3010 A. Dainty K. Kadinsky-Cade R.J. Jih J.F. Lewkowicz D.T. Reiter
2	Defense Special Weapons Agency Nuclear Treaty Programs 1901 N. Moore St., Suite 609 Arlington, VA 22209 R. Alewine S.R. Bratt	5	Arms Control & Disarmament Agency 3020 21st St., NW Washington, DC 20451 R. Cockerham, Rm 5741 P. Cordon, Rm 5499 S. Day, Rm 5741 M. Dreicer, Rm 5643 O.J. Sheaks, Rm 5741
6	Air Force Technical Application Center 1030 South Highway A1A Patrick AFB, FL 32935-3002 N.C. Anderholm, DOB C. McBrearty, TN F.F. Pilotte, TT G.H. Rothe III, TTR D.R. Russell, TTR R. Schult, TTR	5	U.S. Department of Energy Forrestal Building 1000 Independence Avenue, SW Washington, DC 20585 L.A. Casey, NN-20 L. Evanson, NN-40.2A S. Rudnick, NN-20 K.F. Veith, NN-20 R. Waldron, NN-20
1	Center for Monitoring Research AFTAC/TT/CMR 1300 N. 17th Street, Suite 1450 Arlington, VA 22209-2308 R. Blandford		

No. of Copies

18 Lawrence Livermore National
Laboratory
P.O. Box 808
Livermore, CA 94551

M. Denny, MS L-205
W. Dunlop, MS L-200
L.A. Glenn, MS L-200
P. Goldstein, MS L-205
W.J. Hannon Jr., MS L-205
P. Harben, MS L-208
D.B. Harris, MS L-205
F.E. Heuze, MS L-200
S. Larsen, MS L-208
K.M. Mayeda, MS L-205
K.K. Nakaniski, MS L-205
J. Rambo, MS L-200
A.S. Ryall, MS L-205
A.T. Smith, MS L-205
P. Sokkappa, MS L-200
J. Sweeney, MS L-208
W.R. Walter, MS L-205
J. Zucca, MS L-205

9 Los Alamos National Laboratory
P.O. Box 1663
Los Alamos, NM 87545

F.N. App, MS F659
W.M. Brunish, MS F659
A.H. Cogbill, MS F659
H. Hartse, MS C335
H.J. Patton, MS C335
D.C. Pearson, MS C335
B.W. Stump, MS C335
S.R. Taylor, MS C335
T.A. Weaver, MS C335

Copies

8 Sandia National Laboratory
P.O. Box 5800
Albuquerque, NM 87185

D.R. Breding, MS 0655
E.P. Chael, MS 1159
J.P. Claassen, MS 0655
M.W. Edenburn, MS 0425
P.B. Herrington, MS 0655
R.G. Keyser, MS 1138
L.S. Walker, MS 0979
C.J. Young, MS 0750

1 U.S. Department of the Interior
U.S. Geological Survey
Reston, VA 22092

W. Leith, MS 928

2 The IRIS Consortium
1616 N. Ft. Myer Dr., Suite 1050
Arlington, VA 22209

D. Simpson
G. Van Der Vink

1 FCDSWA
1680 Texas Street SE
Kirkland AFB, NM 87117-5669

R.E. Reinke

Onsite
No. of Copies

33 Pacific Northwest National
Laboratory
P.O. Box 999
Richland, WA 99352

D.N. Anderson, K5-12 (10)
K.K. Anderson, K5-12
D.M. Boyd, K5-02
D.L. Eggett, K5-12
N.D. Foote, K7-10
T.R. Fox, K6-48
D.N. Hagedorn, K5-12 (5)
K.T. Higbee, K5-12
N.E. Miller, K5-12
B.A. Pulsipher, K5-12
P.E. Redgate, K5-12
A.C. Rohay, K9-48
R.A. Warner, K6-48
Information Release, K1-06(7)