PNNL-11579

Pacific Northwest National Laboratory

Operated by Battelle for the U.S. Department of Energy

Quadratic Negative Evidence Discrimination

D.N. Anderson
T. Redgate
K.K. Anderson
A.C. Rohay
Technical Editor: F.M. Ryan

May 1997

PNNL-11579

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

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC06-76RLO 1830

Printed in the United States of America

Available to DOE and DOE contractors from the Office of Scientific and Technical Information, P.O. Box 62, Oak Ridge, TN 37831; prices available from (615) 576-8401.

Available to the public from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161

This document was printed on recycled paper.

DISCLAIMER

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

PNNL-11579

Quadratic Negative Evidence Discrimination

D.N. Anderson
T. Redgate
K.K. Anderson
A.C. Rohay
Technical Editor: F.M. Ryan

May 1997

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

Summary

This paper develops regional discrimination methods which use information inherent in phase magnitudes that are unmeasurable due to small signal amplitudes and/or high noise levels. The methods are enhancements to teleseismic techniques proposed by [Elvers, 1974], and are extended to regional discrimination. Events observed at teleseismic distances are effectively identified with the M_s vs m_b discriminant because relative to the pressure wave energy (m_b) of an event, an earthquake generates more shear wave energy (M_s) than does an explosion. For some teleseismic events, the M_s magnitude is difficult to measure and is known only to be below a threshold. With M_s unmeasurable, the M_s vs m_b discriminant cannot be formed. However, if the M_s threshold is sufficiently small relative to a measured m_b, then the event is still likely to be an explosion.[Elvers, 1974] proposed techniques to quantify these concepts.

The methods presented in this report are developed for a single seismic station, and make use of empirical evidence in the regional L_g vs P_g discriminant (see [Pomeroy et al., 1983]). The L_g vs P_g discriminant is analogous to the teleseismic M_s vs m_b discriminant. The methods developed in this paper can also be applied to other regional discriminants.

- We show how to estimate the parameters associated with these methods, using both censored (signal equal to or less than the noise level) and complete data.
- We develop the equations for station-specific false-negative and false-positive error rates. These error rates depend on the required minimum signal-to-noise ratio, and can be adjusted, within limits, to desired levels.
- We show how to combine the source identifications from a network of stations into a single source identification via a scoring framework. The error rate calculations for the network scoring framework are also discussed.

This work is in support of Los Alamos National Laboratory (LANL) and Lawrence Livermore National Laboratory (LLNL), and their efforts to characterize the regional seismicity of Western China and the Middle East and North Africa. The methods presented in this report will be applied to the regional data currently being gathered by LANL and LLNL.

Contents

List of Figures

List of Tables

Acknowledgments

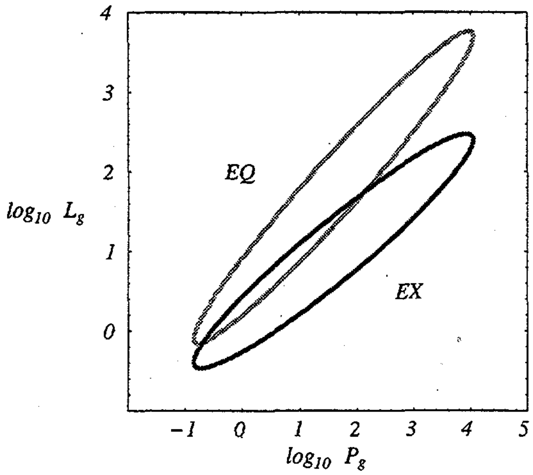
The authors acknowledge the support of Leslie Casey and the Department of Energy's Office of National Security and Non-Proliferation Research and Development (NN-20) for funding this work. We express our appreciation to Dr. Steve Taylor, Los Alamos National Laboratory for seed ideas and valuable technical discussions on the negative evidence concept. The authors assume full responsibility for any errors or omissions.

1 Introduction

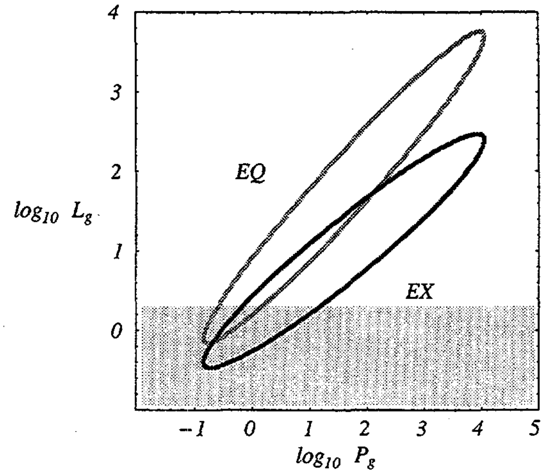
This report presents methods of regional seismic discrimination that could potentially be used when the signal of one phase of a spectral ratio discriminant is buried in noise, thus preventing the construction of the discriminant. The methods presented are enhancements of the M_s vs m_b negative evidence discrimination concept, introduced in [Elvers, 1974]. The enhancements are a direct application of quadratic and linear discrimination analysis (QDA and LDA) [McLachlan, 1992]. For clarity, we develop the negative evidence methods in this paper by jointly modeling phase magnitudes, as opposed to a univariate model of a discriminant. Events observed at teleseismic distances are effectively identified with the M_s vs m_b discriminant because relative to the pressure wave energy (m_b) of an event, an earthquake has more shear wave energy (M_s) than an explosion. For some teleseismic events, the M_s magnitude is difficult to measure and is known only to be below a threshold. With M_s unmeasurable, the M_s vs m_b discriminant cannot be formed. However, if the M_s threshold is sufficiently small relative to a measured m_b, then the event is still likely to be an explosion.

A regional analogue to the M_s vs m_b discriminant is the L_g vs P_g discriminant (see [Pomeroy et al., 1983] for a lucid discussion of regional discriminants). As in the teleseismic setting, the L_g magnitude is difficult to measure for some regional events (especially explosions) and is known only to be below a threshold. With L_g unmeasurable, the L_g vs P_g discriminant cannot be formed. However, if the L_g threshold is sufficiently small relative to a measured P_g, then the event is still likely to be an explosion. This report quantifies these ideas in the context of monitoring compliance with the Comprehensive Nuclear Test Ban Treaty (CTBT).

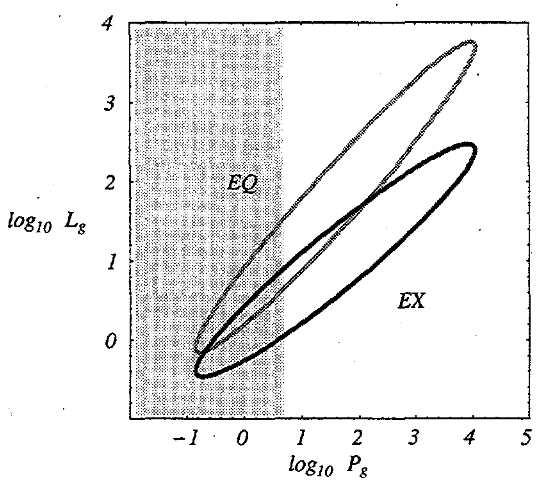
To clarify, the case when both L_g and P_g are measurable is illustrated in Figure 1(a). Here, classical statistical discrimination methods can be used to construct a decision line to identify seismic events. Figure 1(b) illustrates the case when P_g can be measured but L_g is known only to be below a threshold. Figure 1(c) illustrates the case when L_g can be measured but P_g is known only to be below a threshold. Figure 1(d) illustrates the case when both L_g and P_g are buried in noise. In Section 2, the [Elvers, 1974] approach is mildly enhanced with QDA to construct quadratic negative evidence discrimination (QNED) methods. Linear negative evidence discrimination (LNED) is a special case of QNED. LNED can also be constructed



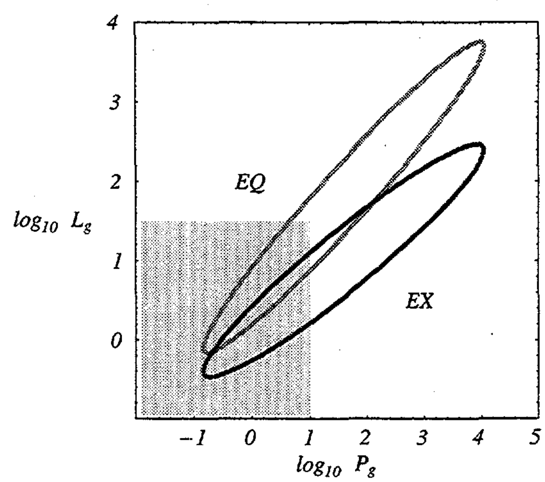
(a) P_g and L_g Measured.



(b) P_g Measured and L_g Buried in Noise.



(c) P_g Buried in Noise and L_g Measured.



(d) P_g Buried in Noise and L_g Buried in Noise.

Figure 1: Earthquake and Explosion 95% Ellipsoids with Shaded Regions that Illustrate P_g and L_g Censoring Cases.

by direct application of LDA to the [Elvers, 1974] approach. Section 2 also develops the equations to compute QNED false-negative and false-positive error rates. A false-negative error occurs when an explosion is erroneously identified as an earthquake (a missed CTBT violation). A false-positive error occurs when an earthquake is erroneously identified as an explosion. In Section 3, QNED is illustrated with fabricated regional explosion and earthquake models. Section 4 is a discussion of future work necessary to apply QNED to a regional CTBT setting. This work is in support of Los Alamos National Laboratory (LANL) and Lawrence Livermore National Laboratory (LLNL) and their efforts to characterize the regional seismicity of Western China [Hartse et al., 1996] and the Middle East and North Africa. A mature QNED method will be applied by Pacific Northwest National Laboratory (PNNL) to these regional data.

2 Quadratic Negative Evidence Discrimination (QNED)

In this section, we develop negative evidence methods in terms of quadratic discrimination analysis. A regional discrimination rule can be constructed with the magnitudes $Y = \log\ L_g$ and $X = \log\ P_g$. It is generally accepted in the seismic community that the probability structure of X, Y can be adequately modeled with a bivariate normal distribution. One criterion of good measurements for X and Y requires an adequate signal-to-noise ratio (SNR), where noise energy is measured immediately preceding event arrival. A pre-event noise magnitude, $Z = \log_{10}(\mathit{noise\,energy})$, can also be adequately modeled with a normal distribution. In terms of magnitudes, the SNR for Y is $Y - Z$, and an SNR $> \log_{10}\kappa$ indicates a good measurement of Y. Here, κ is the minimum acceptable SNR for both X and Y. For event i and station j, we denote the explosion and earthquake magnitudes with the random variable $\mathbf{W}'_{ij} = (X_{ij}, Y_{ij}, Z_{ij})$, and we model $\mathbf{W}_{ij}$ as independent multivariate normal (MVN), with parameters μ_{Exj}, Σ_{Exj} and $\mu_{Eq_j}, \Sigma_{Eq_j}$ respectively (LNED is a special case of QNED with $\Sigma_{Exj} = \Sigma_{Eq_j} = \Sigma_j$). Measured values of $\mathbf{W}_{ij}$ are denoted $\mathbf{w}'_{ij} = (x_{ij}, y_{ij}, z_{ij})$. The normal density functions are denoted $f_{\mathbf{W}}(x, y, z; \mu_{Exj}, \Sigma_{Exj})$ and $f_{\mathbf{W}}(x, y, z; \mu_{Eq_j}, \Sigma_{Eq_j})$.

A weak observed signal-to-noise ratio, $y_{ij} - z_{ij} \leq \log_{10}\kappa$, indicates that the L_g magnitude is buried in noise, that is, y_{ij} is unavailable. In this case, a magnitude threshold can be computed: since $y_{ij} - z_{ij} \leq \log_{10} \kappa$, we can claim that $y_{ij} \leq z_{ij} + \log_{10} \kappa$. Valuable information is lost by not including magnitude thresholds as well as good magnitude measurements in regional discrimination analysis. As noted in Section 1, if X can be accurately measured for an event but Y is known only to be small, then the event is likely to be an explosion, even though a L_g vs P_g discriminant cannot be formed. Combining the discrimination information from m stations is discussed in Section 2.4.

2.1 Likelihood Ratio Discrimination

For an event with unknown source, good measurements of $\mathbf{W}$ can be used to classify the event as explosion or earthquake. In this case, an approach to source identification is to declare the event an explosion if

$$f_{\mathbf{W}}(x_{ij}, y_{ij}, z_{ij}; \mu_{Exj}, \Sigma_{Exj}) > \lambda_1 f_{\mathbf{W}}(x_{ij}, y_{ij}, z_{ij}; \mu_{Eqj}, \Sigma_{Eqj}). \tag{1}$$

In words, this rule is: call the event an explosion if the likelihood of observing x_{ij}, y_{ij}, z_{ij} is appreciably larger for an explosion than for an earthquake, where "appreciably" is defined by the factor λ_1. If the measurements from an event are x_{ij} and z_{ij} with y_{ij} below threshold ($y_{ij} \leq z_{ij} + \log_{10} \kappa$), then the L_g magnitude is buried in noise. In this case we call the event an explosion if the likelihood of observing x_{ij}, z_{ij} and $y_{ij} \leq z_{ij} + \log_{10} \kappa$ is appreciably larger for an explosion than for an earthquake; that is, if

$$\int_{\nu=-\infty}^{z_{ij}+\log_{10}\kappa} f_{\mathbf{W}}(x_{ij}, \nu, z_{ij}; \mu_{Exj}, \Sigma_{Exj})\, d\nu > \lambda_2 \int_{\nu=-\infty}^{z_{ij}+\log_{10}\kappa} f_{\mathbf{W}}(x_{ij}, \nu, z_{ij}; \mu_{Eqj}, \Sigma_{Eqj})\, d\nu. \tag{2}$$

Likelihood theory for censored measurements is well developed in reliability methods (see [Nelson, 1982]). If the measurements are y_{ij} and z_{ij} with x_{ij} below threshold ($x_{ij} \leq z_{ij} + \log_{10} \kappa$), then the P_g magnitude is buried in noise, and we call the event an explosion if

$$\int_{\upsilon=-\infty}^{z_{ij}+\log_{10}\kappa} f_{\mathbf{W}}(\upsilon, y_{ij}, z_{ij}; \mu_{Exj}, \Sigma_{Exj})\, d\upsilon > \lambda_3 \int_{\upsilon=-\infty}^{z_{ij}+\log_{10}\kappa} f_{\mathbf{W}}(\upsilon, y_{ij}, z_{ij}; \mu_{Eqj}, \Sigma_{Eqj})\, d\upsilon. \tag{3}$$

Finally, if the only information available from station j is $x_{ij} \leq z_{ij} + \log_{10}\kappa$, and $y_{ij} \leq z_{ij} + \log_{10}\kappa, z_{ij}$, then we call the event an explosion if

$$\int_{\upsilon=-\infty}^{z_{ij}+\log_{10}\kappa} \int_{\nu=-\infty}^{z_{ij}+\log_{10}\kappa} f\mathbf{w}(\upsilon,\nu,z_{ij};\mu_{Exj},\Sigma_{Exj})\,d\nu\,d\upsilon > \lambda_4 \int_{\upsilon=-\infty}^{z_{ij}+\log_{10}\kappa} \int_{\nu=-\infty}^{z_{ij}+\log_{10}\kappa} f\mathbf{w}(\upsilon,\nu,z_{ij};\mu_{Eqj},\Sigma_{Eqj})\,d\nu\,d\upsilon. \quad (4)$$

The values of $\lambda_1, \ldots, \lambda_4$ and κ are determined by the desired false-negative and false-positive rates (operating characteristics) associated with these four discrimination rules, as explained below.

2.2 Operating Characteristics

In this section, we derive the probability that an earthquake is incorrectly identified as an explosion and the probability that an explosion is incorrectly identified to be an earthquake. Respectively, these probabilities are denoted $P_j(\tilde{Ex} \mid Eq)$ and $P_j(\tilde{Eq} \mid Ex)$, and are the false-positive and false-negative rates of a station discrimination rule (the declared identification of a source is denoted by $\tilde{Ex}$ or $\tilde{Eq}$). Let C_1 denote the case "P_g and L_g measured" (x_{ij}, y_{ij}, z_{ij}), C_2 denote the case "P_g measured and L_g buried in noise" $(x_{ij}, y_{ij} \leq z_{ij} + \log_{10}\kappa, z_{ij})$, C_3 denote the case "P_g buried in noise and L_g measured" $(x_{ij} \leq z_{ij} + \log_{10}\kappa, y_{ij}, z_{ij})$, and C_4 denote the case "P_g and L_g buried in noise" $(x_{ij} \leq z_{ij} + \log_{10}\kappa, y_{ij} \leq z_{ij} + \log_{10}\kappa, z_{ij})$. The key to defining the operating characteristics of QNED is in constructing decision rules that will be applied to future events of unknown source. These decision rules are essentially regions in X, Y, Z space that classify events as earthquakes or explosions. Since we have four cases $C_1, \ldots, C_4$, we need four decision rules. For case C_1, let the explosion decision rule $\Omega_{1_j}(\lambda_1, \kappa)$ be

$$\{x, y, z \text{ such that } x > z + \log_{10}\kappa, y > z + \log_{10}\kappa, f\mathbf{w}(x,y,z;\mu_{Exj},\Sigma_{Exj}) > \lambda_1 f\mathbf{w}(x,y,z;\mu_{Eqj},\Sigma_{Eqj})\}.$$

Note that the earthquake decision rule is the complement of $\Omega_{1_j}(\lambda_1, \kappa)$. Let the explosion decision rule $\Omega_{2_j}(\lambda_2, \kappa)$ be

$$\{x, y, z \text{ such that } x > z + \log_{10} \kappa, y \leq z + \log_{10} \kappa,$$
$$\int_{\nu=-\infty}^{z+\log_{10}\kappa} f\mathbf{w}(x, \nu, z; \mu_{Exj}, \Sigma_{Exj})\, d\nu > \lambda_2 \int_{\nu=-\infty}^{z+\log_{10}\kappa} f\mathbf{w}(x, \nu, z; \mu_{Eqj}, \Sigma_{Eqj})\, d\nu\}.$$

Similarly define $\Omega_{3_j}(\lambda_3, \kappa)$ and $\Omega_{4_j}(\lambda_4, \kappa)$. Note that the decision rules $\Omega_{1_j}(\lambda_1, \kappa), \ldots, \Omega_{4_j}(\lambda_4, \kappa)$ are equivalent to the events $C_1 \cap \tilde{Ex}, \ldots, C_4 \cap \tilde{Ex}$, where for example $C_1 \cap \tilde{Ex}$ denotes the intersection of Case C_1 conditions and an event identified as an explosion. Direct application of basic probability rules gives

$$P_j(\tilde{Ex} \mid Eq) =$$
$$P_j(C_1 \cap \tilde{Ex} \mid Eq) + P_j(C_2 \cap \tilde{Ex} \mid Eq) + P_j(C_3 \cap \tilde{Ex} \mid Eq) + P_j(C_4 \cap \tilde{Ex} \mid Eq), \quad (5)$$

as the false-positive rate (incorrectly identifying an earthquake as an explosion); and

$$P_j(\tilde{Eq} \mid Ex) =$$
$$P_j(C_1 \cap \tilde{Eq} \mid Ex) + P_j(C_2 \cap \tilde{Eq} \mid Ex) + P_j(C_3 \cap \tilde{Eq} \mid Ex) + P_j(C_4 \cap \tilde{Eq} \mid Ex). \quad (6)$$

as the false-negative rate (incorrectly identifying an explosion as an earthquake). In designing a QNED discrimination rule, a compromise among the desired case-specific $(C_1, \ldots, C_4)$ false-negative and false-positive rates defines the values of $\lambda_1, \ldots, \lambda_4$. In other words, fixed values of $\lambda_1, \ldots, \lambda_4$ define the overall false-positive and false-negative rates, $P_j(\tilde{Ex} \mid Eq)$ and $P_j(\tilde{Eq} \mid Ex)$, by specifying the regions $\Omega_{1_j}(\lambda_1, \kappa), \ldots, \Omega_{4_j}(\lambda_4, \kappa)$. Constructing a discrimination rule such that $P_j(\tilde{Ex} \mid Eq)$ and $P_j(\tilde{Eq} \mid Ex)$ are both as small as possible is thus a matter of compromise among the false-negative and false-positive rates associated with the measurement cases $C_1, \ldots, C_4$. This procedure is described in Section 3.

Each of the component probabilities in Equations 5 and 6 can be computed numerically. For example, since $C_2 \cap \tilde{Ex}$ is equivalent to $\Omega_{2_j}(\lambda_2, \kappa)$, we can evaluate $P_j(C_2 \cap \tilde{Ex} \mid Eq)$ by integrating over the region $\Omega_{2_j}(\lambda_2, \kappa)$:

$$P_j(C_2 \cap \tilde{Ex} \mid Eq) = \iiint\limits_{v,\nu,\omega \in \Omega_{2_j}(\lambda_2,\kappa)} f_{\mathbf{W}}(v, \nu, \omega; \mu_{Eqj}, \Sigma_{Eq_j}) dv\, d\nu\, d\omega.$$

This computation, while numerically tractable, is quite time consuming. As an alternative, the operating characteristics of QNED can be obtained through Monte Carlo simulation methods. This approach has advantages, since $P_j(C_2 \cap \tilde{Ex} \mid Eq)$ can be obtained by computing the component probabilities $P_j(C_2 \mid Eq)$ and $P_j(\tilde{Ex} \mid C_2 \cap Eq)$: that is, $P_j(C_2 \cap \tilde{Ex} \mid Eq) = P_j(C_2 \mid Eq)P_j(\tilde{Ex} \mid C_2 \cap Eq)$.

For the Monte Carlo analysis, a large number of values of X, Y, Z can be simulated with an earthquake model. For fixed κ, $P_j(C_2 \mid Eq)$ is estimated to be the proportion of cases in the simulated data where $x > z + \log_{10} \kappa \cap y \leq z + \log_{10} \kappa$. For fixed κ and λ_2, $P_j(\tilde{Ex} \mid C_2 \cap Eq)$ is estimated to be the proportion of C_2 case earthquakes that are identified as explosions. As shown in Section 3, writing the false-negative and false-positive rates in terms of the $C_1, \ldots, C_4$ error rates provides some latitude in designing the operating characteristics of a station. We note that this Monte Carlo analysis to determine QNED operating characteristics is conducted off-line and in advance of the operational use of QNED. Once the Monte Carlo analysis has been completed, QNED can be applied rapidly in a CTBT setting. For an operational setting, efficient QNED algorithms are available for Equations 1 through 4. The QNED methods presented in this report were coded by the authors in Mathematica [Wolfram, 1996].

2.3 Parameter Estimation

The development of QNED to this point assumes that the parameters μ_{Exj}, Σ_{Exj} and $\mu_{Eq_j}, \Sigma_{Eq_j}$ are known or well estimated prior to operational implementation. In this section, we describe how to estimate these parameters using a combination of censored and good measurements. As with events observed in real-time, ground truth data (GTD_{Ex}, GTD_{Eq}) is composed of events where one of the cases $C_1, \ldots, C_4$ holds. The likelihood equation for ground truth data is found by multiplying the individual $C_1, \ldots, C_4$ likelihoods for both earthquakes and explosions (Nelson 1982):

$$
\begin{aligned}
&L\left(\mu_{Exj}, \Sigma_{Exj}, \mu_{Eq_j}, \Sigma_{Eq_j}\right) = \\
&\prod_{i \in C_1 \cap GTD_{Ex}} f\mathbf{w}(x_{ij}, y_{ij}, z_{ij}; \mu_{Exj}, \Sigma_{Exj}) \times \prod_{i \in C_2 \cap GTD_{Ex}} \int_{\nu=-\infty}^{z_{ij}+\log_{10}\kappa} f\mathbf{w}(x_{ij}, \nu, z_{ij}; \mu_{Exj}, \Sigma_{Exj})\, d\nu \times \\
&\qquad \prod_{i \in C_3 \cap GTD_{Ex}} \int_{v=-\infty}^{z_{ij}+\log_{10}\kappa} f\mathbf{w}(v, y_{ij}, z_{ij}; \mu_{Exj}, \Sigma_{Exj})\, dv \times \\
&\qquad \prod_{i \in C_4 \cap GTD_{Ex}} \int_{v=-\infty}^{z_{ij}+\log_{10}\kappa} \int_{\nu=-\infty}^{z_{ij}+\log_{10}\kappa} f\mathbf{w}(v, \nu, z_{ij}; \mu_{Exj}, \Sigma_{Exj})\, d\nu\, dv \times \\
&\prod_{i \in C_1 \cap GTD_{Eq}} f\mathbf{w}(x_{ij}, y_{ij}, z_{ij}; \mu_{Eq_j}, \Sigma_{Eq_j}) \times \prod_{i \in C_2 \cap GTD_{Eq}} \int_{\nu=-\infty}^{z_{ij}+\log_{10}\kappa} f\mathbf{w}(x_{ij}, \nu, z_{ij}; \mu_{Eq_j}, \Sigma_{Eq_j})\, d\nu \times \\
&\qquad \prod_{i \in C_3 \cap GTD_{Eq}} \int_{v=-\infty}^{z_{ij}+\log_{10}\kappa} f\mathbf{w}(v, y_{ij}, z_{ij}; \mu_{Eq_j}, \Sigma_{Eq_j})\, dv \times \\
&\qquad \prod_{i \in C_4 \cap GTD_{Eq}} \int_{v=-\infty}^{z_{ij}+\log_{10}\kappa} \int_{\nu=-\infty}^{z_{ij}+\log_{10}\kappa} f\mathbf{w}(v, \nu, z_{ij}; \mu_{Eq_j}, \Sigma_{Eq_j})\, d\nu\, dv. \qquad (7)
\end{aligned}
$$

Parameter estimates are obtained by choosing values for $\mu_{Exj}, \Sigma_{Exj}, \mu_{Eq_j}, \Sigma_{Eq_j}$ that maximize Equation 7 (For the LNED model, we maximize Equation 7 subject to the constraint $\Sigma_{Exj} = \Sigma_{Eq_j} = \Sigma_j$). This exercise is computationally intensive, but fortunately can be done off-line from an operational QNED framework.

2.4 Network QNED

The measurements from a network of m stations can be synthesized into a single discrimination rule. There are several approaches to network discrimination, each having merit. One approach, appealing for its simplicity, is to count the number of stations that individually identify an event to be an explosion. If the count is greater than some predetermined value n^*, then the network event identification would be explosion. The procedure to choose a value for n^* is described below.

Let

$$I_{ij} = \begin{cases} 1 & \text{if station } j \text{ identifies event } i \text{ to be an explosion} \\ 0 & \text{if station } j \text{ identifies event } i \text{ to be an earthquake} \end{cases}$$

and define

$$N_i = \sum_j^m I_{ij}.$$

In other words, N_i is the number of stations that identify event i as an explosion. For station j, the false-positive rate is $P(I_{ij} = 1 \mid Eq) = P_j(\tilde{Ex} \mid Eq)$ (Equation 5). If the operating characteristics of all stations in a network are such that the false-positive rate $\approx q$ for all stations, then $N_i \mid Eq$ can be approximately modeled as a binomial(m,q) random variable. If the station false-positive rates are substantially different from each other, then the probability structure of $N_i \mid Eq$ must be computed by considering all configurations of stations in which $N_i \mid Eq = k$; $k = 1, 2, \ldots, m$. Good quality algorithms are readily available for these calculations [Wolfram, 1996].

Under a CTBT, a seismic event is assumed a priori to be an earthquake. To declare explosion, we must have persuasive network evidence that the event was an explosion. In terms of statistical inference methods, the CTBT null and alternative hypotheses are

$$\begin{cases} H_0 : \text{event was an earthquake} \quad (N_i \text{ is binomial}(m,q)) \\ H_a : \text{event was an explosion} \end{cases}$$

We call an event an explosion if $N_i \geq n^*$. The critical value n^* is determined from the tolerable network false-positive rate (α), that is, choose n^* so that $P(N_i \geq n^* \mid H_0) \leq \alpha$ (In statistical inference terms, α is the Type I error rate). We can also choose a value of n^* with a false-negative error rate constraint; that is, $P(N_i < n^* \mid H_a) \leq \beta$ (In statistical inference terms, β is the Type II error rate).

3 Example: Constructing a Station Discrimination Rule

This section illustrates how to derive the operating characteristics (error rates) of QNED with simulated regional explosion and earthquake data. The QNED error rates for these models are derived from a Monte Carlo simulation of 20,000 replicates of $Y = \log\ L_g$ and $X = \log\ P_g$ from both earthquake and explosion models (10,000 from each). These data are used to compute the false-negative and false-positive rates developed in Section 2. In this example, the probabilities $P_j(\tilde{Eq} \mid Ex \cap C_4)$ and $P_j(\tilde{Ex} \mid Eq \cap C_4)$ are set to 1 and 0 respectively — we assume under the CTBT that a station source identification will be earthquake when the station cannot measure either P_g or L_g. The models used in the simulation are illustrated in Figure 2. The parameters for $\mathbf{W}_{ij}{}' = (X_{ij}, Y_{ij}, Z_{ij})$ for these

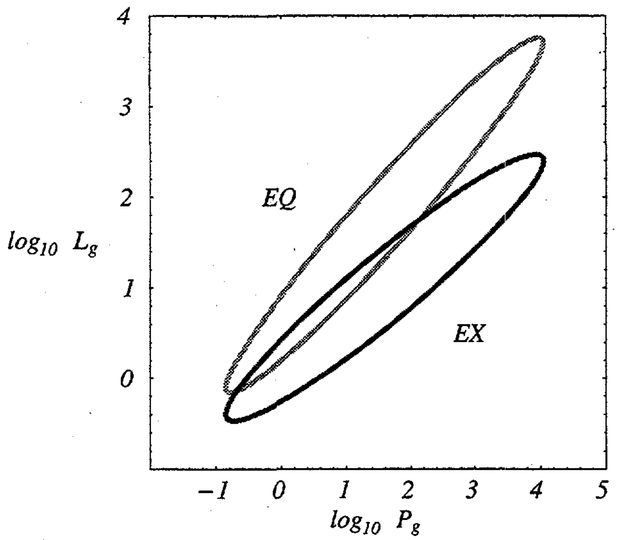


Figure 2: 95% Ellipsoids of Earthquake and Explosion Models.

models are

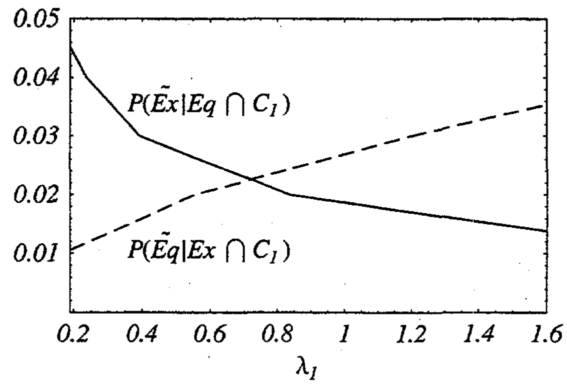
$$\mu_{Exj} = \begin{pmatrix} 1.6 \\ 1.0 \\ 0.75 \end{pmatrix} \quad \Sigma_{Exj} = \begin{pmatrix} 1.0 & 0.57 & 0.33 \\ 0.57 & 0.36 & 0.198 \\ 0.33 & 0.198 & 0.3025 \end{pmatrix}$$
$$\mu_{Eqj} = \begin{pmatrix} 1.6 \\ 1.8 \\ 0.75 \end{pmatrix} \quad \Sigma_{Exj} = \begin{pmatrix} 1.0 & 0.776 & 0.33 \\ 0.776 & 0.64 & 0.264 \\ 0.33 & 0.264 & 0.3025 \end{pmatrix}. \tag{8}$$

An arbitrary value of $\kappa = 2$ was used as a minimal SNR requirement for a good measure of X_{ij} and Y_{ij}. For example, if $x_{ij} - z_{ij} \leq \log_{10} 2$, then x_{ij} is buried in noise. In Section 2, we showed that the component probabilities of Equations 5 and 6 could be broken into smaller component probabilities, by applying basic probability rules. Estimates of these component probabilities are given in Table 1 and Figure 3.

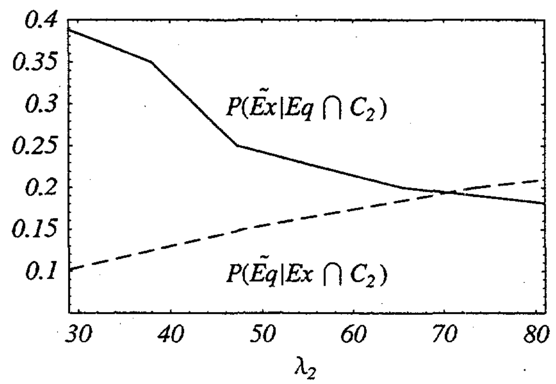
Table 1: Monte Carlo Error Rate Estimates ($P_j(C_i \mid Eq)$ and $P_j(C_i \mid Ex)$, $i = 1, \ldots, 4$).

$P_j(C_1 \mid Ex)$	=	0.4455	$P_j(C_1 \mid Eq)$	=	0.7502
$P_j(C_2 \mid Ex)$	=	0.3056	$P_j(C_2 \mid Eq)$	=	0.0043
$P_j(C_3 \mid Ex)$	=	0.0104	$P_j(C_3 \mid Eq)$	=	0.1234
$P_j(C_4 \mid Ex)$	=	0.2385	$P_j(C_4 \mid Eq)$	=	0.1221

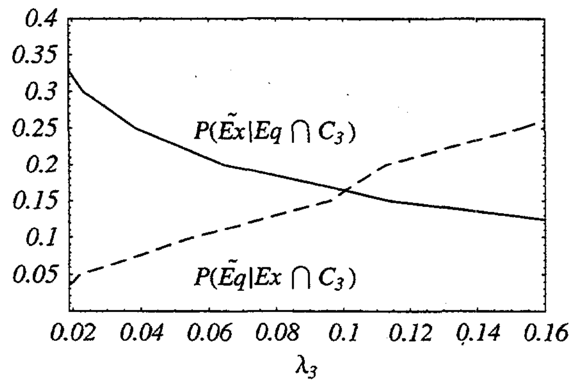
Most technical papers on seismic discrimination summarize operating characteristics only with the probabilities $P_j(\tilde{Ex} \mid C_1 \cap Eq)$ and $P_j(\tilde{Eq} \mid C_1 \cap Ex)$, that is, the false-positive and false-negative error rates associated with the measurement case C_1. QNED provides a realistic assessment of error rates for all measurement cases, $C_1, \ldots, C_4$. In the CTBT setting, we want to minimize the false-negative $(\tilde{Eq} \mid Ex)$ error rate without unduly increasing the false-positive $(\tilde{Ex} \mid Eq)$ error rate. We choose values of $\lambda_1, \lambda_2, \lambda_3$ to achieve a reasonable compromise between these two error rates. From Figure 3(a), a value of $\lambda_1 = 0.2$ gives $P_j(\tilde{Ex} \mid Eq \cap C_1) \approx 0.045$ and $P_j(\tilde{Eq} \mid Ex \cap C_1) \approx 0.010$. From Figure 3(b), a value of $\lambda_2 = 30$ gives $P_j(\tilde{Ex} \mid Eq \cap C_2) \approx 0.38$ and $P_j(\tilde{Eq} \mid Ex \cap C_2) \approx 0.10$. From Figure 3(c), a



(a) Case C_1 False-Negative (- - -) and False-Positive (—) rates.



(b) Case C_2 False-Negative (- - -) and False-Positive (—) rates.



(c) Case C_3 False-Negative (- - -) and False-Positive (—) rates.

Figure 3: Monte Carlo Estimates of False-Negative ($P(\tilde{Eq} \mid Ex \cap C_\ell)$) and False-Positive ($P(\tilde{Ex} \mid Eq \cap C_\ell)$) rates.

value of $\lambda_3 = 0.02$ gives $P_j(\tilde{Ex} \mid Eq \cap C_3) \approx 0.32$ and $P_j(\tilde{Eq} \mid Ex \cap C_3) \approx 0.04$. We determine the aggregate false-positive and false-negative error rates by summing the probabilities determined by these chosen values of $\lambda_1, \lambda_2, \lambda_3$, for all four case conditions. Neglecting case C_4 for the moment, the overall error rates for the first three cases (the first three terms in Equations 5 and 6) are:

$$P_j(C_1 \cap \tilde{Ex} \mid Eq) + P_j(C_2 \cap \tilde{Ex} \mid Eq) + P_j(C_3 \cap \tilde{Ex} \mid Eq) \approx \\ 0.7502 \times 0.045 + 0.0043 \times 0.38 + 0.1234 \times 0.32 = 0.075$$

$$P_j(C_1 \cap \tilde{Eq} \mid Ex) + P_j(C_2 \cap \tilde{Eq} \mid Ex) + P_j(C_3 \cap \tilde{Eq} \mid Ex) \approx \\ 0.4455 \times 0.010 + 0.3056 \times 0.10 + 0.0104 \times 0.04 = 0.035.$$

From Table 1, we have $P_j(C_4 \mid Eq) = 0.1221$ and $P_j(C_4 \mid Ex) = 0.2385$. If we assume that $P_j(\tilde{Eq} \mid Ex \cap C_4) = 1$ and $P_j(\tilde{Ex} \mid Eq \cap C_4) = 0$ (because of the assumption that under the CTBT a station source identification will always be earthquake if X and Y cannot be measured), then

$$P_j(\tilde{Ex} \mid Eq) \approx 0.7502 \times 0.045 + 0.0043 \times 0.38 + 0.1234 \times 0.32 + 0.1221 \times 0 = 0.075$$
$$P_j(\tilde{Eq} \mid Ex) \approx 0.4455 \times 0.010 + 0.3056 \times 0.10 + 0.0104 \times 0.04 + 0.2385 \times 1 = 0.274.$$

This is the result of the a priori distribution of earthquakes and explosions shown in Figure 2. The classification of all C_4 events as earthquakes results in a misclassification of case C_4 explosions. Note that case C_4 events may have magnitudes below the threshold of interest in a CTBT setting. The aggregate error rates $P_j(\tilde{Ex} \mid Eq)$ and $P_j(\tilde{Eq} \mid Ex)$ can be adjusted with different values of $\lambda_1, \ldots, \lambda_3$ — choosing these values is the process of constructing a station discrimination rule.

As another example, we choose larger values of $\lambda_1, \lambda_2, \lambda_3$. This will serve to illustrate the range of possible error rates associated with the models in this section. From Figure 3(a), a value of $\lambda_1 = 1.6$ gives $P_j(\tilde{Ex} \mid Eq \cap C_1) \approx 0.014$ and $P_j(\tilde{Eq} \mid Ex \cap C_1) \approx 0.035$. From Figure 3(b), a value of $\lambda_2 = 80$ gives $P_j(\tilde{Ex} \mid Eq \cap C_2) \approx 0.17$ and $P_j(\tilde{Eq} \mid Ex \cap C_2) \approx 0.21$. From

Figure 3(c), a value of $\lambda_3 = 0.16$ gives $P_j(\tilde{Ex} \mid Eq \cap C_3) \approx 0.12$ and $P_j(\tilde{Eq} \mid Ex \cap C_3) \approx 0.26$. Neglecting case C_4, the overall error rates for the first three cases are:

$$P_j(C_1 \cap \tilde{Ex} \mid Eq) + P_j(C_2 \cap \tilde{Ex} \mid Eq) + P_j(C_3 \cap \tilde{Ex} \mid Eq) \approx$$
$$0.7502 \times 0.014 + 0.0043 \times 0.17 + 0.1234 \times 0.12 = 0.026$$

$$P_j(C_1 \cap \tilde{Eq} \mid Ex) + P_j(C_2 \cap \tilde{Eq} \mid Ex) + P_j(C_3 \cap \tilde{Eq} \mid Ex) \approx$$
$$0.4455 \times 0.035 + 0.3056 \times 0.21 + 0.0104 \times 0.26 = 0.082.$$

In other words, based on the models assumed in this example and those cases $(C_1, \ldots, C_3)$ where there is at least one measurement, adjusting the λ values in the station discrimination rule will result in a false-positive error rate varying from 2.6% to 7.5% and a false-negative error rate that varies from 8.2% to 3.5%. The authors believe that the operating characteristics computed in this section are realistic and possible in a regional (CTBT) setting.

4 Future Developments

As a mature discrimination method, QNED should be applicable to log spectral ratios (discriminants). Denote two different L_g, P_g discriminants as $D_{ij} = X_{ij} - Y_{ij}$ and $D^{\star}{}_{ij} = X^{\star}{}_{ij} - Y^{\star}{}_{ij}$ (d_{ij} and $d^{\star}{}_{ij}$ denote measured values of D_{ij} and $D^{\star}{}_{ij}$). These discriminants could be constructed from any combination of low and high frequency filters applied to the L_g, P_g phases. Methods of removing source and path effects from such discriminants are currently being researched at LANL. If $x_{ij} - z_{ij} > \log_{10}\kappa$ and $y_{ij} - z_{ij} \leq \log_{10}\kappa$ then X_{ij} can be measured and Y_{ij} is buried in noise. In this case, we can claim that $d_{ij} = x_{ij} - y_{ij} \geq x_{ij} - z_{ij} - \log_{10}\kappa$. A similar argument holds for the case where Y_{ij} is measured and X_{ij} is buried in noise. When both X_{ij} and Y_{ij} are buried in noise, we can only claim that $-\infty < d_{ij} < \infty$ and source identification must be based on $D^{\star}{}_{ij}$ alone. With these inequalities, the discriminants D_{ij} and $D^{\star}{}_{ij}$ can be placed into a QNED framework. QNED research will continue with:

- the development of a QNED framework based on two discriminants, D_{ij} and $D^{\star}{}_{ij}$,
- the development of a network discrimination rule that accurately reflects the likelihood contribution to a decision from all cases of station observations,
- the integration of source and path correction methods into the QNED theory, and
- the application of mature QNED methods to Western China, Middle East and North Africa regional data.

References

[Elvers, 1974] Elvers, E. (1974). Seismic Event Identification by Negative Evidence. *Bulletin of the Seismological Society of America*, 64(6):1671–1683.

[Hartse et al., 1996] Hartse, H., Taylor, S., Phillips, W., and Randall, G. (1996). A Preliminary Study of Regional Seismic Discrimination in Central Asia with Emphasis on Western China. Technical Report LAUR-96-2002, Los Alamos National Laboratory, Geophysics Group,EES-3.

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

[Nelson, 1982] Nelson, W. (1982). *Applied Life Data Analysis.* John Wiley & Sons, Inc.

[Pomeroy et al., 1983] Pomeroy, P., Best, W., and McEvilly, T. (1983). Test Ban Treaty Verification with Regional Data-A Review. *Bulletin of the Seismological Society of America*, 72(6):S89–S129.

[Wolfram, 1996] Wolfram, S. (1996). *The Mathematica Book.* Wolfram Media/Cambridge University Press, 3rd edition.

Distribution

No. Of Copies
External

2 U.S. Department of Energy
Forrestal Building
1000 Independence Avenue, SW
Washington, DC 20585

L.A. Casey, NN-20 (2)

2 Arms Control & Disarmament Agency
3020 21st St., NW
Room 5741
Washington, DC 20451

M. Dreicer
R.J. Morrow

2 Air Force Technical Application Center
1030 South Highway A1A
Patrick AFB, FL 32925-3002

F.F. Pilotte, TT
D.R. Russell, TTR

1 R.R. Blanford
AFTAC/TT/CMR
Center for Monitoring Research
1300 North 17th Street, Suite 1450
Arlington, VA 22209-2308

No. Of Copies

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

M.D. Denny, MS L-205
W. Dunlop, MS L-200
P. Goldstein, MS L-205
D.B. Harris, MS L-205
K.M. Mayeda, MS L-205
K.K. Nakanishi, MS L-205
P. Sokkappa, MS L-195
W.R. Walter, MS L-205
J.J. Zucca, MS L-205

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

W.M. Brunish, MS F659
A.H. Cogbill, MS C335
H.E. Hartse, MS C335
H.J. Patton, MS C335
D.C. Pearson, MS C335
W.S. Phillips, MS C335
G.E. Randall, MS D443
B.W. Stump, MS C335
S.R. Taylor, MS C335

1 S.R. Bratt
Nuclear Treaty Programs Office
1901 N. Moore St./Suite 609
Arlington, VA 22209

No. Of Copies

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

D.R. Breding, MS 0655
E.P. Chael, MS 1189
J.P. Claassen, MS 0655
P.B. Herrington, MS 0655
R.G. Keyser, MS 1138
L.S. Walker, MS 0979
C.J. Young, MS 0750

Internal

28 Pacific Northwest Laboratory

P.O. Box 999
Richland, WA 99352
D.N. Anderson, K5-12 (8)
K.K. Anderson, K5-12
D.M. Boyd, K5-02
D.N. Hagedorn, K5-12 (5)
K.T. Higbee, K5-12
B.A. Pulsipher, K5-12
P.E. Redgate, K5-12
A.C. Rohay, K9-48
F.M. Ryan, K5-12
R.A. Warner, K6-40
Information Release, K1-06 (7)