PNNL-11359
UC-705

Statistical Evaluation of CTBT Regional Seismic Monitoring

K. K. Anderson

September 23, 1996

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

Pacific Northwest National Laboratory
Operated for the U.S. Department of Energy
by Battelle

Battelle

PNNL-11359

DISCLAIMER

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

PNNL-11359
UC-705

Statistical Evaluation of CTBT Regional Seismic Monitoring

K. K. Anderson

September 23, 1996

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

Pacific Northwest National Laboratory
Richland, Washington 99352

Executive Summary

This report summarizes FY96 statistical research to support the regionalization efforts of the U.S. Department of Energy's Comprehensive Test Ban Treaty research program. We consider the problems of estimation and uncertainties associated with three aspects of regional seismic monitoring:

- determining station detection capability,
- developing magnitude scales, and
- modeling location accuracy.

The indirect method of estimating a station's regional detection capability is based on measuring the seismic noise level and estimating the signal-to-noise ratio required for detection. A methodology is presented that produces better estimates of the detection capability and the uncertainties of the estimates.

Magnitude scaling is calibrating regional station magnitudes to the worldwide teleseismic m_b scale. It involves estimating the station-dependent distance corrections in the relationship $m_b = a + b\log(amp) + c\log(dist)$, where amp is the frequency-adjusted P-phase amplitude and $dist$ is the great circle distance between the station and the regional event. A magnitude scaling methodology is presented that estimates the three scaling coefficients for multiple regional stations.

In both of these problems, statistically incorporating the uncertainty in worldwide teleseismic m_b magnitudes of observed events improves the estimation and uncertainty analyses.

Seismic network modeling commonly measures location accuracy in terms of the area of the 90% error ellipse associated with the estimated location of an event. An improved Monte Carlo approach is proposed to estimate the network location accuracy at either a hypothesized location or on a global grid of locations.

Acknowledgments

We acknowledge the support of Leslie Casey and the U.S. Department of Energy's Office of National Security and Nonproliferation Research and Development (NN-20) for funding this research as part of the Comprehensive Test Ban Treaty Research and Development Program, ST482C. The research was done at the Pacific Northwest National Laboratory, a multiprogram laboratory operated for the U.S. Department of Energy under Contract DE-AC06-76RLO 1830.

Contents

Tables

Figures

1 Introduction

A global seismic monitoring system under a Comprehensive Test Ban Treaty (CTBT) is judged by its capability to detect, locate, and identify suspicious seismic events. *Performance measures* are those statistical objects that describe these capabilities. *Performance criteria* are the thresholds derived from the overall monitoring system goals, against which the evaluated performance measures are compared. This report proposes statistical objects for performance measurement of detection and location, a continuation of the research of Anderson and Anderson (1995). A statistical methodology for calibrating regional station magnitudes to the worldwide teleseismic m_b scale is also proposed.

Anderson and Anderson (1995) developed a hybrid approach combining a direct estimation and a recurrence curve method to determine regional network detection capability, based on a comparison to a reference seismic system and seismicity. Here we consider an indirect method of estimating an individual station's regional detection capability, based on measuring the seismic noise level and estimating the signal-to-noise ratio required for detection (Ringdal, 1975). We present in Section 2 a methodology that produces better estimates of the detection capability and the uncertainties of the estimates.

In Section 3, we present a methodology for calibrating regional station magnitudes to the worldwide teleseismic m_b scale. The methodology jointly estimates the station-dependent distance corrections in the relationship $m_b = a + b\log(amp) + c\log(dist)$, where amp is the frequency-adjusted P-phase amplitude and $dist$ is the great circle distance between the station and the regional event.

In both of these problems, statistically incorporating the uncertainty in m_b, the worldwide teleseismic magnitudes of observed events, improves the estimation and the associated uncertainty analyses.

Seismic network modeling commonly measures location accuracy in terms of the area of the 90% confidence ellipse (or error ellipse) associated with the estimated location of an event. The CTBT Integrated Verification System Evaluation Model (IVSEM) derives a Monte Carlo estimate of the network location accuracy at a hypothesized location or on a global grid of locations. An improved area estimation methodology, based on rigorous statistical theory, is presented in Section 4 and compared with the current methodology.

2 Uncertainties of Detection Thresholds

The *indirect method* of estimating station detection capability is based on the statistical assumption that observed signal-to-noise ratios divided by the detection thresholds are log-Gaussian distributed (Ringdal 1975). A detection is assumed to occur when

$$SNR > th,$$

where th is the automatic detection threshold. If SNR/th is log-Gaussian distributed, then the probability of a detection is

$$P(SNR > th) = P((\log(SNR/th) - \mu)/\sigma) > -\mu/\sigma) = \Phi(\mu/\sigma), \quad (1)$$

where $\Phi(x)$ is the Gaussian cumulative distribution function and μ and σ^2 are, respectively, the mean and the variance of $\log(SNR/th)$. Usually, μ is assumed to be a function of the magnitude and location of the event, while σ^2 is assumed to be independent of the source of the event (Kelly and Lacoss 1969; Ringdal 1975). Detection thresholds are determined from Equation (1) through inversion. The $p \times 100\%$ detection threshold is the magnitude which satisfies the equation

$$\mu = \sigma\Phi^{-1}(p), \tag{2}$$

where $\Phi^{-1}(x)$ is the inverse Gaussian cumulative distribution function. For example, the 90% detection threshold is the magnitude which satisfies $\mu = 1.28\sigma$.

Sereno and Bratt (1989) modeled μ as a function of range (in km from station to event) and used least-squares to estimate the parameters of this function and to estimate σ^2. A statistical framework for their approach is as follows; adapting their terminology and notation. Let us assume that, for an event with true magnitude m at range r, $\log(SNR/th)$ has a Gaussian distribution with mean

$$\mu = \alpha m + a + b\log(r)$$

and variance σ^2, where a and b are unknown parameters and the parameter α is chosen such that $\log(SNR/th) - \alpha m$ does not depend on the source. Sereno and Bratt (1989) used a local magnitude measurement M_L as an estimate of the true magnitude of the event. Given a set of n events, they used least-squares to regress the values of $\log(SNR/th) - \alpha M_L$ on $\log(r)$, which provided estimates of a and b. However, the mean squared error (MSE) from this regression (the "regression variance") overestimates σ^2 because of the measurement error in M_L. For example, if we assume M_L is independent of $\log(SNR/th)$ and has a Gaussian distribution with mean m and variance σ_L^2, then the mean and variance of $\log(SNR/th) - \alpha M_L$ are

$$a + b\log(r) \qquad \text{and} \qquad \sigma^2 + \alpha^2\sigma_L^2,$$

respectively. The regression MSE is an estimate of the variance $\sigma^2 + \alpha^2\sigma_L^2$ in this case. If M_L is dependent on $\log(SNR/th)$ (which is more likely than not), then the variance of $\log(SNR/th) - \alpha M_L$ is even more complicated. Instead of M_L, Taylor and Hartse (1996) in a study of detection thresholds at the China Digital Seismograph Network station WMQ used a worldwide m_b that can be treated as independent of $\log(SNR/th)$. Thus, the mean and variance of $\log(SNR/th) - \alpha m_b$ is

$$a + b\log(r) \qquad \text{and} \qquad \sigma^2 + \alpha^2\sigma_b^2,$$

respectively, where σ_b^2 is the variance of the worldwide m_b. With this statistical framework in place, estimation of detection thresholds and their uncertainties is possible.

2.1 Estimation of Detection Thresholds

From Equation (2) and the statistical framework just derived, the $p \times 100\%$ detection threshold for events at range r, $m(p, r, \alpha)$, is

$$m(p, r, \alpha) = \left(\sigma\Phi^{-1}(p) - a - b\log(r)\right)/\alpha.$$

The detection thresholds are expressed in this way, $m(p, r, \alpha)$, to make explicit their dependence on the level p, the range r, and the value of α. $m(p, r, \alpha)$ is estimated by inserting estimates of σ, a, and b:

$$\hat{m}(p, r, \alpha) = \left(\hat{\sigma}\Phi^{-1}(p) - \hat{a} - \hat{b}\log(r)\right)/\alpha. \tag{3}$$

Estimates of a and b are provided by the regression of $\log(SNR/th) - \alpha m_b$ on $\log(r)$. An estimate of σ is

$$\hat{\sigma} = \sqrt{s^2 - \alpha^2\hat{\sigma}_b^2},$$

where s^2 is the regression MSE and $\hat{\sigma}_b^2$ is an estimate of the variance of a worldwide m_b.

2.2 Uncertainty Analysis for Estimates of Detection Thresholds

The uncertainty associated with estimates of detection thresholds $\hat{m}(p, r, \alpha)$ in Equation (3) can be estimated in a straightforward manner from the variances of $\hat{\sigma}$, $\hat{a}$, and $\hat{b}$.

If we model the uncertainty in the estimate of the variance of the worldwide m_b by assuming that $K\hat{\sigma}_b^2/\sigma_b^2$ has a Chi-squared distribution with K degrees of freedom, then the variance of $\hat{\sigma}$ is approximately

$$\hat{V} = \frac{1}{2\hat{\sigma}^2}\left(\frac{s^4}{(n-2)} + \frac{\alpha^4\hat{\sigma}_b^4}{K}\right).$$

(The larger K is, the more certain one is of σ_b^2.) The variance of the estimate of the $p \times 100\%$ detection threshold can be estimated by

$$\frac{1}{\alpha^2}\left(\hat{V}\left(\Phi^{-1}(p)\right)^2 + s_a^2 + 2s_{ab}\log(r) + s_b^2(\log(r))^2\right), \tag{4}$$

where

$$\begin{pmatrix} s_a^2 & s_{ab} \\ s_{ab} & s_b^2 \end{pmatrix} = s^2 \begin{pmatrix} n & \sum_{j=1}^n \log(r_j) \\ \sum_{j=1}^n \log(r_j) & \sum_{j=1}^n (\log(r_j))^2 \end{pmatrix}^{-1}$$

is the covariance matrix of the regression estimates of a and b.

This formulation does not take into account the uncertainty in α. If we treat the uncertainty in α as independent of the data (the values of $\log(SNR/th) - \alpha m_b$) and assume its variance equal to σ_α^2, then

$$(\hat{m}(p, r, \alpha)\sigma_\alpha/\alpha)^2$$

should be added to the estimated variance in Equation (4) to take into account the uncertainty in α.

Approximate confidence limits can be placed on the detection threshold estimates $\hat{m}(p, r, \alpha)$ by assuming the Gaussian distribution and using their estimated variances from Equation (4). Approximate 95% confidence limits are $\hat{m}(p, r, \alpha)$ plus/minus 2 times the standard deviation (the square root of Equation (4)).

2.3 Simulation Study

We simulated data to compare our adjusted estimator in Equation (3) with the unadjusted procedures used by Sereno and Bratt (1989) and Taylor and Hartse (1996). The magnitudes m were randomly generated using a Gutenberg-Richter (1941) magnitude-frequency relation with a "b value" of 0.87. The ranges were randomly generated from a uniform distribution between 200 km and 1200 km. The other model parameters used in our simulation are given in Table 1.

Table 1: Regional Detection Thresholds Simulation Parameters

Parameter	Value
a	-3.0
b	5.0
α	1.0
th	2.0
σ^2	0.0225
σ_b^2	0.04

After the true magnitude m and range r were generated for each simulated event, the SNR and m_b were generated. If $SNR > th$, the event was considered detected and used in the subsequent regression of $\log(SNR/th) - \alpha m_b$ on $\log(r)$. These simulated data are plotted in Figure 1, where the size of each circle is proportional to the m_b of its associated event.

The true, adjusted, and unadjusted 90% detection threshold curves are plotted in Figure 2. The adjusted estimator in Equation (3) produces curves which are closer to the true curve. By overestimating the 90% detection thresholds, the unadjusted estimator underestimates the station's detection capability. The unadjusted estimator overestimates the thresholds by over 0.1 magnitude units for the model considered here.

When σ^2 is large compared with σ_b^2, a potential bias problem exists for estimating a and b. Because events with small SNR are not detected, and hence are not available for the regression, the estimate of b can be biased high. With real data, $\hat{\sigma}^2$ can be compared with $\hat{\sigma}_b^2$. If $\hat{\sigma}^2$ is larger than $\hat{\sigma}_b^2$, the unadjusted estimator may give better results because the bias in $\hat{b}$ cancels out the bias when using the regression root mean squared error s instead of $\hat{\sigma}$ in Equation (4).

3 Regional Magnitude Scaling

CTBT monitoring of events in a new region of interest will require calibration of new (or not previously used) stations. Detection and identification of events in the region may be based entirely on data from these uncalibrated stations. Regional magnitude scaling involves calibrating the amplitudes from the regional stations to the worldwide m_b magnitudes using a model very similar to that described in the previous section of this report. That is, we want to determine the station-dependent values of a, b, and c such that, for each station,

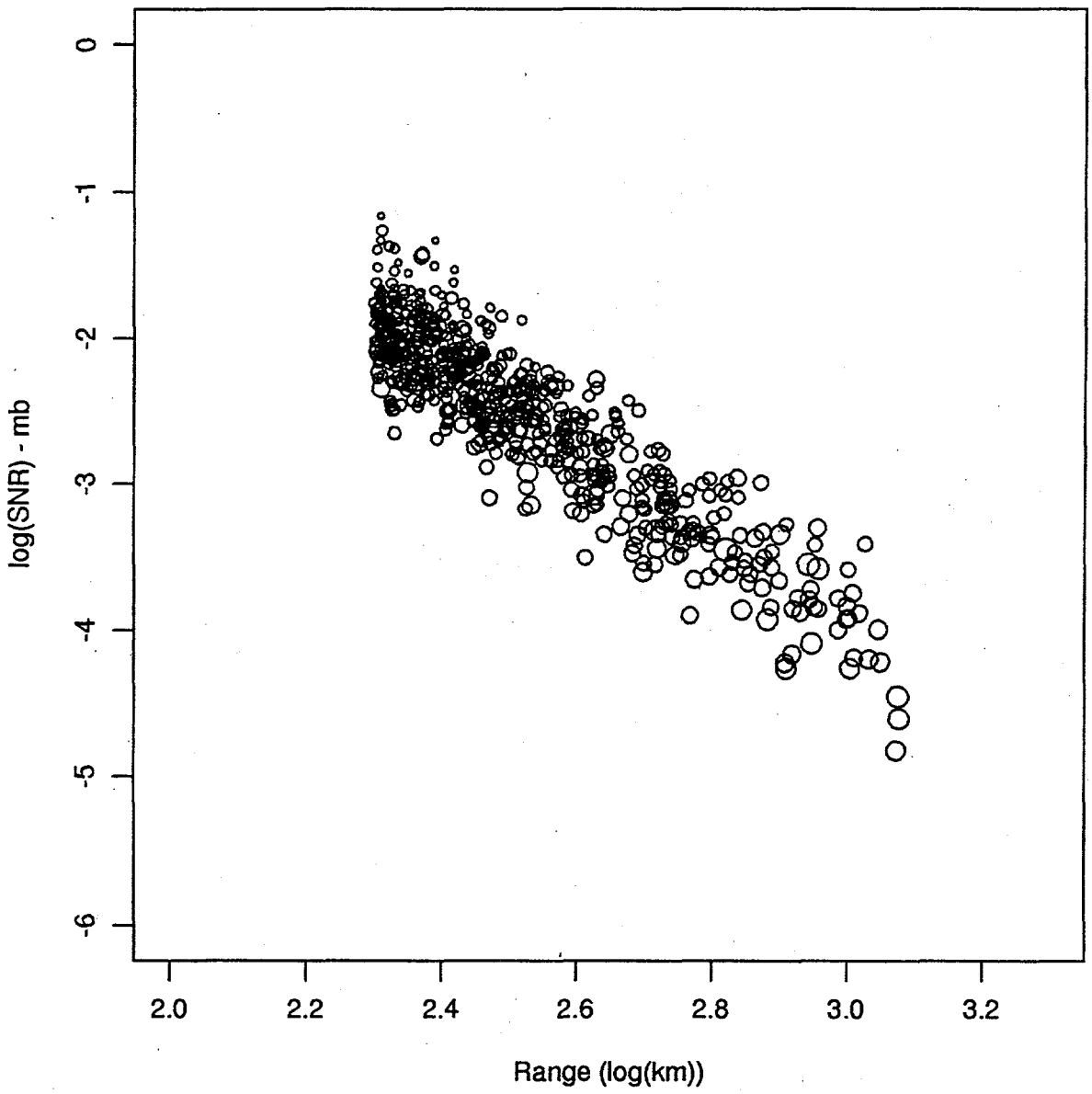

Figure 1: Simulated Signal-to-noise versus Range Data. Simulated $\log(SNR/th) - \alpha m_b$ versus $\log(r)$ data are plotted. The size of each circle is proportional to m_b.

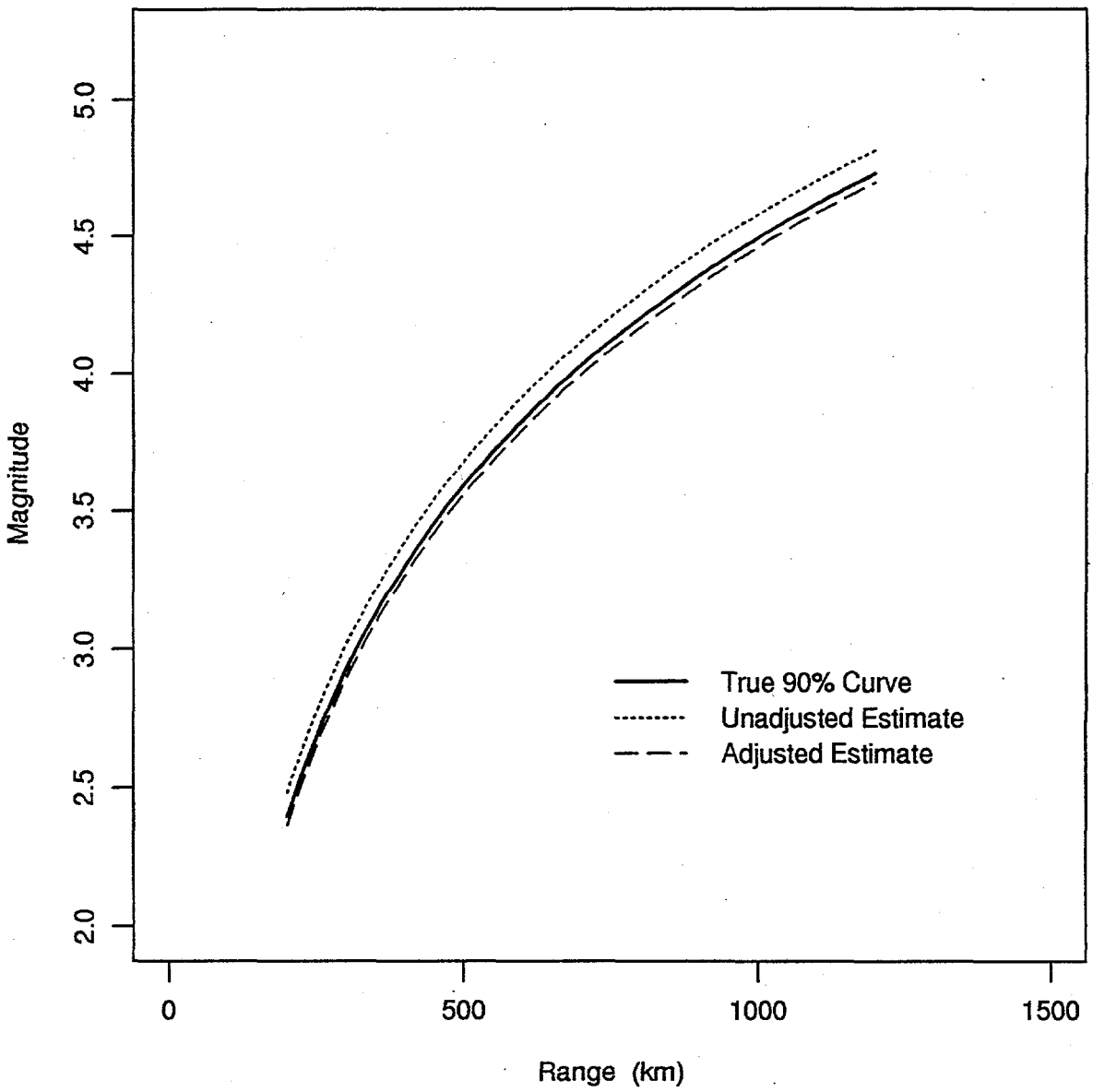

Figure 2: Detection Thresholds for Simulated Data. The estimated 90% detection thresholds from simulated data were calculated using both the adjusted and unadjusted regression variances. The bold, solid line is the true 90% detection threshold curve.

$$a + b\log(A) + c\log(D) \tag{5}$$

is scaled to the worldwide m_b magnitude scale, where A is the frequency-adjusted station amplitude and D is the distance from the station to the event. A regional path effect term could be included if an analysis of the residuals after scaling shows spatial structure. The parameter b is often assumed to be constant across stations. However, modeling b as station-dependent permits us to examine this assumption.

This section considers the problem of calibrating a set of regional stations to the worldwide m_b magnitude for a large set of events. The statistical approach taken is known as measurement error modeling (Fuller 1987) and is similar to the approach used by Anderson and Nicholson (1993) to determine relative regional magnitude-yield relationships. The approach takes into account the uncertainty (or measurement error) in the m_b values and in the distances (due to location errors).

3.1 Statistical Model

A number of regional stations can be scaled jointly. It makes sense to do so because the stations observe many of the same events, especially with well-balanced data. The statistical model includes the station-dependent parameter b as presented in Equation (5). Scaling the stations jointly allows us to test whether parameter b is the same for each station and therefore only one value is needed.

Table 2: Regional Magnitude Scaling Data Table: Example

	Station 1 Amplitude	Station 2 Amplitude	Station 3 Amplitude	m_b	Event Location	
Event 1	A_{11}	A_{12}	A_{13}	m_{b1}	LON_1	LAT_1
Event 2	A_{21}	A_{22}	A_{23}	m_{b2}	LON_2	LAT_2
Event 2	A_{31}	A_{32}	A_{33}	m_{b3}	LON_3	LAT_3
Event 4	A_{41}	A_{42}	A_{43}	m_{b4}	LON_4	LAT_4
Event 5	A_{51}	A_{52}	A_{53}	m_{b5}	LON_5	LAT_5

We assume that data from n events are available to scale K regional stations whose locations are known. The necessary data fill a table such as Table 2 (shown with five events and three stations). The station amplitude data are modeled as

$$\log(A_{ij}) = (m_i - a_j - c_j\log(d_{ij}))/b_j + \epsilon_{ij}, \quad i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, K, \tag{6}$$

where d_{ij} are the true great circle distances, m_i are the true event magnitudes, and ϵ_{ij} are independent Gaussian random variables with zero means and with variances σ_j^2 that may depend on the station.

The observed event locations are modeled as the true locations (lon_i, lat_i) plus independent measurement error,

$$LON_i = lon_i + v_{1i}, \qquad LAT_i = lat_i + v_{2i}, \tag{7}$$

where (v_{1i}, v_{2i}) are independent bivariate Gaussian random variables with zero mean and common covariance Σ. The great circle distance (ignoring depth and elevation) between two points on the earth is a well-known function of the longitudes and latitudes of the points. The assumption of unbiased errors in the event locations is questionable. In many regions, systematic errors in location are observed; however, those errors generally are not well understood. Our initial studies show that the errors in the log-distances are nearly negligible compared to the errors in the magnitudes and that reasonable results are achieved by ignoring errors in log-distances.

The worldwide m_b measurements for these n regional events are assumed to be independent of the individual station amplitudes and distances and are direct measurements of the true event magnitudes m_i,

$$m_{bi} = m_i + u_i, \tag{8}$$

where u_i are independent Gaussian random variables with zero mean and common variance σ_b^2.

The unknown parameters a_j, b_j, and c_j are the scaling constants for station j. The true event magnitudes m_i and the event locations are considered *nuisance parameters* in this problem; they must be estimated at the same time as the scaling constants, but are of secondary interest (the nuisance parameters can be viewed as being re-estimated during this scaling process, since the magnitudes and locations were already estimated). For estimation purposes, the variances σ_j^2, σ_b^2, and Σ must be known at least to a constant multiple; that is, we need to know their relative sizes.

3.2 Joint Estimation of the Scaling Constants

Assuming the independence and Gaussian distribution of all of the measurement errors in Equations (6-8), the maximum likelihood estimates of the unknown scaling constants a_j, b_j, and c_j, the true event locations (lon_i, lat_i), and the true event magnitudes m_i can be found by least-squares estimation. That is, minimize the weighted sum of squares objective function

$$\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{K}(\log(A_{ij}) - (m_i - a_j - c_j \log(d_{ij}))/b_j)^2/\sigma_j^2 + \frac{1}{2}\sum_{i=1}^{n}(m_{bi} - m_i)^2/\sigma_b^2 + \tag{9}$$

$$\frac{1}{2}\sum_{i=1}^{n}\begin{pmatrix} LON_i - lon_i \\ LAT_i - lat_i \end{pmatrix}' \Sigma^{-1} \begin{pmatrix} LON_i - lon_i \\ LAT_i - lat_i \end{pmatrix}$$

with respect to the model parameters. The three major sums contributing to the objective function correspond to the measurement errors in the log-amplitudes, in the worldwide m_b magnitudes, and in the event locations, respectively. All three contributions are minimized together. The variances σ_j^2, σ_b^2, and Σ serve as the weights.

The weighted sum of squares in Equation (9) is a nonlinear function of the parameters and requires an iterative optimization scheme. However, the possibly large number

of parameters may cause some difficulty for optimization routines. Some iterative shortcut may be needed, like the one used by Anderson and Nicholson (1993). The covariance of the estimates can be calculated using the classical maximum likelihood approach (via the Hessian of the likelihood function).

3.3 Model Extensions

The procedure given above assumes complete amplitude/magnitude data; i.e., that m_b magnitudes and amplitudes are available for all events from all stations. This is not necessary, however, as Anderson and Nicholson (1993) demonstrated with a similar analysis of amplitudes of Soviet underground nuclear tests. Data missing because of inoperable stations is not a problem. If amplitudes and m_b magnitudes are missing because of low signal-to-noise ratios, then the parameter estimates could be biased. In this case, a maximum likelihood procedure, such as that introduced by Ringdal (1986) for network magnitudes, provides better estimates. Further, since the data and models of Section 2 and this section are so similar, the two analyses, estimation of regional station detection thresholds and regional magnitude scaling, could be performed jointly. Future research will merge the maximum likelihood approach to magnitude scaling with the joint estimation of regional station detection thresholds.

4 Location Accuracy

This section describes a statistical approach to improve the location accuracy modeling in the CTBT Integrated Verification System Evaluation Model (IVSEM). Network location accuracy is measured in terms of the area of the 90% confidence ellipse or error ellipse associated with the estimated location of an event. IVSEM derives a Monte Carlo estimate of the network location accuracy at a hypothesized location or on a global grid of locations. An improved area estimation methodology is presented and compared with the current methodology.

4.1 Statistical Theory

The multivariate Gaussian distribution is parameterized by a mean vector μ and a covariance matrix Σ. The usual confidence regions are ellipsoids centered at μ with shapes controlled by Σ. That is, with a confidence level of $p \times 100\%$, a multivariate Gaussian random variable X is expected to fall within the $p \times 100\%$ confidence ellipsoid. In two dimensions (the bivariate Gaussian case), the confidence regions are ellipses.

The $p \times 100\%$ confidence ellipses for a bivariate Gaussian random variable X are derived from the following statistical theory. The quadratic form $(X-\mu)'\Sigma^{-1}(X-\mu)$ has a Chi-squared distribution with two degrees of freedom (Anderson 1958). Hence, the probability that X falls within the ellipse defined by $E_p = \{x : (x-\mu)'\Sigma^{-1}(x-\mu) = c_p\}$ is

$$P((X-\mu)'\Sigma^{-1}(X-\mu) \leq c_p) = p$$

where c_p is the quantile of the Chi-squared distribution with two degrees of freedom corresponding to the desired confidence level. Some values of c_p are given in Table 3. The ellipse E_p is referred to as the $p \times 100\%$ confidence ellipse for X. From analytic geometry, the area of E_p is

$$\pi c_p \sqrt{\sigma_{11}\sigma_{22} - \sigma_{12}^2},$$

where σ_{11}, σ_{12}, and σ_{22} are the elements of the covariance matrix Σ. Note that $\sigma_{11}\sigma_{22} - \sigma_{12}^2$ is the determinant of Σ.

Table 3: Quantiles of the Chi-squared Distribution with Two Degrees of Freedom

Confidence Level p	c_p
0.75	2.772589
0.80	3.218876
0.85	3.794240
0.90	4.605170
0.95	5.991465
0.96	6.437752
0.97	7.013116
0.98	7.824046
0.99	9.210340

4.2 Estimating the Area of Confidence Ellipses

When Σ is unknown, the area of the $p \times 100\%$ confidence ellipse, $\pi c_p \sqrt{\sigma_{11}\sigma_{22} - \sigma_{12}^2}$, can be estimated by *plugging in* estimates of the variances and covariance. Our approach is to estimate Σ using the sample covariance matrix S and then use the determinant of S in the area formula. Given a random sample of size n from the bivariate Gaussian, the sample covariance matrix is

$$S = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})(X_i - \bar{X})',$$

where $\bar{X} = (1/n) \sum_{i=1}^{n} X_i$. The estimated area of the $p \times 100\%$ confidence ellipse is

$$\hat{A}_p = \pi c_p \sqrt{s_{11}s_{22} - s_{12}^2},$$

where s_{11}, s_{12}, and s_{22} are the elements of S.

An alternate formulation for S which takes into account knowledge of μ is

$$S = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \mu)(X_i - \mu)'.$$

When μ is known, it does not have to be estimated by $\bar{X}$.

4.3 Application to IVSEM

Location accuracy in IVSEM is defined to be the area of the 90% confidence ellipse for the location of an event. The location accuracy calculations in IVSEM are based on a Monte Carlo approach, in which each event is simulated N (currently 100) times. At each grid point, the location data (onset times and/or azimuths) are simulated for each detecting station and used to estimate the location (latitude and longitude) of the epicenter of the simulated event. Assuming that the bivariate Gaussian distribution is the appropriate one, the N simulated locations can be translated to a two-dimensional coordinate system (centered at the grid point and in kilometers) and used to calculate the sample covariance matrix S and the area of the 90% confidence ellipse $\hat{A}_{0.9}$. IVSEM then contours the $\hat{A}_{0.9}$ values over the global grid.

Currently, IVSEM estimates the area of the 90% confidence ellipse using $\pi D_{0.90}^2$, where $D_{0.90}$ is the value such that 90% of the distances $D_i = \| X_i - X_0 \|$ are less than $D_{0.90}$ and X_0 is the grid point of interest. This approach yields reasonable estimates of the area when the covariance matrix Σ is proportional to the identity matrix I, the identity matrix, or in other words, when the confidence ellipses are circles. When this is not the case, $\pi D_{0.90}^2$ usually overestimates the true area. The approach described here, based on the eigenvalues of the inverse of the sample covariance matrix, provides a better area estimate (and hence a better estimate of location accuracy estimate) for all bivariate Gaussian cases, and furthermore can account for any Monte Carlo location bias.

The Monte Carlo approach used by IVSEM allows the assumption of Gaussian errors in location to be examined. The Gaussian nature of the 100 simulated locations can be tested. If deviations from the bivariate Gaussian distribution are due to a small percentage of outliers, then a robust method of computing the sample covariance matrix S can be used. Outliers cause S to overestimate Σ and hence overestimate $A_{0.9}$. We plan to examine IVSEM Monte Carlo locations and apply several robust covariance matrix estimation techniques (Rubin, 1983).

5 Conclusions and Recommendations

The statistical methods presented in this report can provide better estimates of individual station regional detection thresholds and magnitude distance/station corrections (a regional magnitude scale) by taking into account the uncertainties in event magnitudes. The author intends to apply these statistical methods to data from the two regions of interest in the DOE CTBT research program – western China and the Middle East/North Africa.

Future work will extend the regional detection threshold estimation to include estimation of α by merging the analysis with the multi-station magnitude scaling. We will also incorporate small magnitude biases in the worldwide m_b measurements by using the maximum likelihood approach of Ringdal (1986).

We have also presented a Monte Carlo estimate of location accuracy based on rigorous statistical theory. This approach should offer an improvement over the current IVSEM approach.

6 References

Anderson, K. K., and W. L. Nicholson (1993). A relative magnitude model for measurements of Soviet underground nuclear explosions from regional stations, *Bull. Seism. Soc. Am.* **83**, 1563-1573.

Anderson, K. K., and D. N. Anderson (1995). Statistical Issues in CTBT Verification: Seismic Monitoring Performance Measures, Pacific Northwest Laboratory, Richland, Washington, Technical Report PNL-10805.

Anderson, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons, New York.

DOE (1994), *Comprehensive Test Ban Treaty Research and Development FY95-96 Program Plan*, United States Department of Energy, Report DOE/NN-0003.

Fuller, W. A. (1987). *Measurement Error Models*, John Wiley & Sons, New York.

Gutenberg, B., and C. F. Richter (1941). Seismicity of the Earth, *Geol. Soc. Am., Spec. Pap.* **34**, 1-133.

Kelly, E. J., and R. T. Lacoss (1969). Estimation of seismicity and network detection capability, Technical Note 1969-41, Lincoln Laboratory, MIT, Lexington, Massachusetts.

Ringdal, F. (1975). On the estimation of seismic detection thresholds, *Bull. Seism. Soc. Am.* **65**, 1631-1642.

Ringdal, F. (1986). Study of magnitudes, seismicity, and earthquake detectability using a global network, *Bull. Seism. Soc. Am.* **76**, 1641-1659.

Rubin, D. B. (1983). Iteratively reweighted least squares, *Encyclopedia of the Statistical Sciences, Volume 4*, John Wiley & Sons, New York.

Sereno, T. J., and S. R. Bratt (1989). Seismic detection capability at NORESS and implications for the detection threshold of a hypothetical network in the Soviet Union, *J. Geophys. Res.* **94**, 10,397-10,414.

Taylor, S.R., and H.E. Hartse (1996), Regional phase detection thresholds at WMQ, Los Alamos National Laboratory, Los Alamos, New Mexico, LAUR-96-395.

Distribution

No. Of Copies **Copies**

2 DCI/ACIS
Arms Control Intelligence Staff
Washington, DC 20505

J. Filson
L.S. Turnbull

2 Defense Special Weapons Agency
Nuclear Treaty Programs
1901 N. Moore St., Suite 609
Arlington, VA 22209

R. Alewine
S.R. Bratt

8 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

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

R. Blandford

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

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

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

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

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

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

Copies

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

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
D.R. Breding, MS 0655

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
K.K. Anderson, K5-12 (10)
D.M. Boyd, K5-02
D.L. Eggett, K5-12
T.R. Fox, K6-48
D.N. Hagedorn, K5-12 (5)
R.C. Hanlen, K6-40
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
F.M. Ryan, K5-12
Information Release, K1-06(7)