AL-84-7

AFOSR-TR- 85-0241

AD-A151 162

Final Technical Report

15 November 1982 - 30 December 1984

TESTING THE HYPOTHESIS OF TTBT COMPLIANCE, AND MAGNITUDE-YIELD REGRESSION FOR EXPLOSIONS IN GRANITE

by

R.H. Shumway and D.W. Rivers

Teledyne Geotech
Alexandria Laboratories
P.O. Box 334
Alexandria, VA 22313

Sponsored by
Advanced Research Projects Agency (DOD)
ARPA Order No. 4493
Monitored by AFOSR/NP Under Contract #F49620-83-C-0040

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AL-84-7	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 85-0241
6a. NAME OF PERFORMING ORGANIZATION Teledyne Geotech 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research
6c. ADDRESS (City, State and ZIP Code) 314 Montgomery Street Alexandria, Virginia 22314	7b. ADDRESS (City, State and ZIP Code) Bolling Air Force Base Building 420 Washington, D. C. 20332
8a. NAME OF FUNDING/SPONSORING ORGANIZATION DARPA 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620-83C-0040
8c. ADDRESS (City, State and ZIP Code) 1400 Wilson Boulevard Arlington, Virginia 22209	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO.: 61101E PROJECT NO.: 4493 TASK NO.: 00 WORK UNIT NO.: Task B
11. TITLE (Include Security Classification) (See Block 16)	

12. PERSONAL AUTHOR(S)
R. H. Shumway and D. W. Rivers

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Final	FROM 11/15/82 TO 12/30/84		

16. SUPPLEMENTARY NOTATION
Testing the Hypothesis of TTBT Compliance, and Magnitude-Yield Regression for Explosions in Granite

17. COSATI CODES

FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
nuclear explosion; Threshold Test Ban Treaty (TTBT); seismic magnitude; explosive yield; hypothesis testing

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Verification of the Threshold Test Ban Treaty (TTBT) cannot be performed merely by estimating explosive yields on the basis of observed seismic magnitudes and concluding that a violation has occurred if one or more yield estimates exceed the TTBT limit of 150 KT. It is necessary to take into account the uncertainties in the seismic magnitudes, in the magnitude - yield relation, and especially in the magnitude bias between the test site at which the magnitude - yield calibration explosions were detonated and the test site of the explosions being monitored. For monitoring one explosion at a time, these uncertainties can be taken into account by placing confidence limits around the yield estimates. For verifying TTBT compliance of an ensemble of explosions considered as a whole, however, this technique cannot be used, since the confidence limits placed around the yield estimates of different explosions are correlated due to the use in every case of the same values of the parameters relating magnitude to yield. In order to examine TTBT compliance for groups of explosions, a test

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr Radoski	22b. TELEPHONE NUMBER (Include Area Code) (202) 767-4906 22c. OFFICE SYMBOL NP

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE. UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

can be performed of the hypothesis that the yields have some fixed distribution in which all the values are less than 150 KT. This compliance hypothesis would be rejected if some minimum number of yields which are calculated using the observed magnitudes exceed a certain threshold which is a function of both the minimum number of exceedences and the maximum allowed probability of issuing a false alarm. For a small false alarm probability, the threshold at which the compliance hypothesis is rejected will far exceed the TTBT limit of 150 KT. This particular hypothesis test is not optimum, but it is preferable to more powerful ones in that it is not susceptible to manipulation by the party conducting the explosions.

The magnitude - yield relation may be calibrated from a set of explosions with known yields by means of a least-squares regression which treats the magnitudes as uncertain quantities and the yields as fixed or by a maximum-likelihood regression which treats both magnitudes and yields as uncertain. For the purposes of TTBT monitoring, it is important to determine not only the best estimates of the parameters relating magnitude to yield but also the variances of these estimates.

A-1

ABSTRACT

Verification of the Threshold Test Ban Treaty [TTBT] cannot be performed merely by estimating explosive yields on the basis of observed seismic magnitudes and concluding that a violation has occurred if one or more yield estimates exceed the TTBT limit of 150 KT. It is necessary to take into account the uncertainties in the seismic magnitudes, in the magnitude - yield relation, and especially in the magnitude bias between the test site at which the magnitude - yield calibration explosions were detonated and the test site of the explosions being monitored. For monitoring one explosion at a time, these uncertainties can be taken into account by placing confidence limits around the yield estimates. For verifying TTBT compliance of an ensemble of explosions considered as a whole, however, this technique cannot be used, since the confidence limits placed around the yield estimates of different explosions are correlated due to the use in every case of the same values of the parameters relating magnitude to yield. In order to examine TTBT compliance for groups of explosions, a test can be performed of the hypothesis that the yields have some fixed distribution in which all the values are less than 150 KT. This compliance hypothesis would be rejected if some minimum number of yields which are calculated using the observed magnitudes exceed a certain threshold which is a function of both the minimum number of exceedences and the maximum allowed probability of issuing a false alarm. For a small false alarm probability, the threshold at which the compliance hypothesis is rejected will far exceed the TTBT limit of 150 KT. This particular hypothesis test is not optimum, but it is preferable to more powerful ones in that it is not susceptible to manipulation by the party conducting the explosions.

The magnitude - yield relation may be calibrated from a set of explosions with known yields by means of a least-squares regression which treats the mag-

nitudes as uncertain quantities and the yields as fixed or by a maximum-likelihood regression which treats both magnitudes and yields as uncertain. For the purposes of TTBT monitoring, it is important to determine not only the best estimates of the parameters relating magnitude to yield but also the variances of these estimates.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

Throughout the 25-year history of Project VELA Uniform, considerable effort has been expended on research related to the accurate measurement of the seismic magnitudes of underground nuclear explosions and the estimation of explosive yields using those magnitudes. This research (much of which has been summarized by Bache *et al.*, 1981, and Bache, 1982) has addressed such issues as instrumentation, signal measurement (both analog and digital), amplitude-distance relations, station magnitude corrections, test site magnitude biases, properties of emplacement media, effects of depth of burial, and source coupling. Although in this report we shall examine both the determination of magnitudes and the estimation of yields, we shall concentrate primarily on a different issue which is the logical successor to these topics : given the uncertainties in seismic magnitudes and in the yield estimation procedure, how can the resulting yield estimates be used in an effective manner for monitoring a Threshold Test Ban Treaty [TTBT] ? We shall examine this question by placing confidence limits around the estimates of individual yields and by performing classical hypothesis tests for TTBT compliance of ensembles of events. We shall next examine the technique for estimating those yields which were used in the TTBT compliance testing procedures. We shall place special emphasis on the uncertainty in those estimates, since it is this uncertainty which limits our capability of verifying the TTBT.

TESTING TTBT COMPLIANCE

We shall first assume that the seismic magnitudes m_b have been measured for a suite of explosions and that the yields of these explosions have been estimated using these values of m_b and some specified formula which relates magnitude to yield. We shall later examine this yield estimation procedure, but for now we shall merely take it to be given, and we shall investigate how the estimated yields and the uncertainties in both the magnitudes and in the yield estimates may be used to quantify the confidence with which it can be said that the suite of events was, or was not, in compliance with a TTBT.

1. Statistical Model of the Magnitude - Yield Relation

TTBT compliance for a set of n explosions of unknown yields will be tested using the n measured seismic magnitudes m_1, m_2, ..., m_n. Although we could use surface-wave magnitudes for yield estimation, we shall herein assume that the magnitudes are all m_b rather than M_s. We assume that these observations are linearly related to the logarithms of the yields :

$$m_j = \alpha + \beta W_j + e_j \quad (1)$$

where $W_j = \log Y_j$ and Y_j is the yield of the jth event. Estimates of the intercept and slope of this relation, which we shall denote as α_0 and β_0, can be obtained from a linear fit of observed magnitudes to the known yields of calibration explosions. A technique for performing this calibration magnitude-yield regression will be presented later. The intercept α in equation (1) includes an additive constant ε which represents the magnitude bias between the test site at which the magnitude-yield calibration events were detonated and the test site of the n events with the unknown yields. This magnitude bias ε is due to differences in the amount of anelastic attenuation of seismic waves which takes place in the crust and mantle beneath the two test sites. For those events of unknown yield,

α and β are subject to considerable uncertainty which we model in terms of their expectations α_0 and β_0 from the calibration set. The uncertainty in α and β relates to coupling and the unknown characteristics of the test site as well as the variance among the calibration events. We assume that this can be incorporated into the variance-covariance structure of α and β, say by the covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_\alpha^2 & \sigma_{\alpha\beta} \\ \sigma_{\alpha\beta} & \sigma_\beta^2 \end{pmatrix} \tag{2}$$

The error term e_j in equation (1) is assumed to have variance σ_j^2 which comes from two sources. The first is the error in the measurement of m_b (after any available station corrections have been applied) mainly due to stochastic propagation effects which are not adequately compensated for by the station corrections, and the second source of error is the scatter in m_b for a fixed yield which is due to differences in the coupling efficiency at the n source locations within the test site. We shall later consider these two sources of magnitude error separately; for now, we shall simply add the variances to obtain the single variance σ_j^2 for the magnitude of the jth event.

Although we shall later discuss the determination from calibration data of the parameters α_0 and β_0 as well as the uncertainties σ_α, σ_β, and $\sigma_{\alpha\beta}$, we shall for now assume that these are fixed quantities which are known *a priori*. For purposes of illustrating the compliance testing procedures, we shall use as values of these parameters certain results which will be derived in a subsequent section of this report. We emphasize that these values are for illustration purposes only; the determination of the parameters which should be used in practice for TTBT monitoring must be derived by a detailed analysis of a large data set of carefully selected calibration explosions with known yields. Although we shall investigate the methodology which should be used in ıch a study, in this report we shall not attempt to accumulate the data set which would be required to carry out the calibration in practice, and the values which we shall use here are therefore unreliable. They will, however, serve to show how TTBT compliance can be tested, given some better set of values for these same parameters. The values which we shall use herein are listed in Table I. It is important to note that for testing TTBT compliance of the n events we must use values of α_0 and σ_α given by

$$\begin{aligned} \alpha_0 &= \alpha_{0,cal} + \varepsilon_0 \\ \sigma_\alpha^2 &= \sigma_{\alpha,cal}^2 + \varphi^2 \end{aligned} \tag{3}$$

where ε_0 represents the estimate of the magnitude bias of the test site of the n events relative to the site of the magnitude-yield calibration events, and where φ is the standard deviation of this bias estimate. The value of the magnitude bias between US and USSR test sites and the uncertainty in that value have both long been controversial topics, and rather than address them here we shall simply choose several different trial values for illustration purposes. We shall use 3 different possible values of ε_0 and 2 of φ, as is shown in Table I. We shall thus test compliance under 6 different assumptions about the magnitude bias between test sites, namely that ε_0 = 0.20 ± 0.05, 0.20 ± 0.10, 0.25 ± 0.05, 0.25 ± 0.10, 0.30 ± 0.05, and 0.30 ± 0.10. This means that, in accordance with equation (3), we shall test TTBT compliance for the 6 possibilities α_0 = 4.082 ± 0.059, 4.082 ± 0.105, 4.132 ± 0.059, 4.132 ± 0.105, 4.182 ± 0.059, and 4.182 ± 0.105. This range of values will serve to show how strongly the question of compliance is affected by the estimate of the magnitude bias and by the uncertainty in that estimate.

It will be recalled that the parameter σ_j in Table I is the standard deviation of the error term e_j in equation (1). In Table I we have adopted a constant value σ_j = 0.07 for j = 1, ..., n, and we shall now attempt to justify this procedure. If

TABLE I

Values of paramcters to be used in this study for testing TTBT compliance

Shagan River – calibration site m_b bias $\varepsilon_0 = \begin{cases} 0.20 \\ 0.25 \\ 0.30 \end{cases}$ (3 trial values)

standard deviation of the preceding estimate $\varphi = \begin{cases} 0.05 \\ 0.10 \end{cases}$ (2 trial values)

intercept of the calibration magnitude – yield relation $\alpha_{0,cal} = 3.882$

hence for Shagan $\alpha_0 = \begin{cases} 4.08 & \text{if } \varepsilon_0 = 0.20 \\ 4.13 & \text{if } \varepsilon_0 = 0.25 \\ 4.18 & \text{if } \varepsilon_0 = 0.30 \end{cases}$

slope of the magnitude – yield relation $\beta_0 = 0.890$

standard deviation of the calibration intercept $\sigma_{\alpha,cal} = 0.031$

hence at Shagan $\sigma_\alpha = \begin{cases} 0.059 & \text{if } \varphi = 0.05 \\ 0.105 & \text{if } \varphi = 0.10 \end{cases}$

standard deviation of the slope $\sigma_\beta = 0.013$

covariance of the intercept and slope $\sigma_{\alpha\beta} = -0.000393$

standard deviation of m_b for each event $\sigma_j = 0.07$

These values are for illustration purposes only and are not intended to be used in practice.

the n magnitudes $m_1, m_2, \ldots, m_n$ are determined using a large network of globally distributed stations, for which the individual station corrections (for the given test site) average out to zero, then the uncertainty in the calculated magnitude m_j is just the standard deviation of the mean of the magnitudes m_{ij} measured at each of the N_j stations i. If the magnitudes are determined using only a few stations, then the uncertainty in m_j is larger than this on account of the uncertainty in a possible network magnitude bias. Since the events which are of interest from the point of view of investigating TTBT compliance are large ones which are detected by many stations, we can assume that any network bias is small and that the standard deviation of m_j is likewise small. Although the uncertainty in the calculated value for m_j is small for these large-magnitude events, a larger uncertainty is required to account for the scatter about the magnitude-yield relation due to variations in coupling efficieny. (Recall that the variance σ_j^2 is the sum of both these uncertainties.) We shall present later a technique for estimating the coupling scatter, but for now we take it to be known *a priori* as $\sigma_c = 0.07$. This value of the magnitude scatter (for events of the same yield) due to coupling is larger than the magnitude scatter (for a single event) due to measurement of m_b across the network, since the events of interest for TTBT compliance are recorded at many stations N_j. The magnitudes which we shall use to illustrate TTBT compliance testing are listed in Table II. These data are taken from Sykes and Cifuentes (1984), who used magnitudes reported to the International Seismological Center [ISC] bulletin and then applied to them station corrections which were derived from the 8 largest events. We shall neglect the slight differences among the listed standard deviations of m_b for the $n = 22$ events and set $\sigma_j^2 = \sigma_c^2 + \sigma_{m_j}^2 \approx 0.07^2 + 0.02^2$ for all events, thus obtaining $\sigma_j \approx 0.07$ for $j = 1, \ldots, n$. Although the preceding calculation is an approximate one, it would not be worthwhile to replace it by a more nearly exact one in which the different values of σ_{m_j} for different events are

TABLE II

Magnitudes of nuclear explosions at eastern Kazakh which will be used to illustrate TTBT compliance tests (data taken from Sykes and Cifuentes, 1984)

Date	m_b
Aug. 29, 1978	5.967 ± 0.012
Sep. 15, 1978	5.963 ± 0.015
Nov. 4, 1978	5.576 ± 0.018
Nov. 29, 1978	5.996 ± 0.017
Jun. 23, 1979	6.215 ± 0.013
Jul. 7, 1979	5.839 ± 0.020
Aug. 4, 1979	6.161 ± 0.013
Aug. 18, 1979	6.170 ± 0.015
Oct. 28, 1979	5.990 ± 0.016
Dec. 2, 1979	5.998 ± 0.013
Dec. 23, 1979	6.170 ± 0.017
Jun. 29, 1980	5.707 ± 0.019
Sep. 14, 1980	6.213 ± 0.030
Oct. 12, 1980	5.918 ± 0.019
Dec. 14, 1980	5.954 ± 0.019
Dec. 27, 1980	5.872 ± 0.023
Apr. 22, 1981	5.954 ± 0.015
Sep. 13, 1981	6.064 ± 0.017
Oct. 18, 1981	6.033 ± 0.019
Dec. 27, 1981	6.242 ± 0.028
Apr. 25, 1982	6.089 ± 0.021
Jul. 4, 1982	6.222 ± 0.026

taken into account, since these small differences are less than the uncertainty in the value of 0.07 which is used for the parameter σ_c. Although we shall use a single value of σ_j for all events, the formalism which we shall derive for testing TTBT compliance does not depend on the use of this approximation.

2. Confidence Intervals for Yield Estimates

The estimated log yield of the j th event, W_j^{est}, can be determined from the j th observed magnitude as

$$W_j^{est} = \frac{m_j - \alpha_0}{\beta_0} \quad (4)$$

The estimated yield is normal with mean W_j and variance

$$\mathrm{var}(W_j^{est}) = (\sigma_\alpha^2 + 2\sigma_{\alpha\beta} W_j + \sigma_\beta^2 W_j^2 + \sigma_j^2) / \beta_0^2 \quad (5)$$

which depends on the true log yield W_j. In order to derive a 100(1-α)% confidence interval, we substitute the right-hand side of equation (5) into the denominator of the inequality

$$\frac{| W_j^{est} - W_j |}{\sqrt{\mathrm{var}(W_j^{est})}} \le Z_{\alpha/2} \quad (6)$$

where $Z_{\alpha/2}$ is the upper $\alpha/2$ percentage point on the normal distribution. (Note that α in this context is a different variable from the α which denotes the intercept in equation (1)). This substitution results in a quadratic expression for W_j, the two roots of which bound the confidence interval. Since we are really interested only in the lower magnitude bound and not on the upper bound on the confidence interval, we shall replace the right-hand side of equation (6) with Z_α so that the confidence interval is one-sided. A possible compliance testing procedure would then be to challenge any yield for which the lower bound on the 95% confidence interval was greater than the maximum yield which is permitted by the TTBT, namely 150 KT.

TABLE III

95% confidence intervals for yield estimates

(1) Yield (KT) estimated using $\varepsilon_0 = 0.20$
(2) Lower bound on (1) if $\varphi = 0.05$
(3) Lower bound on (1) if $\varphi = 0.10$
(4) Yield (KT) estimated using $\varepsilon_0 = 0.25$
(5) Lower bound on (4) if $\varphi = 0.05$
(6) Lower bound on (4) if $\varphi = 0.10$
(7) Yield (KT) estimated using $\varepsilon_0 = 0.30$
(8) Lower bound on (7) if $\varphi = 0.05$
(9) Lower bound on (7) if $\varphi = 0.10$

event	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Aug. 29, 1978	131	91	78	115	80	68	101	70	60
Sep. 15, 1978	130	90	77	114	79	68	100	69	60
Nov. 4, 1978	48	33	28	42	29	25	37	25	22
Nov. 29, 1978	141	98	84	124	86	74	109	76	65
Jun. 23, 1979	249	172	148	219	152	130	192	133	114
Jul. 7, 1979	94	65	56	83	57	49	73	50	43
Aug. 4, 1979	217	150	129	190	132	113	167	116	99
Aug. 18, 1979	222	154	132	195	135	116	171	119	102
Oct. 28, 1979	139	96	83	122	85	73	108	74	64
Dec. 2, 1979	142	98	84	125	86	74	110	76	65
Dec. 23, 1979	222	154	132	195	135	116	171	119	102
Jun. 29, 1980	67	46	40	59	41	35	52	36	31
Sep. 14, 1980	248	172	147	218	151	129	191	132	114
Oct. 12, 1980	116	80	69	102	70	60	89	62	53
Dec. 14, 1980	127	88	75	111	77	66	98	68	58
Dec. 27, 1980	103	71	61	90	62	54	79	55	47
Apr. 22, 1981	127	88	75	111	77	66	98	68	58
Sep. 13, 1981	169	117	100	148	103	88	130	90	77
Oct. 18, 1981	156	108	92	137	95	81	120	83	71
Dec. 27, 1981	267	185	159	235	163	139	206	143	123
Apr. 25, 1982	180	124	107	158	109	94	139	96	82
Jul. 4, 1982	254	176	151	223	154	132	196	136	116

Table III shows the results of applying this test to each of the events in Table II, using the values of α_0, β_0, σ_α, σ_β, $\sigma_{\alpha\beta}$, and σ_j taken from Table I. We see that for 3 of our 6 trial values of $\varepsilon_0 \pm \varphi$, thc lower limit on the confidence interval surrounding the yield estimate exceeds 150 KT for at least one event, and for the other 3 trial values it does not. (The confidence limits on the yields are of course functions not only of the trial values of ε_0 and φ but also of the values which were selected *a priori* for the other parameters in Table I.) The procedure works very well for testing one yield at a time, but it would be difficult to extend it to multiple magnitude readings, such as determining whether *all* the events in Table II are in compliance, since the estimators are highly correlated through the variance-covariance structure for α and β which appears in equation (5). That is, since the 22 lower bounds on the yield estimates are not independent of one another, we cannot simply count the number of times in Table III for which the lower limit on the yield estimate (for a given value of ε_0 and a given value of φ) exceeds 150 KT and conclude that this tally represents the number of events for which we are 95% certain that the 150 KT threshold was exceeded. For purposes of monitoring the TTBT, we are interested in determining whether the pattern of testing as a whole, as represented by all 22 events, has been in compliance. We must therefore use another approach as an alternative to this one of surrounding every yield estimate with a confidence interval.

3. Hypothesis Testing

Since the compliance testing scheme based on confidence intervals around the yield estimates cannot be used for testing ensembles of events, we shall examine as an alternative a test of the null hypothesis that all events in the data set are in compliance with the TTBT. Before we apply this hypothesis test to the data in Table II, we must adjust the test so that the false alarm probability (the probability that the compliance hypothesis will be rejected when in fact it is true) is some specified value, say 5%. If we choose to make the hypothesis test one of rejecting compliance if a certain configuration of estimated log yields W_1^{est}, W_2^{est}, ..., W_n^{est} exceed simultaneously some given threshold T, then the false alarm probability for this test will be determined by the value of T, and we should therefore adjust T so that the false alarm probability equals 5% or whatever other value we have specified for it. The threshold T for which the false alarm probability equals some chosen value such as 5% will depend on the particular configuration of threshold exceedences upon which the test is based (either 1 or more yields exceeding T, or 2 or more yields exceeding T, etc.). No matter what configuration of exceedences is used for the test, if the false alarm probability is to be low, then almost certainly we must set the threshold T for rejecting the compliance hypothesis to be some value which far exceeds the TTBT limit of 150 KT. That is, since the yield estimates are uncertain on account of the standard deviations of the parameters which are shown in Table I, there is a significant probability that any event for which we calculate the yield to be greater than 150 KT has in reality a yield that is less than 150 KT. In order not to make the error more often than, say, 5% of the time of wrongly concluding that the 150 KT limit has been exceeded, we must monitor the TTBT at some threshold T which is greater than 150 KT, and only when that greater threshold has been exceeded by some particular configuration of calculated yields can we reject the null hypothesis that the true yields are in compliance with the 150 KT limit. Although we shall later examine a more powerful hypothesis test involving a different configuration of threshold exceedences, let us for now adopt the suboptimal test of rejecting compliance for a set of n magnitudes if m or more calculated log yields exceed T.

3.1. Hypothesis Testing of at Least m Threshold Exceedences

We first choose to focus our attention on rejection regions where at least one of the estimated yields $W_1^{est}, \ldots, W_n^{est}$ exceeds T_α. The false alarm probability is then defined as

$$P_{FA} = 1 - \text{Prob}\,(\,W_1^{est} \le T,\ W_2^{est} \le T,\ \ldots,\ W_n^{est} \le T\,) \qquad (7)$$

given that the magnitudes are drawn from a particular compliance yield pattern $W_1, W_2, \ldots, W_n$, where $W_i \le 150$ KT for $i = 1, \ldots, n$. We shall examine first the case of tesing compliance for single events (which can also be done as by placing confidence limits around the estimated yields, as we have already shown), and then we shall examine the more important case of testing compliance for ensembles of events.

3.1.1. Null Hypothesis of a Single Compliance Yield

For a single event, the false alarm probability (7) is given by

$$P_{FA} = 1 - \text{Prob}\,(W_1^{est} \le T) = \Phi\left[\frac{W_1 - T}{(\sigma_{W_1}^2 + \sigma_1^2)^{1/2} / \beta_0}\right] \qquad (8)$$

where $\Phi(x)$ is the normal probability distribution function and where, by equation (5), we have

$$\sigma_{W_1}^2 = \sigma_\alpha^2 + 2W_1\sigma_{\alpha\beta} + W_1^2\sigma_\beta^2 \qquad (9)$$

For a given value of the single compliance log yield W_1 in equation (8), say W_1 = log(150 KT), one can plot the false alarm probability against the log yield threshold T to determine a threshold T_α for any given false alarm probability α. Once the threshold T has been set at T_α, the probability of "detecting" a violation of the form W', say

$$P_d(W') = \text{Prob}\,(W_1^{est} > T_\alpha) = \Phi\left[\frac{W'_1 - T_\alpha}{(\sigma_{W'_1}^2 + \sigma_1^2)^{1/2} / \beta_0}\right] \tag{10}$$

can then be plotted as a function of W'_1 to determine the power of the procedure.

3.1.2. Null Hypothesis of a Common Compliance Yield

We now turn to the more important case of testing the compliance hypothesis for an ensemble of events. As a particular example of the hypothesis test based on a single exceedence of the threshold T, let us assume that the null hypothesis specifies that all the tests were conducted at a single compliance yield, *i.e.*, that $W_1 = W_2 = \ldots = W_n = W \leq 150$ KT in equation (1). The false alarm probability defined by (7) then depends on the fact that the $W_1^{est}, \ldots, W_n^{est}$ defined by (4) have a joint multivariate normal distribution with means W and

$$\text{cov}(W_j^{est}, W_k^{est}) = (\delta_{jk}\,\sigma_j^2 + \sigma_W^2) / \beta_0^2 \tag{11}$$

with $\delta_{jk} = 1$ for $j = k$ and zero otherwise and

$$\sigma_W^2 = \sigma_\alpha^2 + 2W\,\sigma_{\alpha\beta} + W^2\sigma_\beta^2 \tag{12}$$

For an ensemble of events, we cannot take the product of terms like the right-hand side of equation (8), since the covariance structure (11) causes the yield estimates of the individual events within the ensemble to be mutually dependent. Instead, we note that the ensemble of estimated yields has a joint multivariate normal distribution, so we can evaluate equation (7) as (Johnson and Katz, 1972, p. 47)

$$P_{FA} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2/2} \prod_{j=1}^{n} \Phi\left[\frac{h_j - \lambda_j x}{\sqrt{1-\lambda_j}}\right] dx \tag{13}$$

where

$$h_j = \frac{W - T}{(\sigma_W^2 + \sigma_j^2)^{1/2} / \beta_0} \tag{14}$$

and

$$\lambda_j = [\, 1 + (\sigma_j^2 / \sigma_W^2)\,]^{-1/2} \tag{15}$$

The integral in equation (13) can be evaluated numerically using Gaussian quadrature. Note that equation (13) reduces to a product of the independent probabilities (8) if $\sigma_W^2 = 0$, *i.e.*, if the yield estimates are not correlated in equation (11).

3.1.3. Null Hypothesis of Multiple Compliance Yields

Although we have thus far considered only the null hypothesis that all the events have the same log yield W, we shall now consider more nearly realistic compliance scenarios as null hypotheses. For these cases as well, the correlation of the estimated yields for the different events will exert a strong influence on the calculation of the probabilities, but now we will not be able to use the preceeding formalism to take the correlation into account since equations (11) - (12) will no longer be applicable.

In the case of probabilities computed for multiple events, the false alarm and signal (*i.e.*, violation) detection probabilities are all functions of the following integral over all possible true (as opposed to estimated) values of the slope and intercept in equation (1), as it is evaluated under a fixed yield configuration $W_1, \ldots, W_n$:

$$\begin{aligned} &\text{Prob}\,(W_1^{est} \le T, \ldots, W_n^{est} \le T) \\ &\quad = \int\int \text{Prob}\,(W_1^{est} \le T, \ldots, W_n^{est} \le T \mid \alpha,\beta)\, f(\alpha,\beta)\, d\alpha\, d\beta \\ &\quad = \int\int \text{Prob}\,(m_1 \le m_T, \ldots, m_n \le m_T \mid \alpha,\beta)\, f(\alpha,\beta)\, d\alpha\, d\beta \\ &\quad = \int\int \prod_{j=1}^{n} \Phi\left[\frac{(\alpha_0 + \beta_0 T) - (\alpha + \beta W_j)}{\sigma_j}\right] f(\alpha,\beta)\, d\alpha\, d\beta \end{aligned} \tag{16}$$

Here m_T is the magnitude threshold corresponding to the yield threshold T and

$f(\alpha,\beta)$ is the joint density of α and β, which we assume to be bivariate normal with means α_0 and β_0 and with variances σ_α^2 and σ_β^2 and covariance $\sigma_{\alpha\beta}$ as given in Table I. We may thus substitute into equation (16)

$$f(\alpha,\beta) = f_1(\alpha) f_2(\beta \mid \alpha) \tag{17}$$

where

$$f_1(\alpha) = \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp\left\{-\frac{(\alpha-\alpha_0)^2}{2\sigma_\alpha^2}\right\} \tag{18}$$

$$f_2(\beta \mid \alpha) = \frac{1}{\sqrt{2\pi[\sigma_\beta^2 - (\sigma_{\alpha\beta})^2/\sigma_\alpha^2]}} \exp\left\{-\frac{[(\beta-\beta_0) - \sigma_{\alpha\beta}(\alpha-\alpha_0)/\sigma_\alpha^2]^2}{2[\sigma_\beta^2 - (\sigma_{\alpha\beta})^2/\sigma_\alpha^2]}\right\} \tag{19}$$

The double integral in equation (16) can then be evaluated using the following approximation :

$$\int_{-\infty}^{\infty} e^{-x^2} f(x)\, dx \approx \sum_{i=1}^{m} h_i f(x_i) \tag{20}$$

where the constants x_i are the zeroes of the mth order Hermite polynomial and the coefficients h_i are a tabulated weight function (Abramowitz and Stegun, 1964, p. 924). The double integral then reduces to the double sum

$$\frac{1}{\pi}\sum_{i=1}^{m}\sum_{k=1}^{m} h_i h_k \prod_{j=1}^{n} \Phi\left[\frac{(\alpha_0 + \beta_0 T) - (A_i + B_{i,k} W_j)}{\sigma_j}\right] \tag{21}$$

where

$$A_i = \alpha_0 + \sqrt{2}\sigma_\alpha x_i \tag{22}$$

$$B_{i,k} = \beta_0 + \sqrt{2}(\sigma_{\alpha\beta}/\sigma_\alpha)x_i + \sqrt{2[\sigma_\beta^2 - (\sigma_{\alpha\beta})^2/\sigma_\alpha^2]}\,x_k \tag{23}$$

We note that the double integral (16) is more computationally intensive than is the single integral for the special case (13), which can also be evaluated using the approximation (20). We note also that although in Table I we have chosen to use a single value of σ_j for all 22 events, equation (21) shows that the hypothesis testing procedure could incorporate a different value of σ_j for each event.

When the fixed set of yields is entirely in compliance, *i.e.*, when $W_i \leq 150$ KT for $i = 1, \ldots, n$, then the complement of the probability in equation (16) gives the false alarm probability P_{FA}, *i.e.*, the probability that none of the estimated yields for this "compliance" distribution exceeds the threshold T. In order to compute the probability that the estimated yield of the kth event exceeds this threshold while the other estimated yields are less than T, we re-write equation (16) as

$$\text{Prob}\,(W_1^{est} \leq T, \ldots, W_k^{est} \geq T, \ldots, W_n^{est} \leq T) =$$

$$\int\int \prod_{j=1}^{n} \Phi \left[c_j \frac{(\alpha_0 + \beta_0 T) - (\alpha + \beta W_j)}{\sigma_j} \right] f(\alpha,\beta) d\alpha d\beta \qquad (24)$$

where $c_j = 1$ for $j \neq k$ and $c_k = -1$ for the kth event. In order to compute the probability that this distribution of yields results in exactly one estimated yield which exceeds the threshold T, we sum the n individual probabilities as follows :

$$\text{Prob}\,(N_{exc} = 1) = \sum_{k=1}^{n} \text{Prob}\,(W_1^{est} \leq T, \ldots, W_k^{est} \geq T, \ldots, W_n^{est} \leq T) \qquad (25)$$

where N_{exc} is the number of exceedences of T and where each term in the sum is given by equation (24). The probability that exactly 2 estimated yields exceed T is

$$\text{Prob}\,(N_{exc} = 2) = \sum_{j=1}^{n} \sum_{\substack{k=1 \\ k \neq j}}^{n} \text{Prob}\,(W_1^{est} \leq T, \ldots, W_j^{est} \geq T, \ldots, W_k^{est} \geq T, \ldots, W_n^{est} \leq T) \qquad (26)$$

and so on for greater numbers of exceedences of T. For hypothesis testing, we will not be concerned so much with the probabilities of *exactly* N exceedences of the threshold T as with the probabilities of *at least* N exceedences; these probabilities are given by

$$\begin{aligned} \text{Prob}\,(N_{exc} \geq 1) &= 1 - \text{Prob}\,(N_{exc} = 0) \\ \text{Prob}\,(N_{exc} \geq 2) &= \text{Prob}\,(N_{exc} \geq 1) - \text{Prob}\,(N_{exc} = 1) \\ \text{Prob}\,(N_{exc} \geq 3) &= \text{Prob}\,(N_{exc} \geq 2) - \text{Prob}\,(N_{exc} = 2) \end{aligned} \qquad (27)$$

etc., where Prob ($N_{exc} = 0$) is given by equation (16). These same calculations, of course, can be performed for any given distribution of the yields W_i, not just the "compliance" distribution which is used as the null hypothesis in the test. In particular, they can be used to compute the signal detection probability for any given "violation" distribution, in which one or more of the W_i are greater than 150 KT.

The main items of interest from the compliance hypothesis testing point of view would be first to find a threshold T_α which gives a false alarm probability

$$\alpha = P_{FA} = 1 - \text{Prob}\ (N_{exc} = 0) = \text{Prob}\ (N_{exc} \geq 1) \qquad (28)$$

with W_1, W_2, ..., W_n a set of compliance yields, and then to compute the probability of a single exceedence Prob ($N_{exc} \geq 1$) with the threshold T set at T_α and with $W_k = W' > 150$ KT for the kth event. This would determine the signal detection probability (for the given false alarm probability α) for a single violation of yield size W'. This can be plotted as a function of W' to determine the probabilities under current knowledge of detecting single violations of log yield W' with a false alarm probability equal to α. Of course, this is only one type of alternative and one should design the test so that the most useful (for weapons development) pattern of violations will be detected with high probability. In this report we have performed tests for only two patterns of violations, namely that of one or more yields exceeding the threshold or that of two or more yields exceeding it.

3.1.4. Application to Observed Magnitudes

We shall now apply the hypothesis testing procedure to the observed magnitudes listed in Table II. We must first select a "compliance" distribution of assumed yields to use as the null hypothesis. As an example of the simple case which we have discussed of a common yield for all the events, we shall make our null hypothesis be that all the events are at the limit of TTBT compliance : W_1 =

W_2 = ... = W_{22} = log(150KT). We choose to make the hypothesis test one of rejecting this null hypothesis test if one or more estimated log yields W_i^{est} exceed some threshold T. Substituting different values of T into the formula (21), we find by trial and error those values which make the false alarm probability α in equation (28) take on the following desired values : 5%, 10%, 20%, and 50%. Rather than expressing the results of these calculations in terms of the log yield thresholds T_α, we list in Table IV values of the yield thresholds Y_T which are obtained by raising 10 to the power T_α for each of the 4 chosen values of α. As is shown in equations (22) and (23), the computation of α for any given value of T is strongly affected by the value which is used for σ_α and hence, through equation (3), by the standard deviation φ of the estimate of the USSR - US magnitude bias. Table IV thus lists a separate set of results for each of the 2 trial values which were chosen for this parameter. We can now accept or reject the null hypothesis, at any of the 4 tabulated levels of significance α, by seeing whether one or more of the calculated yields in Table III exceed the corresponding threshold Y_T. An equivalent but simpler procedure is not to convert the observed magnitudes into estimated yields but to convert the yield thresholds into magnitude thresholds and then see whether any of the observed magnitudes exceed these values. The magnitude thresholds m_T corresponding to T_α, and hence to Y_T, are listed in Table IV. The conversion from yield to magnitude has been performed for each of the three trial values of the magnitude bias ε_0, in accordance with equation (3). Comparing the magniutdes listed in Table II with the thresholds m_T, we see that we would reject the null hypothesis that all the yields are 150 KT with a false alarm probability of 10% if we use ε_0 = 0.20 and φ = 0.05 and with a false alarm probability of 20% if we use ε_0 = 0.20 and φ = 0.10. For the other 4 trial values of the estimate of the magnitude bias and its uncertainty $\varepsilon_0 \pm \varphi$, we can reject the null hypothesis only with a false alarm probability greater than 20%.

TABLE IV

Yield thresholds Y_T at which the TTBT can be monitored with a given false alarm probability α. The null hypothesis is that all yields are 150 KT; it will be rejected if 1 or more calculated yields exceed the threshold Y_T.

Case 1 : $\varphi = 0.05$

α	Y_T	m_T ($\varepsilon_0 = 0.20$)	m_T ($\varepsilon_0 = 0.25$)	m_T ($\varepsilon_0 = 0.30$)
5%	278.0	6.257	6.307	6.357
10%	261.3	6.233	6.283	6.333
20%	242.7	6.205	6.255	6.305
50%	211.6	6.152	6.202	6.252

Case 2 : $\varphi = 0.10$

α	Y_T	m_T ($\varepsilon_0 = 0.20$)	m_T ($\varepsilon_0 = 0.25$)	m_T ($\varepsilon_0 = 0.30$)
5%	334.5	6.329	6.379	6.429
10%	303.0	6.290	6.340	6.390
20%	267.0	6.242	6.292	6.342
50%	212.3	6.153	6.203	6.253

TABLE V

3 assumed distributions of yields (KT) to use as the null (compliance) hypothesis

if $\varepsilon_0 = 0.20$:	if $\varepsilon_0 = 0.25$:	if $\varepsilon_0 = 0.30$:
131.	115.	101.
130.	114.	100.
48.	42.	37.
141.	124.	109.
150.	150.	150.
94.	83.	73.
150.	150.	150.
150.	150.	150.
139.	122.	108.
142.	125.	110.
150.	150.	150.
67.	59.	52.
150.	150.	150.
116.	102.	89.
127.	111.	98.
103.	90.	79.
127.	111.	98.
150.	148.	130.
150.	137.	120.
150.	150.	150.
150.	150.	139.
150.	150.	150.

The null hypothesis which we have used in Table IV represents the "worst possible case" compliance scenario for TTBT monitoring, namely that every event has the maximum allowed yield. The thresholds T_α computed for this case are thus larger than they would be for a more nearly realistic compliance scenario in which only a fraction of the 22 events approach the 150 KT limit. We may therefore take the values listed in Table IV to represent upper bounds to the thresholds which would be computed using other, more plausible, compliance yield distributions as the null hypothesis.

We thus see that a critical step in performing the hypothesis testing calculations is the assigning of assumed values to the fixed yield distribution which is to be used for the null (compliance) hypothesis. It should be noted that this step would not be a problem under a Bayesian formulation with a uniform prior on the yields, wherein the false alarm probability calculation becomes a simple posterior probability. We shall now repeat the hypothesis test, this time using a different null hypothesis which we believe to be closer to the truth than the assumption that all 22 explosions have a yield of 150 KT. For the fixed yield distribution $W_1, \ldots, W_{22}$ we suggest that the observed magnitudes m_j of the USSR explosions be converted to a hypothetical distribution of compliance yields by setting $W_j \equiv (m_j - \alpha_0)/\beta_0$ if this calculated yield is less than 150 KT and setting $W_j \equiv$ 150 KT if it is greater than that. The set of values of W_j which is obtained in this manner is listed in Table V for our 3 trial values of the parameter ε_0. This fixed distribution of assumed yields resembles the calculated yield distribution of Table III except that all yields are forced to be in compliance. We believe that it is a more realistic assumption than the previously discussed case in which all the hypothesized yields are equal. An alternative approach would be to shift all the calculated yields down by a constant amount so that every event is below the 150 KT threshold. This approach, however, is tantamount to assuming that the value of α_0 is in error, so we would then be inconsistent in using that value in equations (21) - (23). We therefore reject the alternative of creating the hypothetical "compliance" yield distribution by displacing all the computed yields downward by a fixed amount in favor of the approach of truncating the calculated yields at 150 KT. One shortcoming of our approach is that it would caus number of the yields to be bunched up at the 150 KT cut-off limit if some of the observed magnitudes were to exceed the magnitude threshold corresponding to that yield threshold. This is unlikely to be a valid model of the true distribution of yields, and it results in larger values of the thresholds T_α than those which would be calculated if fewer compliance yields were as large as 150 KT. This problem is of course less severe than it was in the previously discussed case in which *all* the yields were assumed to be 150 KT. It could perhaps be addressed by setting the compliance yields equal to, say, 140 KT if the calculated yields are greater than that. We have not investigated this possibility. It is possible to formulate a number of other schemes whereby the observed magnitude distribution can be used to generate a hypothesized distribution of compliance yields. Herein we shall consider only the approach of truncating the distribution of calculated yields at 150 KT.

It should be noted that choosing the set of compliance yields on the basis of the observed data, as we have described above, will in fact violate the assumptions under which the false alarm probability has been calculated. In order for these calculations to be valid, the fixed set of compliance yields should be chosen independently of the data under examination. The statistical implications of choosing the compliance yields for the null hypothesis on the basis of the observations, as we have done herein, have not been examined. We have calculated the false alarm rates and signal detection probabilities in the manner described previously, as if the null hypothesis had been chosen in advance of examination of the data. It should be remembered, however, that the fixed set of compliance yields was actually not selected in advance, so these calculations are not strictly valid. This criticism, of course, does not apply to the results shown in Table IV, which were computed using an *a priori* assumption about the yield distribution, albeit a rather implausible one.

Table VI shows the magnitude and yield thresholds for rejection of the null hypothesis that the yield distribution is as listed in Table V. As in Table IV, the thresholds differ for the 2 trial values of φ, but now they also differ for the 3 trial values of ε_0, since the null hypotheses in Table V are different for these 3 cases. We thus have listed in Table VI 6 sets of thresholds for rejecting the null (compliance) hypothesis with various false alarm probabilities α. As we had anticipated, the thresholds are lower than are those in Table IV which were computed using the null hypothesis that all the events are barely in compliance. Since the largest magnitude in Table II is 6.242, we see that the null hypothesis is rejected with a false alarm probability of 5% if the magnitude bias is taken to be 0.20 ± 0.05 and with a false alarm probability of 20% if it is taken to be either 0.20 ± 0.10 or 0.25 ± 0.05. For the other 3 cases considered in Table VI, the null hypothesis that the yields are all in compliance cannot be rejected unless the false alarm probability is allowed to be greater than 20%.

The hypothesis tests which have been applied so far have all had as the criterion for rejection that one or more estimated log yields exceed a threshold T. We shall now examine a different test, which has as its criterion for rejection that 2 or more estimated log yields exceed a threshold T. The threshold T will of course now be different (for the same false alarm probability α), since it is less likely that at least 2 yields exceed some value than that at least one yield exceeds it. The procedure for evaluating the thresholds T_α for this new hypothesis test were shown in equations (24) - (27). Using the null hypotheses that the yields (all in compliance with the TTBT) are as listed in Table V, we find that the thresholds for applying this test are as shown in Table VII. The 2 largest magnitudes in Table II are 6.222 and 6.242, so we see that the compliance

TABLE VI

Yield thresholds Y_T at which the TTBT can be monitored with a given false alarm probability α. The null hypothesis is that the yields are as listed in Table V; it will be rejected if 1 or more calculated yields exceed the threshold Y_T.

Case 1 : $\varepsilon_0 = 0.20$ and $\varphi = 0.05$

α	Y_T	m_T
5%	266.0	6.240
10%	249.5	6.215
20%	231.4	6.186
50%	201.1	6.132

Case 2 : $\varepsilon_0 = 0.20$ and $\varphi = 0.10$

α	Y_T	m_T
5%	319.0	6.310
10%	288.8	6.272
20%	254.5	6.223
50%	201.7	6.133

Case 3 : $\varepsilon_0 = 0.25$ and $\varphi = 0.05$

α	Y_T	m_T
5%	261.7	6.284
10%	245.3	6.259
20%	227.2	6.229
50%	197.0	6.174

Case 4 : $\varepsilon_0 = 0.25$ and $\varphi = 0.10$

α	Y_T	m_T
5%	313.5	6.354
10%	283.5	6.315
20%	249.8	6.266
50%	197.6	6.175

TABLE VI
(continued)

Case 5 : $\varepsilon_0 = 0.30$ and $\varphi = 0.05$

α	Y_T	m_T
5%	257.5	6.328
10%	241.3	6.302
20%	223.2	6.272
50%	193.0	6.216

Case 6 : $\varepsilon_0 = 0.30$ and $\varphi = 0.10$

α	Y_T	m_T
5%	307.8	6.397
10%	278.3	6.358
20%	245.2	6.309
50%	193.6	6.217

TABLE VII

Yield thresholds Y_T at which the TTBT can be monitored with a given false alarm probability α. The null hypothesis is that the yields are as listed in Table V; it will be rejected if 2 or more calculated yields exceed the threshold Y_T.

Case 1 : $\varepsilon_0 = 0.20$ and $\varphi = 0.05$

α	Y_T	m_T
5%	237.3	6.196
10%	224.3	6.174
20%	209.6	6.148
50%	184.3	6.098

Case 2 : $\varepsilon_0 = 0.20$ and $\varphi = 0.10$

α	Y_T	m_T
5%	288.0	6.271
10%	262.0	6.234
20%	230.7	6.185
50%	184.6	6.099

Case 3 : $\varepsilon_0 = 0.25$ and $\varphi = 0.05$

α	Y_T	m_T
5%	232.0	6.237
10%	219.1	6.215
20%	204.5	6.189
50%	179.5	6.138

Case 4 : $\varepsilon_0 = 0.25$ and $\varphi = 0.10$

α	Y_T	m_T
5%	281.0	6.311
10%	255.5	6.275
20%	225.2	6.226
50%	179.8	6.139

TABLE VII
(continued)

Case 5 : $\varepsilon_0 = 0.30$ and $\varphi = 0.05$

α	Y_T	m_T
5%	227.0	6.279
10%	214.2	6.256
20%	199.7	6.229
50%	174.9	6.178

Case 6 : $\varepsilon_0 = 0.30$ and $\varphi = 0.10$

α	Y_T	m_T
5%	274.3	6.352
10%	249.5	6.315
20%	219.8	6.266
50%	175.2	6.179

hypothesis is rejected at a significance level of α = 5% if the magnitude bias is 0.20 ± 0.05, at α = 10% if it is 0.25 ± 0.05, at α = 20% if it is 0.20 ± 0.10, and at less significant levels for the other 3 possibilities examined in that table.

Having applied the hypothesis testing procedure to the observed magnitudes, we now ask the question : if in fact there was a TTBT violation, how likely is it that this procedure would have detected it ? For each of the hypothesis tests which we have examined, we may answer this question of effectiveness by again using the formalism in equations (24) - (27), this time setting the threshold T to the value T_α which we have already found and using a hypothesized distribution of yields W_1, W_2, ..., W_{22} which violates the TTBT in some way. We demonstrate this procedure in Table VIII, using the hypothesis test that was used to generate Table VI (null hypothesis of yields as listed in Table V, rejection if one or more log yields exceed T). We shall evaluate the power of this hypothesis test for detecting a TTBT violation consisting of a single event with yield greater than 150 KT, the other 21 events having the (compliance) yields which were hypothesized for them in Table V. We shall take the violating event to be one of those which was assumed in Table V to be 150 KT. In Table VIII we show the increasing power of the hypothesis test to detect the violation as we increase its yield from 175 KT to 500 KT. For example, the table shows that if the magnitude bias is 0.20 ± 0.05 and if we displace one of the 150 KT yields in Table V upward to 175 KT, then there is a probability of 0.070 that at least one observed magnitude will exceed the threshold of m_T = 6.240 which was set in Table VI for monitoring TTBT compliance with a false alarm probability of 5%. If the violating event has a yield of 200 KT, then the probability that the compliance hypothesis will be rejected at the significance level α = 5% is 0.132. If we choose to monitor the TTBT with a false alarm probability of 10% and if the yield of the single violation is 200 KT, then there is a probability of 0.217 that at least one observed magnitude will exceed the threshold m_T = 6.215 and that the

TABLE VIII

"Signal" (violation) detection probabilities for TTBT monitoring with various false alarm probabilities α when only one event is not in compliance. The hypothesis test is the one which was used in Table VI.

Case 1 : $\varepsilon_0 = 0.20$ and $\varphi = 0.05$

α	signal detection probability if the yield (KT) of the single violation is : 175.	200.	250.	300.	400.	500.
5%	.070	.132	.402	.709	.966	.998
10%	.132	.217	.521	.800	.983	.999
20%	.246	.352	.661	.882	.993	1.000
50%	.556	.659	.869	.968	.999	1.000

Case 2 : $\varepsilon_0 = 0.20$ and $\varphi = 0.10$

α	signal detection probability if the yield (KT) of the single violation is : 175.	200.	250.	300.	400.	500.
5%	.061	.092	.227	.426	.763	.922
10%	.117	.162	.334	.551	.848	.958
20%	.226	.288	.493	.702	.923	.983
50%	.533	.599	.772	.898	.985	.998

Case 3 : $\varepsilon_0 = 0.25$ and $\varphi = 0.05$

α	signal detection probability if the yield (KT) of the single violation is : 175.	200.	250.	300.	400.	500.
5%	.074	.144	.429	.733	.971	.998
10%	.137	.232	.549	.820	.986	.999
20%	.254	.372	.688	.897	.994	1.000
50%	.566	.677	.885	.974	.999	1.000

TABLE VIII
(continued)

Case 4 : $\varepsilon_0 = 0.25$ and $\varphi = 0.10$

α	signal detection probability if the yield (KT) of the single violation is :					
	175.	200.	250.	300.	400.	500.
5%	.063	.098	.243	.447	.779	.929
10%	.121	.171	.354	.574	.862	.963
20%	.231	.300	.514	.722	.931	.986
50%	.539	.613	.789	.909	.987	.998

Case 5 : $\varepsilon_0 = 0.30$ and $\varphi = 0.05$

α	signal detection probability if the yield (KT) of the single violation is :					
	175.	200.	250.	300.	400.	500.
5%	.078	.157	.456	.756	.975	.998
10%	.144	.248	.576	.838	.988	.999
20%	.263	.392	.713	.911	.995	1.000
50%	.577	.696	.899	.978	.999	1.000

Case 6 : $\varepsilon_0 = 0.30$ and $\varphi = 0.10$

α	signal detection probability if the yield (KT) of the single violation is :					
	175.	200.	250.	300.	400.	500.
5%	.066	.105	.260	.470	.796	.937
10%	.125	.181	.375	.596	.874	.968
20%	.237	.313	.536	.741	.939	.988
50%	.546	.627	.805	.919	.989	.999

violation will thereby be detected.

The probabilities which are listed in Table VIII conform to a pattern which we would expect. When the yield of the single violation is small, the distribution of the 21 assumed yields differs only slightly from the null hypothesis, so the probability that one or more estimated log yields exceeds the threshold T is only slightly greater than the false alarm probability. When the yield of the single violation is great, the probability that the compliance hypothesis will be rejected approaches unity. We are also able to estimate the signal detection probability in one intermediate case. If we set the threshold T so high that the probability of a false alarm is negligible, then the probability of detecting a single violation which has log yield exactly equal to T is 50%. We may verify this for "Case 1" in Tables VI and VIII. If we set the false alarm probability α at 1%, we find that the yield threshold for rejecting the null hypothesis is 300.0 KT. If we change the hypothesized yield distribution in Table V so that one of the events has a yield of 300 KT rather than 150 KT, then we find that the probability that one or more estimated yields exceed 300 KT is 0.503.

Note that the "detection" in Table VIII is a detection that the TTBT has been violated by the ensemble of the 22 events as a whole rather than by the single event which we have chosen to be in violation. That is, there is a chance that the event which we have set to have the yield greater than 150 KT will have a magnitude less than the threshold m_T, and there is a chance (very nearly equal to the false alarm probability α) that one or more of the other events have a magnitude greater than m_T, so we cannot definitely say that it was the violating event which led to rejection of the compliance hypothesis. Of course, as the yield of the violating event is allowed to increase, most of the power of the test shown in Table VIII is due to the probability that the magnitude of that particular event will exceed m_T.

We can, of course, use other violation patterns such as one event with a yield of 200 KT and another with a yield of 250 KT as an alternative to the single-event violation pattern which was used in Table VIII. We can also calculate the "signal" (*i.e.*, violation) detection probability for the different hypothesis tests which were used in Tables IV and VII. The violation pattern which is used in these calculations need not be the same one as that upon which rejection of the hypothesis is based; for instance, we can calculate the signal detection probability of the "2 or more exceedences" test shown in Table VII using a single-event violation pattern as we did in Table VIII. Of course, we do not expect that this particular hypothesis test would be very powerful for detecting this particular violation pattern. We thus see that the hypothesis test should be designed to reject the compliance hypothesis for some TTBT violation pattern which we expect would be useful for weapons development and that the "one or more exceedences" test which was used in Tables IV and VI may not be the most powerful one for detecting this pattern of violations. We shall next examine the question of what is the most powerful hypothesis test in general.

3.2. Optimal Hypothesis Test

It is clear that the threshold value T_α can be determined for any number of events n by setting the false alarm probability at α, but it is not clear what kind of violation we would like to detect. Any assumed pattern of violations can be converted to a signal detection probability for the given false alarm probability α. For example, in the hypothesis test in which the log yields are all equal to the same compliance value W, one pattern might be systematic cheating, *i.e.*, $W_j = W + \Delta$ for all j, or another pattern might be a selective outlier violation, *i.e.*, $W_j = W$ for $j \neq k$ and $W_k = W + \Delta$ for the kth event. So far we have set thresholds T_α by considering only the particular violation patterns of one or more, or two or more, estimated yields W_j^{est} exceeding T_α. We now consider more general violation patterns.

One point of interest is that statistical theory gives an optimal test for any given null hypothesis

$$H_1 : \mathbf{W} = \mathbf{W}_1$$

against

$$H_2 : \mathbf{W} = \mathbf{W}_2$$

where $\mathbf{W} = (W_1, W_2, \ldots, W_n)'$ is the nx1 yield vector. The Neyman-Pearson lemma gives the most powerful test of H_2 against H_1 for a given false alarm probability. This means that the probability of detecting a violation of the form H_2 is highest for the given false alarm probability. This test, which is given below, is not equivalent to the one discussed earlier with false alarm probability (7) which is based on rejection if one or more magnitudes exceed T. The optimal test, however, leads to intractable computations unless certain approximations are made.

Suppose that we wish to develop the test of H_2 against H_1 which implies that under H_i, the distribution of $\mathbf{m} = (m_1, m_2, \ldots, m_n)'$ is multivariate normal with mean

$$\mu_i = \mathbf{1}\, \alpha_0 + \mathbf{W}_i\, \beta_0 \tag{29}$$

and covariance matrix

$$\Sigma_i = \begin{bmatrix} \mathbf{1} \\ \mathbf{W}_i \end{bmatrix} \Sigma\, (\mathbf{1}', \mathbf{W}_i') + \mathrm{Diag}\, (\sigma_1^2, \ldots, \sigma_n^2) \tag{30}$$

where $\mathbf{1}' = (1, 1, \ldots, 1)$ and $\mathbf{W}_i' = (W_{i1}, \ldots, W_{in})$ and $\mathrm{Diag}\, (\sigma_1^2, \ldots, \sigma_n^2)$ denotes an $n x n$ diagonal matrix. If we develop the test, it leads to rejecting compliance when

$$Q(\mathbf{m}) = \frac{1}{2} \mathbf{m}' \,(\Sigma_1^{-1} - \Sigma_2^{-1})\, \mathbf{m} + (\mu'_2 \Sigma_2^{-1} - \mu'_1 \Sigma_1^{-1})\, \mathbf{m} \tag{31}$$

exceeds some threshold. The distribution theory is difficult for this case and even if one assumes $\Sigma = \Sigma_1 \equiv \Sigma_2$, the resulting linear discriminant function still depends on the assumed yield vectors $\mathbf{W}_1$ and $\mathbf{W}_2$, *i.e.*,

$$Q'(\mathbf{m}) = \beta_0(\mathbf{W}_2 - \mathbf{W}_1)' \Sigma^{-1} \mathbf{m} \quad . \tag{32}$$

The false alarm and signal detection probabilities can be specified for any $\mathbf{W}_1$ and $\mathbf{W}_2$ since $Q'(\mathbf{m})$ will have a univariate normal distribution with means

$$\gamma_i = \beta_0(\mathbf{W}_2 - \mathbf{W}_1)' \Sigma^{-1} \mu_i \tag{33}$$

under H_i, $i = 1,2$ and variance

$$\sigma_Q^2 = \beta_0^2(\mathbf{W}_2 - \mathbf{W}_1)' \Sigma^{-1} (\mathbf{W}_2 - \mathbf{W}_1) \tag{34}$$

under both hypotheses. In this case, the problem with the critical region defined by (32) is that the value of $Q'(\mathbf{m})$ can be manipulated by running a series of tests at a low level to keep the linear function below the critical value for rejecting the null hypothesis, as can be seen by considering the behavior of the right-hand side of (32) when several of the n components of $\mathbf{m}$ are small. This is in contrast to the hypothesis test discussed previously for which the null hypothesis was rejected if there were one or more exceedences of the threshold T_a. Since we are not playing against nature, it would seem to be preferable to stick with suboptimal critical regions of the form implied by the false alarm rate (7), since they will be sensitive to one-time only violations and are not subject to manipulation by an adversary.

We have thus far considered the coefficients α_0 and β_0 to be given quantities, and we have not addressed the problem of how they might be calculated from the calibration data of magnitudes for explosions with known yields. We shall now present a technique for determining these parameters which treats both the magnitudes and the yields of the calibration explosions as random variables. First, we shall consider the simpler approach of treating only the magnitudes as random variables. We shall finally apply both of these techniques to a particular data set in order to obtain the results which were listed in Table I and which were used in the last section of this report for testing TTBT compliance.

1. Weighted Least-Squares Solution

We adopt equation (1) as a model of the magnitude-yield relation, where we consider the log yields W_j to be fixed quantities and the magnitudes m_j to be measurements characterized by errors e_j of size σ_j. That is, the W_j are assumed to be known with arbitrary precision whereas the m_j are treated as random variables having normal distributions with means $\alpha + \beta W_j$ and variances σ_j^2. We may then obtain estimates of the parameters α and β by the standard technique of weighted least squares (where the weights are the reciprocals of the variances σ_j^2). We define :

$$\begin{aligned} S_1 &= \sum 1/\sigma_j^2 \\ S_2 &= \sum W_j/\sigma_j^2 \\ S_3 &= \sum W_j^2/\sigma_j^2 \\ S_4 &= \sum m_j/\sigma_j^2 \\ S_5 &= \sum m_j W_j/\sigma_j^2 \end{aligned} \tag{35}$$

We obtain estimates of α and β from

$$\begin{bmatrix} S_1 & S_2 \\ S_2 & S_3 \end{bmatrix} \begin{bmatrix} \alpha^{est} \\ \beta^{est} \end{bmatrix} = \begin{bmatrix} S_4 \\ S_5 \end{bmatrix} \tag{36}$$

which results in

$$\alpha^{est} = \frac{S_3S_4 - S_2S_5}{S_1S_3 - S_2^2} \tag{37}$$

$$\beta^{est} = \frac{S_1S_5 - S_2S_4}{S_1S_3 - S_2^2} \tag{38}$$

The variance-covariance structure of the least-squares estimates is given by

$$\begin{bmatrix} \sigma_\alpha^2 & \sigma_{\alpha\beta} \\ \sigma_{\alpha\beta} & \sigma_\beta^2 \end{bmatrix} = \begin{bmatrix} S_1 & S_2 \\ S_2 & S_3 \end{bmatrix}^{-1} = \frac{1}{S_1S_3 - S_2^2} \begin{bmatrix} S_3 & -S_2 \\ -S_2 & S_1 \end{bmatrix} \tag{39}$$

As in equation (3), we must add to σ_α^2 a term φ^2 to account for the variance in the estimate of the magnitude bias between the calibration test site and the test site at which the magnitude-yield relation is to be applied. Since we are concerned here only with the calibration data, we shall not include this additional term except when we make a comparison with the values which were computed using equation (3).

2. Maximum Likelihood Estimation

We shall now consider an improved model of the magnitude-yield relation which treats the yields of the calibration explosions as being measured quantities which are subject to some uncertainty. Another feature of this improved model is the incorporation of a term which explicitly takes coupling into account rather than treating it as a component of the magnitude measurement error. The new model of the magnitude-yield relation is

$$m_j = \alpha + \beta W_j + e_j + c_j \tag{40}$$

The term e_j represents the error in the magnitude m_j. As before, this random variable is normally distributed with mean zero and variance σ_j^2. The log yield W_j is now considered to be a random variable also. We define its mean and variance as μ_j and τ_j^2, respectively. Equation (40) differs from equation (1) in that an additional term, c_j, has been included to treat the variability of source coupling among the n calibration events. In equation (1), for which the mean values α_0 and β_0 of the random variables α and β were specified *a priori*, the magnitude scatter due to coupling was incorporated into the magnitude error term e_j. As was explained in the discussion of Equation (1), the variance σ_j^2 of the error term e_j comes from two sources : the error in the measurement of m_b, and the scatter in m_b for a fixed yield due to variations in coupling efficiency. In order to derive new estimates α^{est} and β^{est} as alternatives to the *a priori* values α_0 and β_0, in equation (40) it is necessary to separate the two components of the variance, since the magnitude measurement variance is known for each event (given the number of detecting stations and the m_b residual at each station) but the coupling scatter is unknown. We denote the variance of the magnitude error e_j as σ_j^2 and the variance of the coupling term c_j as σ_c^2. Note that we assume this variance to be a single constant for the n calibration explosions. The model would require further modification if it were believed that the coupling variability were greater among some subsets of the n explosions than among others. This would be the case, for example, if the calibration explosions were detonated at geologically dissimilar regions within the test site, and it would be especially important if the source media were not the same for all the explosions. In this study, we shall assume that the geological heterogeneity among the n sources can be characterized by the single value σ_c.

The regression which is treated by the new model differs from the one which was solved by weighted least squares in that now both the magnitudes and the yields of the explosions are regarded as random variables. (The model described in the previous section corresponds to the case $\tau_j^2 \equiv 0$.) The model of the magnitude-yield relation in equation (40) is thus characterized by the known means and variances μ_j and τ_j^2 of the log yields W_j, the known variances σ_j^2 of the magnitude errors e_j, and the unknown variance σ_c^2 of the coupling scatter

c_j. We obtain maximum-likelihood estimates for the parameters α and β in this model by an iterative approach. As a first approximation, we take α and β to be the values which were found using the weighted least-squares method described previously. An initial value will be guessed for the coupling variance σ_c^2. We then define

$$\eta_j^2 = \beta^2 \tau_j^2 + \sigma_j^2 + \sigma_c^2 \tag{41}$$

The likelihood function for estimating α and β is given by

$$\log L = -\frac{1}{2}\sum_{j=1}^{n} \log(\eta_j^2) - \frac{1}{2}\sum_{j=1}^{n} (m_j - \alpha - \beta\mu_j)^2 / \eta_j^2 \tag{42}$$

We shall obtain estimates for α and β (simultaneously with estimates for the log yields W_j) by maximizing this function. The variances and covariances of the estimates of W_j and e_j are given by

$$\begin{aligned} \sigma_{W_j}^2 &= \tau_j^2(1 - \beta^2\tau_j^2/\eta_j^2) \\ \sigma_{e_j}^2 &= \sigma_j^2(1 - \sigma_j^2/\eta_j^2) \\ \sigma_{W_j e_j} &= -\beta^2\tau_j^2\sigma_j^2/\eta_j^2 \end{aligned} \tag{43}$$

The estimates of the random variables W_j and e_j themselves are given by

$$W_j^{est} = \mu_j + (m_j - \alpha - \beta\mu_j)\beta\tau_j^2/\eta_j^2 \tag{44}$$

$$e_j^{est} = (m_j - \alpha - \beta\mu_j)\sigma_j^2/\eta_j^2 \tag{45}$$

We denote the averages over the n explosions of W_j^{est} and e_j^{est} as W_{ave}^{est} and e_{ave}^{est} and the average of m_j as m_{ave}. We then obtain as new estimates of the parameters β and α

$$\beta = \frac{\sum_{j=1}^{n}(W_j^{est} - W_{ave}^{est})(m_j - m_{ave}) - \sum_{j=1}^{n}[(W_j^{est} - W_{ave}^{est})(e_j^{est} - e_{ave}^{est}) + \sigma_{W_j e_j}]}{\sum_{j=1}^{n}[(W_j^{est} - W_{ave}^{est})^2 + \sigma_{W_j}^2]} \tag{46}$$

$$\alpha = m_{ave} - \beta W_{ave}^{est} - e_{ave}^{est} \tag{47}$$

Finally, we obtain a new value for the variance of the coupling scatter as

$$\sigma_c^2 = \frac{1}{n} \sum_{j=1}^{n} [(m_j - \alpha - \beta W_j^{est})^2 + \beta^2 \sigma_{W_j}^2 + 2\beta \sigma_{W_j e_j} + \sigma_{e_j}^2] \tag{48}$$

The sequence of equations (41) and (43) - (48) is repeated until the estimates of all the parameters converge and the likelihood function (42) stabilizes.

3. Application to Calibration Data

We shall apply the regression analysis which we have just described to the data listed in Table IX. As is indicated in that table, the events which we shall use for calibration were detonated at physically distinct test sites and in different media. In order for the calibration to have relevance for the Shagan River events in Table II, we should restrict the calibration events to those detonated in granite, since this source medium is presumed to have coupling properties similar to those of the emplacement medium at Shagan. Due to the paucity of data for explosions in granite, however, we shall ignore the restriction of a single source medium and we shall use for calibration the collection of events listed in Table IX. Although this procedure may prove inadequate for calibrating the magnitude-yield regression in practice, it will serve our needs of generating the values which were listed in Table I and which were used for illustration purposes in our examples of TTBT compliance testing. The magnitudes and yields which are listed in Table IX were taken from Blandford *et al.* (1983). The magnitudes were computed by means of a generalized linear model [GLM] which permits the calculation of station magnitude effects simultaneously with the event magnitudes and corrections for the clustering of events at a common test site. (This latter effect leads to correlations in the data.) Corrections were also made to the magnitudes for the attenuation parameter t^* and for interference between the P and pP pulses. The t^* corrections reduce the data to an equivalent single test site (NTS), so the magnitude bias ε_0 which we have used for

TABLE IX

Events used for magnitude - yield regression
(data taken from Blandford *et al.*, 1983)

event	date	site	medium	Y (KT)	m_b	σ_j
SHOAL	10/26/63	Fallon, Nev.	granite	12.2	4.85	0.03
REX	2/24/66	Pahute Mesa	tuff	19.0	4.94	0.03
RUBIS	10/20/63	Algeria	granite	52.0	5.48	0.04
PILEDRIVER	6/2/66	Climax Stoek	granite	56.0	5.52	0.03
LONG SHOT	10/29/65	Amehitka	andesite	80.0	5.55	0.03
SAPHIR	2/27/65	Algeria	granite	120.0	5.78	0.04
SCOTCH	5/23/67	Pahute Mesa	tuff	155.0	5.88	0.03
BILBY	9/13/63	Yueea Flat	tuff	235.0	5.95	0.03
MILROW	10/2/69	Amehitka	lava	1000.0	6.55	0.04
BENHAM	12/19/68	Pahute Mesa	tuff	1150.0	6.60	0.03
HANDLEY	3/26/70	Pahute Mesa	tuff	1200.0	6.74	0.03
BOXCAR	4/26/68	Pahute Mesa	rhyolite	1300.0	6.63	0.05

testing the TTBT compliance of the Shagan River events was relative to NTS.

In order to perform the simplest regression which could be carried out using these data, one would assume that all the yields are known exactly (*i.e.*, that $\tau_j^2 = 0$ for all j), and then one would perform a weighted least-squares fit with all the weights σ_j^2 equal to one another. As can be seen from equations (35) - (38), the estimated values of α and β are independent of the value chosen for σ_j^2, so long as σ_j^2 is the same for each event. As equation (39) shows, however, the variances and covariance of these variables are a function of the constant σ_j^2, so we cannot choose this constant arbitrarily. The straightforward approach would be to choose some value for σ_j^2, compute estimates of α and β using equations (37) and (38), and then compute the magnitude residual $R_j = m_j - \alpha - \beta W_j$ for each event. The common magnitude variance for all events would then be given by

$$\sigma_j^2 = \sigma^2 = \frac{1}{n-2} \sum_{j=1}^{n} (m_j - \alpha - \beta W_j)^2 \qquad (49)$$

and this value would be used in equation (39) to compute σ_α^2, σ_β^2, and $\sigma_{\alpha\beta}$ for the least-squares estimates. Instead of this procedure, however, we shall simply use the values for σ_j which were given by Blandford *et al.* (1983) and which are listed in Table IX. Note that these values are not exactly the same for every event.

When we apply the least-squares technique of equations (37) - (39), we obtain $\alpha_0 = 3.864$, $\beta_0 = 0.899$, $\sigma_\alpha = 0.031$, $\sigma_\beta = 0.013$, and $\sigma_{\alpha\beta} = -0.000393$. We have used these last 3 values in Table I. In order to apply the maximum-likelihood technique of equations (46) - (48), we must assign a standard deviation τ_j to the value of the log yield of each calibration event, which is now taken to be a quantity subject to error. We shall assume that the yields Y_j are known to within some precision ΔY_j, and we set

$$\tau_j \approx \log\left(\frac{Y_j + \Delta Y_j}{Y_j}\right) \approx \log\left(\frac{Y_j}{Y_j - \Delta Y_j}\right) \qquad (50)$$

If the fractional precision of the measured yield $\Delta Y_j / Y_j$ is suitably small, we can approximate the distribution of the log yield W_j as being normal with standard deviation τ_j. We shall arbitrarily assume that the yields are known to within a fractional precision $\Delta Y_j / Y_j$ of 10%, and we shall substitute this value into equation (89) in order to find a common value of τ_j which will be used for every event in the maximum-likelihood regression. This procedure results in values of α_0 = 3.882, β_0 = 0.890, and σ_c = 0.07. These values were used in Table I. (Recall that the value of σ_j in Table I is given by the square root of the sum of the variances $\sigma_j^2 + \sigma_m^2$, where the magnitude standard deviation σ_m was taken to be 0.02 for every event.) We have thus mixed the values of the slope and intercept of the maximum-likelihood regression with the variance-covariance structure of those values which were obtained using the least-squares regression. We have done this because we have not extended the maximum-likelihood approach to evaluate σ_α, σ_β and $\sigma_{\alpha\beta}$. We believe that only a small error is introduced by using the least-squares estimates of these parameters, however, since the values of α_0 and β_0 which were obtained by both techniques are so nearly the same.

CONCLUSIONS

We have applied a weighted least-squares regression and an iterative least-squares regression to obtain estimates of the parameters which relate seismic magnitudes to yields. The results of these 2 analyses were quite similar. The least-squares regression permitted the determination of the variance-covariance structure of the slope and intercept of the magnitude-yield relation, and the maximum-likelihood regression, which treats the calibration yields as uncertain values, permitted the calculation of the magnitude scatter due to source coupling. The data which were used to calibrate the magnitude-yield relation were not uniform with respect to test site and source medium, so the values which were determined here ought to be used only for illustrating the TTBT compliance testing procedures and not for monitoring the TTBT in practice.

The parameters which resulted from the regression analyses were applied first to the computation of confidence limits surrounding estimates of the yields for 22 Shagan River explosions and then to a test of the hypothesis that all 22 explosions were in compliance with the TTBT limit of 150 KT. The assigning of confidence limits to the yields is effective for considering each one of the 22 events separately, but it cannot be used to consider all the events collectively since the confidence limits are correlated among events. This correlation comes about from the uncertainty in the parameters which are used for the magnitude-yield relation and especially from the uncertainty in the size of the USSR - US magnitude bias. It was demonstrated that testing the null hypothesis that all 22 events were in compliance is an effective way of monitoring the TTBT. Tests were performed which rejected the null hypothesis of compliance if one or more, or 2 or more, estimated log yields exceeded some threshold T. Neither of these tests is optimum, but it was shown that the optimum test is impractical because it is susceptible to failure if a series of low-yield explosions are included among the events being monitored.

It was shown that on account of the uncertainties in the seismic magnitudes, in the magnitude-yield relation, and in the USSR - US magnitude bias, it is necessary to set the threshold T for rejecting the null (compliance) hypothesis at a level which far exceeds the TTBT limit of 150 KT in order to be able to perform the test with a suitably low probability of issuing a false alarm of TTBT violation. The power of one particular hypothesis test was measured by using a distribution of yields which had a single violation of the 150 KT limit and computing the probability that the null hypothesis would be rejected for the estimates of those yields. This power testing procedure can also be applied to the other hypothesis tests and to other yield distributions containing violations. An important step in designing a hypothesis test is the assumption of the distribution of yields, all in TTBT compliance, which is to be used as the null hypothesis. Two methods for creating this yield distribution were used. One was the assumption that all 22 events had yields of 150 KT, and the other was the creation of an *ad hoc* yield distribution using the observed magnitudes. The latter method, while not strictly valid from a theoretical viewpoint, may be more useful in practice.

ACKNOWLEDGEMENTS

The portions of this work devoted to TTBT compliance testing and magnitude - yield regression were performed in conjunction with meetings of a Statistical Working Group organized by DARPA to contribute to a report on TTBT verification. We appreciate the contributions made by that panel to this study. The other members of the Working Group were R. Alewine, R. Blandford, T. Eisenhauer, H. Gray, J. Hannon, E. Herrin, T. Jordan, G. Leies, G. McCartor, J. Murphy, W. Nicholson, and D. Westervelt.

REFERENCES

Abramowitz, M., and I. A. Stegun (1964). *Handbook of Mathematical Functions*, National Bureau of Standards, Washington, D. C.

Bache, T. C. (1982). Estimating the yield of underground nuclear explosions, *Bull. Seism. Soc. Am.*, *72*, S131 - S168.

Bache, T. C., W. J. Best, R. R. Blandford, G. V. Bulin, D. G. Harkrider, E. J. Herrin, A. Ryall, and M. J. Shore (1981). A technical assessment of seismic yield estimation, *DARPA-NMR-81-02*, Defense Advanced Research Projects Agency, Arlington, Virginia.

Blandford, R. R., R. H. Shumway, and R. Wagner (1983). Magnitude - yield for nuclear explosions at several test sites with allowance for effects of truncated data, amplitude correlation between events within test sites, absorption, and pP, *VSC-TR-83-15 (SDAC-TR-83-6)*, Teledyne Geotech, Alexandria, Virginia.

Johnson, N. L., and S. Katz (1972). *Continuous Multivariate Distributions*, John Wiley and Sons, New York.

Sykes, L. R., and I. L. Cifuentes (1984). Yields of Soviet underground nuclear explosions from seismic surface waves : Compliance with the Threshold Test Ban Treaty, *Proc. Natl. Acad. Sci. USA*, *81*, 1922 - 1925.