ATOMIC WEAPONS RESEARCH ESTABLISHMENT

AWRE REPORT NO. O22/86

Some Simulation Studies on Seismic Magnitude Estimators

R C Lilwall

Recommended for issue by

A Douglas, Superintendent

Approved by

B L Elphick, Head of Division

ISBN 0 85518176 1

CONTENTS

SUMMARY

Conventional seismic magnitude estimates are frequently biased because of the absence of station readings resulting from poor signal-to-noise ratio on the seismograms. Several magnitude estimation techniques, designed to overcome this threshold bias problem, have been published and in this report some of their properties are described using simulation experiments. All these alternative estimators successfully eliminate this bias, except at magnitudes below the detection thresholds of the more sensitive stations where deficiencies are apparent either in terms of bias or large variance. The presence of bias in the joint least squares analysis method (LSMF), where both magnitudes and station terms for several sources are computed simultaneously, is also investigated. A marginal reduction in magnitude bias is found but bias is present in the estimated station terms. A suggested joint "maximum likelihood" approach provides an attractive alternative.

1. INTRODUCTION

Although the magnitude of a seismic disturbance can be estimated using observations from a single station, the presence of scatter in the observations makes it desirable to use data from many stations. It is well known however that simple averaging of the results from a network of stations, as used by the International Seismological Centre (ISC) and in the United States Geological Survey earthquake data reports often results in biased estimates. An important cause of bias is data censoring, arising at higher magnitudes from saturated or clipped recordings, but more frequently at lower magnitudes because of detection or amplitude reporting thresholds. The magnitude at which the latter becomes significant can be reduced by using a network of stations all with low detection thresholds, but with the existing world network, alternative analysis techniques need to be employed, even at moderate magnitudes. Several statistical procedures, usually referred to as "maximum likelihood methods", have been published and recently applied to data in international bulletins (Ringdal (1), Lilwall (2)). Anticipating more routine use of these methods, especially to the existing world network, it is desirable to make a comparative study to find the merits of each. Here a simple simulation study is described assuming the type of data at present available through international agencies. Another technique used for magnitude estimation is the "joint method" where magnitudes and other parameters such as station or distance corrections are computed for a group of sources. This is particularly useful for the estimation of relative magnitudes of nuclear explosions within a test site but the extent of any bias resulting from censoring in either the estimated magnitudes or other parameters is not obvious. In the second part of this report a simulation experiment is described which investigates the problem of bias when applying an existing joint least squares and a possible alternative joint maximum likelihood technique to groups of underground nuclear explosions.

2. SINGLE MAGNITUDE ESTIMATORS

2.1 Theoretical outlines

Suppose a seismic disturbance with "true" magnitude m_t occurs within a network of stations. Furthermore let N_s stations be at distances suitable for the estimation of m_t. If the ground amplitude (in terms of Log^A/T) at the ith station is a_i then the station magnitude can be defined as:

$$m_i = a_i + B_i(\Delta,h) \qquad \ldots (1)$$

where B_i is the distance (Δ)-depth(h) correction. If S_i is the station amplitude term then the true magnitude is given by:

$$m_t = m_i - S_i + \epsilon_i \qquad \ldots (2)$$

where ϵ_i is a random variable which, following Freedman (3) is usually assumed to be normally distributed with variance σ_i^2. The probability density function for m_i is therefore

$$P(m_i|m_t) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-0.5\left[\frac{m_i-S_i-m_t}{\sigma_i}\right]^2} \qquad \ldots (3)$$

The conventional magnitude estimator is the mean of the <u>observed</u> m_i.

$$\bar{m} = \sum_{i=1}^{i=N_s} (m_i-S_i)/ND \quad i \epsilon D \qquad \ldots (4)$$

where ND is the number of observed m_i and D is the corresponding set of indices. Clearly $\bar{m}$ should be unbiased if the m_i is sampled randomly from the normal population (equation 3) but it has been appreciated for some time (Herrin and Tucker (4), Evernden and Kohler (5), Ringdal (6), Christoffersson et al (7)) that loss or censoring of lower values of m_i resulting from the presence of station reporting thresholds give rise to a positive bias in $\bar{m}$. Alternative estimators have been published however which are designed to allow for this data loss (Christoffersson et al (7), Ringdal (6)). These methods modify the density function (equation 3) and can also use the information implicit in the stations not able to measure m_i. What follows summarises the theory.

Suppose for the ith station we have the following additional information concerning the thresholds for measuring a_i and hence m_i. Let

g_i = mean (50%) threshold for measuring an amplitude (a_i)

G_i = $B_i(\Delta,h)+g_i$ be the corresponding threshold in terms of magnitude for a given source. ... (5)

γ_i^2 = variance of actual threshold, assumed to be a random normally distributed variable with mean g_i (or G_i).

The threshold g_i need not be the detection threshold and will generally be higher since the station analyst will frequently only measure amplitudes when there is a good signal to noise ratio.

The data from the N_s stations are split into two sets of observations; "m_i not measured, $i \notin D$" and "m_i measured, $i \in D$". The conditional probabilities of each observation as given by Christoffersson et al (7) are:

$$P(m_i|m_t) = \begin{cases} \Phi\left[\dfrac{m_i-G_i}{\gamma_i}\right] \theta\left[\dfrac{m_i-S_i-m_t}{\sigma_i}\right] & i \in D \quad \ldots (6A) \\[2ex] \Phi\left[\dfrac{G_i-S_i-m_t}{\sqrt{\sigma_i^2+\gamma_i^2}}\right] & i \notin D \quad \ldots (6B) \end{cases}$$

where θ is the normal density function given by equation 3 and Φ the cumulative normal distribution given by:

$$\Phi(y) = \int_{-\infty}^{y} \theta(x)\,dx \qquad \ldots (7)$$

The likelihood function for the set of observations depends on whether we wish to include "m_i not measured" observations. If not, then the likelihood is the product of (ND) terms such as equation (6A).

$$L(m_t) = \prod_{\substack{i=1 \\ i \in D}}^{N_s} \frac{\Phi\left[\dfrac{m_i-G_i}{\gamma_i}\right] \theta\left[\dfrac{m_i-S_i-m_t}{\sigma_i}\right]}{\left[\dfrac{S_i+m_t-G_i}{\sqrt{\sigma_i^2+\gamma_i^2}}\right]} \qquad \ldots (8)$$

The bottom line is required to make the probabilities conditional to each station measuring m_i. (See Christoffersson et al (7), Christoffersson and Ringdal (8)). A point estimate $\bar{m}$ of m_t can be found by maximising this function for the set of observed m_i and assumed values for G_i, S_i, σ_i and γ_i. This estimator appears little used, but since it requires the observed measurements only, it is more strictly comparable to the conventional mean estimator given by equation 4 than the following alternatives.

If the information present in the "m_i not measured" observations is to be used then the likelihood function is the product of terms as in 6A and 6B where appropriate. The probability must in this case be made conditional that there is at least one observed m_i. The probability of at least one observation is:

$$P_1(\text{NOBS } m_i \geqslant 1) \quad 1.0 - \prod_{i=1}^{N_s} \Phi \left[\frac{G_i - S_i - m_t}{\sqrt{\sigma_i^2 + \gamma^2}} \right] \qquad \ldots (9)$$

Hence the likelihood function for the full set of N_s observations is:

$$L(m_t) = \frac{1}{P_1} \prod_{\substack{i=1 \\ i \in D}}^{N_s} \Phi \left[\frac{m_i - G_i}{\gamma_i} \right] \theta \left[\frac{m_i - S_i - m_t}{\sigma_i} \right] \prod_{\substack{i=1 \\ i \in D}}^{i=N_s} \Phi \left[\frac{G_i - S_i - m_t}{\sqrt{\sigma_i^2 + \gamma_i^2}} \right] \qquad \ldots (10)$$

When maximising this function with respect to m_t clearly the first term $\Phi\ ((m_i - G_i)/\gamma_i)$ can be omitted.

The last estimator to be considered here is that described by Ringdal (6). His likelihood function is essentially

$$L(m_t) = \prod_{\substack{i=1 \\ i \in D}}^{N_s} \theta \left[\frac{m_i - S_i - m_t}{\sigma_i} \right] \prod_{\substack{i=1 \\ i \in D}}^{N_s} \Phi \left[\frac{G_i - S_i - m_t}{\sqrt{\sigma_i^2 + \gamma_i^2}} \right] \qquad \ldots (11)$$

This is essentially the same as equation (10) but omitting the factor $1/P_1$ conditioning the probability to at least one measured m_i being present.

The likelihood equations 8, 10 and 11 represent alternative estimation procedures to using the simple mean (equation 4). The use of the extra information contained in the "non observed m_i" would appear to give likelihood equations 10 and 11 an advantage provided that such negative observations are definitely the result of the station thresholds (G_i). There are many other reasons for such a result however such as station "down time", lost/spoilt recordings, or clipping with larger magnitudes. Unless detailed station operation records are available these extra causes of non observation must be allowed for in the likelihood equations. For bulletin data such modifications are important and can be attempted as in Ringdal (1) and Lilwall (2). Equation (8) is clearly free from this problem and appears to be an attractive alternative especially for routine determinations using bulletin data. The optimum choice also depends on the relative performance of the estimators and it is hoped to clarify this here using the following comparative study using simulated data.

2.2 Data simulation study

The object is to compare the four magnitude estimators described in the previous section. The method used is to generate simulated sets of observations which obey the underlying statistical model and apply each of the four estimators to the resulting data.

Two hypothetical networks of stations were used in the investigation. Network 1 is identical to that used in an original study by Ringdal (6) using the likelihood formulation in equation 11. This network consists of 10 stations each with thresholds (G_i) in steps of 0.1 units from 4.1 to 5.0 inclusive (see figure 1). Ringdal showed that his method was effective in removing the bias present when using the standard mean estimator for simulated true magnitudes down to 4.0. He did not investigate lower magnitudes but no problem resulting from the omission of the factor $1.0/P_1$ present in equation 10 was encountered. Network 2 is intended to model a typical network of stations submitting data to the International Seismological Centre (ISC) during the period 1978-81. Values of G_i for this period are computed from the thresholds g_i published in a previous report (2) using equation (5) and $B(\Delta,h)$ corrections published by Marshall, Bingham and Young (9) for a surface focus source in the Kuriles. The distribution of G_i is also given in figure 1.

Station magnitude data m_i were generated as follows for a simulated true magnitude m_t:

(1) Generate a set of N_s station magnitudes m_i by adding normally distributed random numbers with zero mean, variance σ_i^2 to m_t.

(2) Generate a set of N_s station reporting thresholds by adding normally distributed random numbers with zero mean, variance γ_i^2 to the mean thresholds G_i.

(3) For each station, signal a measured station magnitude m_i if it is above the threshold, otherwise signal a "non observation".

(4) Using the four estimation procedures under examination compute the magnitude using the observed m_i and non observations where appropriate.

(5) Repeat a large number of times to obtain the distribution of the estimates for each m_t.

Values of γ_i and σ_i were fixed at 0.2 and 0.35 respectively, values typical of those found by several workers using Bulletin data and widely distributed seismic sources (eg, North (10), Ringdal (1), Lilwall (2)). Station terms were set to zero. The normal distributions were truncated at ±4 standard deviations, partly for computational convenience, but also because there is no evidence that the true distributions are unbounded.

For each true magnitude the simulations were repeated until 500 sets of estimates were obtained. For lower magnitudes several times this number were necessary because of the large number which had no detections.

2.3 Discussion of the simulation results

Figure 2 shows a summary of the results for network 1. They are presented in terms of the bias as a function of simulated true magnitude. Bias here is the difference of the mean estimated value, using each set of the 500 observations, from the "true" value. It is also given in terms of the median value. The standard deviation of the estimates about the mean is used as a measure of the variance.

Figure 2A illustrates the usual bias present when the straight mean is used. This bias increases monotonically with reducing magnitude and is significant (0.1 units) even at magnitudes corresponding to the least sensitive station. Figure 2B shows the results of applying the maximum likelihood equation (8) again using only the observed values of m_i. Clearly this estimator is relatively free from bias over the whole magnitude range but also has the highest variance. For magnitudes at or below the average station threshold the variance rapidly becomes large, and exceeds the value which would be obtained from single observations sampled randomly from the underlying normal population (equation 3 with $\sigma = 0.35$). Figures 2C and 2D illustrate the results obtained using equations 10 and 11 which include the information from the "non-observations". In figure 2C the results obtained by Ringdal (6) are essentially repeated for magnitudes down to 4.0. Below this level, which corresponds approximately to the threshold of the most sensitive station, the estimates become progressively positively biased. This is the result of the approximation in equation 11 compared with 10 and accounts for the apparent reduction in the difference between magnitudes determined by the maximum likelihood and mean estimators observed by Ringdal (1). When all the terms are included (equation 10) the positive bias at low magnitudes is removed. It is replaced by a much smaller negative value resulting partly from the truncation of the simulated data to ±4 standard deviations but also apparently to an inherent property of this estimator as implemented. When measured in terms of the median this negative bias is reduced.

An interesting observation is that although the estimates in figures 2A and 2C become positively biased at low magnitudes the variance does not increase but may actually decrease. This is a direct consequence of the reduction in the variance of the observed station magnitudes given by equation 8 compared with the underlying distribution, equation 3. This is illustrated in figure 3 which shows that even for magnitudes at the 50% detection level the standard deviation of the observed m_i falls to 0.75 of the true value.

Results for the larger Network 2 (figure 4) are similar except that the addition of a large number of less sensitive stations increases the bias present using the mean estimator and extends it to much higher magnitudes. The reduction in the standard errors at high magnitudes merely reflects the larger amount of data contributing, but little difference is apparent at the lower end where few of the stations contribute.

3. JOINT MAGNITUDE ESTIMATION

3.1 Introduction

Magnitude determination for several seismic disturbances together permits the estimation of additional parameter values such as station terms and amplitude distance terms. In techniques such as used by North (10) the problem is dealt with piecewise in that the magnitudes are first determined individually and then station terms estimated from the resulting station magnitude residuals. Methods involving the simultaneous estimation of all the unknowns as described in Douglas (11) are more rigorous and do not require iteration. This least squares technique has been used in several studies (Carpenter et al (12), Booth et al (13), Marshall et al (9)) and is routinely used in the determination of the magnitudes of explosions in nuclear test sites (Marshall et al (14,15)). These methods do not address the problem of bias resulting from data censoring and it is inevitable that bias is present in the resulting estimates. It is not obvious however how the bias is partitioned between the magnitudes and other estimated parameters. This problem is investigated here by again simulating data and applying both the joint least squares and joint maximum likelihood algorithms.

3.2 Theoretical outline of joint methods

Consider N_e seismic disturbances and an observational network of N_s stations. Let m_{tj} be the true magnitude of the jth event and m_{ij} be the station magnitude (as defined by equation 2) of the jth event at the ith station. Equation (2) can be generalised:

$$m_{tj} = m_{ij} - S_i + \epsilon_{ij} \qquad \ldots (12)$$

From the set of observed m_{ij}, least squares can be used to estimate the S_i and m_{tj} provided an additional constraint

$$\sum_{i=1}^{i=N_s} S_i = 0.0 \qquad \ldots (13)$$

is used. This is essentially the "Least Squares by Matrix Factorisation" technique (LSMF) described by Douglas (11) and has been generalised further to include the estimation of amplitude-distance terms. The form given above however is the most appropriate for closely spaced sources where perturbations from the assumed amplitude-distance curve can be described by a single term S_i for each station.

The LSMF method takes no account of data censoring and essentially is similar to use of the straight mean in conventional magnitude estimates. All three of the maximum likelihood equations 8, 10 or 11 can be generalised to the joint estimation of magnitudes and station terms. To compare directly with LSMF it is necessary to use the observed station magnitudes only and hence generalise equation 8. The likelihood is therefore

$$L(m_{tj}, S_i) = \prod_{\substack{j=1 \\ i=1 \\ ij\epsilon D}}^{\substack{N_s \\ N_e}} \frac{\Phi\left[\dfrac{m_{ij}-G_i}{\gamma_i}\right] \theta \left[\dfrac{m_{ij}-S_i-m_{tj}}{\sigma_i}\right]}{\Phi\left[\dfrac{S_i+m_{tj}-G_i}{\sqrt{\gamma_i^2+\sigma_i^2}}\right]} \qquad \ldots (14)$$

Where the products are made only for the combinations ij corresponding to the observed m_{ij} set D. Point estimates of m_{tj}, S_i and σ_i can be found by maximising this function using known values for G_i and γ_i. Again the additional constraint given by equation 13 is required. The maximisation subject to variation of all the σ_i gives rise to numerical problems so they are all assumed to be a single variable σ. In practice the value of σ and approximate values form m_{tj} and S_i can be found using a piecewise iterative scheme maximising the likelihood for the magnitudes and then the stations separately and examining the overall variation with σ. Starting with the values of m_{tj} and S_i so obtained a final maximisation can be made by Newton-Raphson iteration. The constraint equation 13 can be applied using the method of Lagrange Multipliers (eg, Edwards (16), Aitchison and Silvey (17)). The stable determination of σ is possible because the data from a large number of observations are pooled when using joint techniques. This is a distinct advantage over single determinations where it is usually assumed (eg, 0.3 to 0.4 for earthquakes). Error in the assumed value results in bias even when using the maximum likelihood estimators.

Confidence limits on the point estimates for LSMF can be obtained from the inverted normal equation using the standard least squares theory as described by Douglas (11). For the maximum likelihood method they can be obtained by exploring the variation of the likelihood around its maximum. Approximate confidence limits can be found easily however from the results of the Newton-Raphson method which requires the inverted matrix of second derivatives of the likelihood function. This matrix approximates the variance matrix for the distributions of the estimates (eg, Edwards (16)).

3.3 Data simulation study of joint methods

The simulations described here are not intended to be a comprehensive study on the effect of station thresholds on all the various possibilities for joint estimation schemes. Results will be dependent not only on the relative distribution of magnitudes and station thresholds within the group but also on the nature and spatial distribution of the sources which influences the value of the variance σ^2. Instead this study concentrates on the use of such methods as applied to groups of sources from underground nuclear explosion test sites.

The relatively uniform source radiation and close spacing of the epicentres within a test site make LSMF a powerful technique for determining the magnitudes. The application of the least squares method to real data (Marshall et al (14,15)) indicates that the station magnitude variance σ^2 is smaller than found for widely distributed earthquake data.

Typically σ is found to be between 0.1 to 0.2 for bulletin data but in the presence of threshold effects, these will be underestimated, a value of 0.2 appears appropriate. This relatively small value for σ should minimise the effect of station thresholds on LSMF but since the accurate determination of magnitudes is crucial to yield estimation it is interesting to base the simulation study on this problem.

Figure 5 shows the distribution of magnitudes of presumed explosions in the two adjacent Soviet test sites at Degelen Mt (Site A) and Shagan River (Site B) as determined using LSMF by Marshall et al (14). The distributions differ in that the Shagan magnitudes tend to be larger. Also shown is the distribution of station thresholds G_i for a hypothetical network of stations. This network is based on the real world network submitting amplitude data to the ISC during 1978-81 for which amplitude (Log^A/T) thresholds g_i have been computed by Lilwall (2). Only stations in the distance range Δ = 30 to 90^o are included and Soviet stations deleted as they do not report amplitude data for these sources. The magnitude thresholds in figure 5 are obtained by adding distance factors B_i as published by Marshall et al (9). Simulated data sets were produced in a similar way to that described in section 2 assuming a distribution of true magnitudes and hypothetical network as shown in figure 5. Both σ_i and γ_i were set to 0.2. Station terms derived from a normal distribution with zero mean and SD = 0.2, were included in the generation of the data.

The least squares (LSMF) and joint maximum likelihood (JML) techniques described in the previous section were applied to the two sets of data. In addition the magnitudes were determined singly using the means. Figure 6 presents the results for the magnitudes in terms of the average difference between the estimated and "true" value as a function of the true value. Clearly even with a relatively low value for σ the bias problem using the mean is still substantial especially for site A (Degelen simulation). Use of LSMF gives only moderate reductions in the bias (triangles) whilst as expected the JML results (dots) are unbiased. The effect of bias on the validity of the confidence limits is striking. For LSMF the standard limits include the true value in 3% and 29% of cases for sites A and B respectively whilst these percentages rise to 62% and 68% for JML. It is evident that the use of JML if possible may provide a useful improvement in routine analysis of real data.

Figure 7 illustrates the deviation of the estimated station terms for site A from the simulated values as a function of the station threshold G_i. Again the JML results are as expected with zero bias and increasing variance with increasing threshold. A clear trend is present in the LSMF results with a negative bias at low thresholds changing to a positive value at high thresholds. The total range of the bias is considerable (0.2 to 0.3 units). This trend is clearly the result of the threshold censoring of low readings at the less sensitive stations but affects all the estimations through the constraint that the total sum of station terms is zero.

The results for site B are shown in figure 8. Again the LSMF values have a clear but much reduced trend. It is however surprisingly large considering that 80% of the magnitudes are 5.5 and above. The validity of confidence bounds for the station terms for both sites parallels those for magnitudes, for LSMF the percentages for inclusion of the true value are 3% and 34% whilst for JML they are 68% and 58% for A and B respectively.

4. CONCLUSIONS

This report examines the performance of several statistical procedures for magnitude estimation in the presence of data censorship resulting from station thresholds. For magnitudes estimated singly, all three estimators studied show little bias for magnitudes above the 50% detection threshold of the most sensitive station in the observing network. If the origin of the observations giving no amplitude measurement can be correctly identified then the two methods which use this information (equation 10 and 11) have a clear advantage in terms of reduced variance. For routine magnitude determinations from bulletin data however equation 8 gives similarly unbiased results and is clearly preferable to the conventional mean. At lower true magnitudes there appears to be a trade off between bias and variance. Under an unlimited normal law for the station magnitude and noise distributions the large variance found for the unbiased estimators shows the information content in magnitude observations below the threshold is low. It is worth noting that if the performance of each estimator is assessed in terms of the rms deviation from the true magnitude Ringdal's formulation (equation 11) gives the best results for magnitudes down to 3.2 in spite of the bias. Associated confidence limits, if computed however, will be unrealistically low.

Joint least squares analysis, when applied to world network data such as received by the ISC, should give a moderate reduction in the magnitude bias present in standard determinations. Significant bias is still present however, both in the estimated magnitudes and station terms. It is smaller if larger (with respect to station thresholds) magnitudes only are included but this reduces the value of the technique when applied to many problems using Bulletin data. Care must therefore be taken when interpreting the station terms resulting from least squares analysis since bias contribution can be as great at ±0.1 units. Joint maximum likelihood is an attractive alternative to least squares providing reliable values for the thresholds are available. Finally it is worth restating that the effect of station thresholds will be greater for widely spaced earthquake sources which result in an increased variance σ. For such data use of maximum likelihood rather than least squares technique has even more advantages than for the analysis of closely spaced explosion sources.

REFERENCES

1. F Ringdal: "Study of Magnitudes, Seismicity and Earthquake Detectability Using a Global Network."
Centre for Seismic Studies Technical Report C84-01 (1984)

2. R C Lilwall: "Redetermination of Body-Wave Magnitudes (m_b) Using ISC Bulletin Data."
AWRE Report O21/85, HMSO, London (1985)

3. H Freedman: "Estimating Earthquake Magnitude."
Bull Seism Soc Am, 57, 747-760 (1967)

4. E Herrin and W Tucker: "On the Estimation of Body Wave Magnitudes."
Tech Rep to AFOSR (1972)

5. J F Evernden and W M Kohler: "Bias in Estimates of m_b at Small Magnitudes."
Bull Seism Soc Am, 66, 1887-1904 (1976)

6. F Ringdal: "Maximum Likelihood Estimation of Seismic Magnitude."
Bull Seism Soc Am, 66, 789-802 (1976)

7. L A Christoffersson, R T Lacoss and M A Chinnery: "Statistical Models for Magnitude Estimation."
Lincoln Lab SATS, TR-75-335, 2-5 (1975)

8. A Christoffersson and F Ringdal: "Optimum Approaches to Magnitude Measurements."
In: Identification of Seismic Sources - Earthquake or Underground Explosion. Proc of NATO Advanced Study Institute, edited by E S Husebye and S Mykkelveit, p 575-587. Reidel Publishing Co (1980)

9. P D Marshall, J Bingham and J B Young: "An Analysis of P-wave Amplitudes Recorded by Seismological Stations in the USSR."
Geophys J R astr Soc., 84, 71-91 (1986)

10. R G North: "Station Magnitude Bias - Its Determination, Causes and Effects."
Lincoln Lab Tech Note 1977-24 (1977)

11. A Douglas: "A Special Purpose Least Squares Programme."
AWRE Report O54/66, HMSO, London (1966)

12. E W Carpenter, P D Marshall and A Dougl s: "The Amplitude-Distrance Curve for Short Period Teleseismic P-waves."
Geophys J R astr Soc, 13, 61-70 (1967)

13. D C Booth, P D Marshall and J B Young: "Long and Short Period P-wave Amplitudes from Earthquakes in the Range 0 to 114^{o}."
Geophy J R astr Soc, 39, 523-537 (1974)

14. P D Marshall, T C Bache and R C Lilwall: "Body Wave Magnitudes and Locations of Soviet Underground Explosions at the Semipalatinsk Test Site."
AWRE Report O16/84, HMSO, London (1984)

15. P D Marshall, R C Lilwall and P J Warburton: "Body Wave Magnitudes and Locations of French Underground Explosions at Mururoa Test Site."
AWRE Report O12/85, HMSO, London (1985)

16. A W F Edwards: "Likelihood."
Cambridge University Press (1972)

17. J Aitchinson and S D Silvey: "Maximum Likelihood Estimation of Parameters subject to Restraints."
App Math Statist, 29, 813-828 (1958)

LIST OF FIGURES

FIGURE 1 FREQUENCY DISTRIBUTIONS OF THE 50% REPORTING THRESHOLDS FOR THE TWO NETWORKS USED IN THE SIMULATION STUDY. Network 1 is a simple hypothetical network as used by Ringdal (6). Network 2 simulates the world network reporting to the ISC during the period 1978-81 for a source in the Kuriles (Δ = 30 to 90).

FIGURE 2 SIMULATION RESULTS FOR NETWORK 1. Dots and crosses are the mean and median difference of each set of 500 determinations from each simulated true magnitude and thus gives the bias if present. Dotted lines represent ±1 SD about the mean. A and B are obtained when using the simple mean and equation 8 respectively and employ the observed station magnitudes (m_i) only. C and D include the "non observations" and likelihood equations 11 and 10 respectively.

FIGURE 3(a) PROBABILITY DENSITY FUNCTIONS (GIVEN BY EQUATION 8) FOR OBSERVED STATION MAGNITUDES AT A STATION FOR SOURCES OF TRUE MAGNITUDE m_t = 3.2, 3.6, 4.0 AND 4.4. Complete distribution (observed and non observed) of station magnitudes assumed normally distributed about m_t with σ = 0.3. Station detection threshold G_i = 4.0 with SD γ = 0.2. Notice that not only do the observed magnitudes become biased with lower true magnitude m_t but the SD of the observations is reduced.

(b) VARIATION IN SD OF OBSERVED m_i FOR VARIOUS TRUE σ AS FUNCTION OF TRUE MAGNITUDE m_t. Estimators of σ based on the observed magnitude population will tend to be low.

FIGURE 4 SIMULATION RESULTS FOR NETWORK 2. Dots and crosses are the mean and median difference of each set of 500 determinations from each simulated true magnitude and thus give the bias if present. Dotted lines represent ±1 SD about the mean. A and B are obtained when using the simple mean and equation 8 respectively and employ the measured station magnitudes (m_i) only. C and D include the "non observations" and likelihood equations 11 and 10 respectively.

FIGURE 5(a) DISTRIBUTION OF 50% REPORTING THRESHOLDS COMPUTED FOR THE SOVIET SEMIPALATINSK TEST SITE FOR STATIONS IN THE DISTANCE RANGE Δ = 30 TO 100^{o}. Network is based on stations reporting to the ISC during 1978-81 excluding those in the Soviet Union.

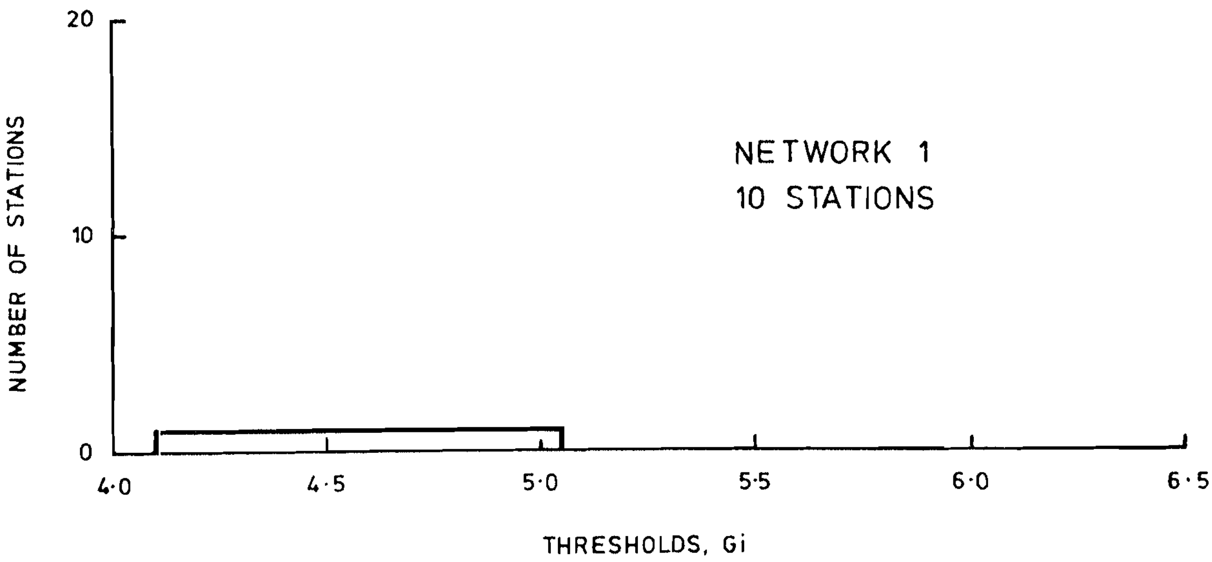
(b) DISTRIBUTION OF SIMULATED MAGNITUDES FOR SITE B BASED ON THOSE OBSERVED FOR THE SHAGAN RIVER REGION OF THE TEST SITE.

(c) DISTRIBUTION OF SIMULATED MAGNITUDES FOR SITE A BASED ON THOSE OBSERVED FOR DEGELEN MOUNTAIN REGION OF THE TEST SITE.

FIGURE 6 MEAN DIFFERENCES FROM THE (SIMULATED) TRUE MAGNITUDES FOR SITES A AND B AS A FUNCTION OF THE TRUE MAGNITUDE. Crosses, triangles and dots are results using the simple means, LSMF and joint maximum likelihood (JML) procedures respectively. LSMF gives a moderate reduction in bias whilst JML effectively removes all the bias.

FIGURE 7 ERRORS IN COMPUTED STATION TERMS FOR SITE A (SIMULATED DEGELEN MOUNTAIN) PLOTTED AS A FUNCTION OF STATION THRESHOLD. Errors using joint maximum liklihood are evenly distributed about zero whilst those using LSMF show a variable bias as a function of the threshold.

FIGURE 8 ERRORS IN COMPUTED STATION TERMS FOR SITE B (SIMULATED SHAGAN RIVER) PLOTTED AS A FUNCTION OF STATION THRESHOLD. As in figure 7 the joint maximum likelihood results show no bias whilst although smaller some bias for LSMF is still visible.



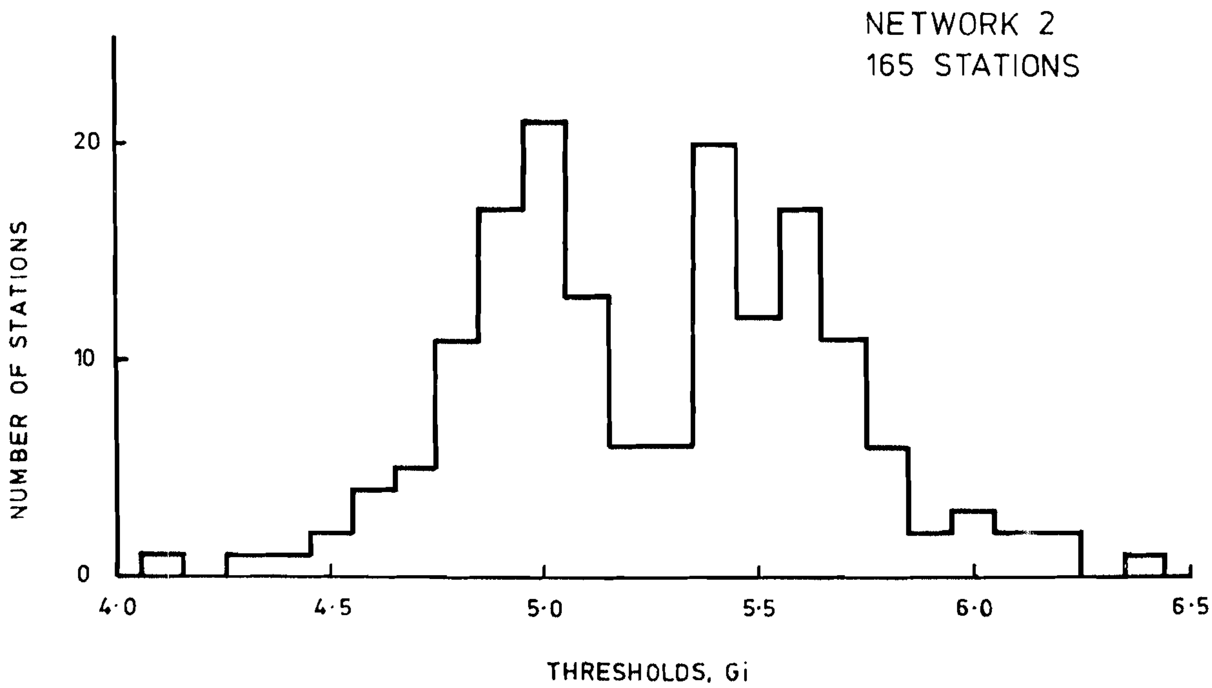


FIGURE 1. FREQUENCY DISTRIBUTIONS OF THE 50% REPORTING THRESHOLDS FOR THE TWO NETWORKS USED IN THE SIMULATION STUDY.

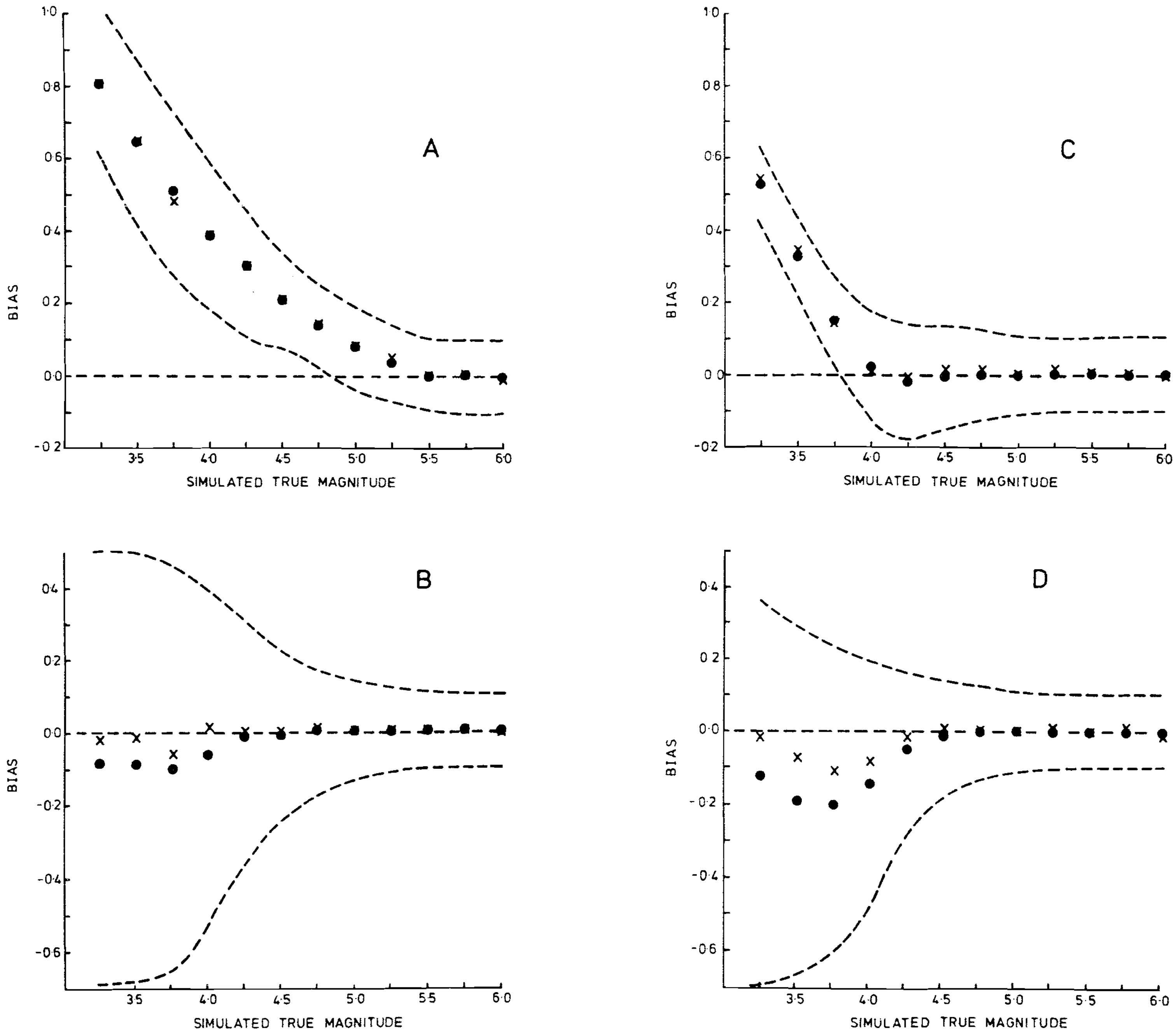


FIGURE 2. SIMULATION RESULTS FOR NETWORK 1.

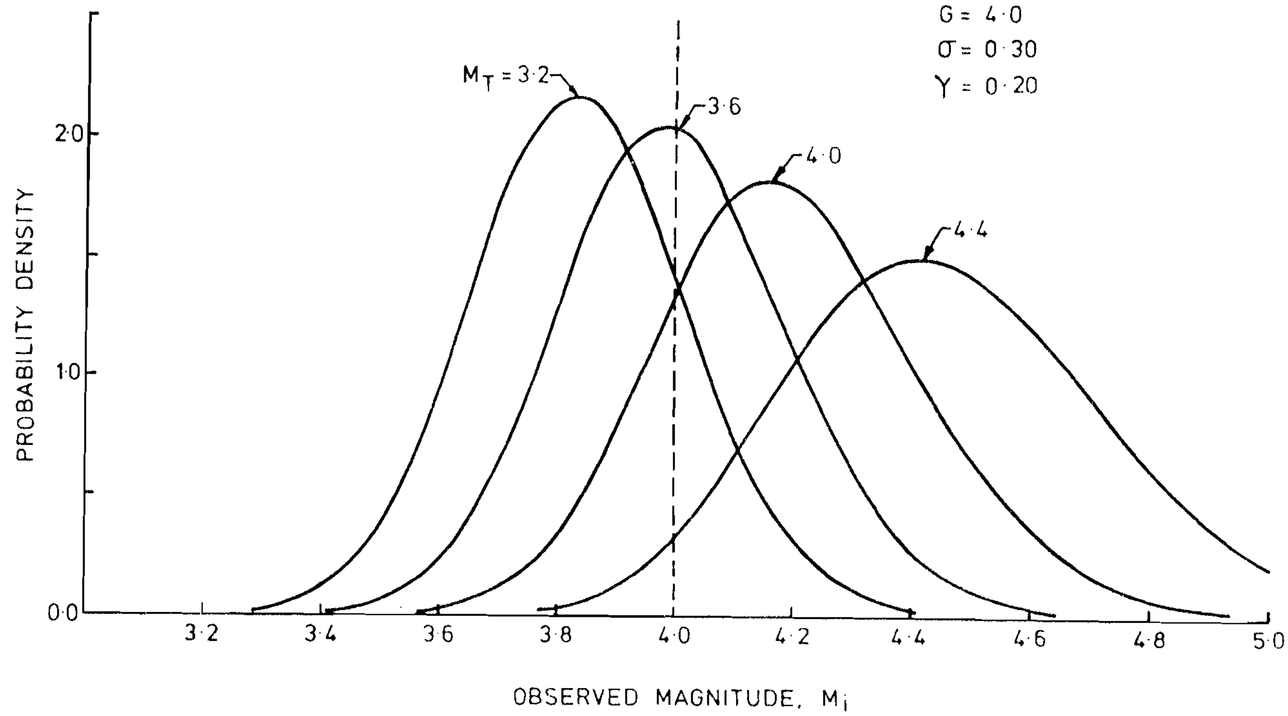


FIGURE 3(a). PROBABILITY DENSITY FUNCTIONS (GIVEN BY EQUATION 8) FOR OBSERVED STATION MAGNITUDES AT A STATION FOR SOURCES OF TRUE MAGNITUDE

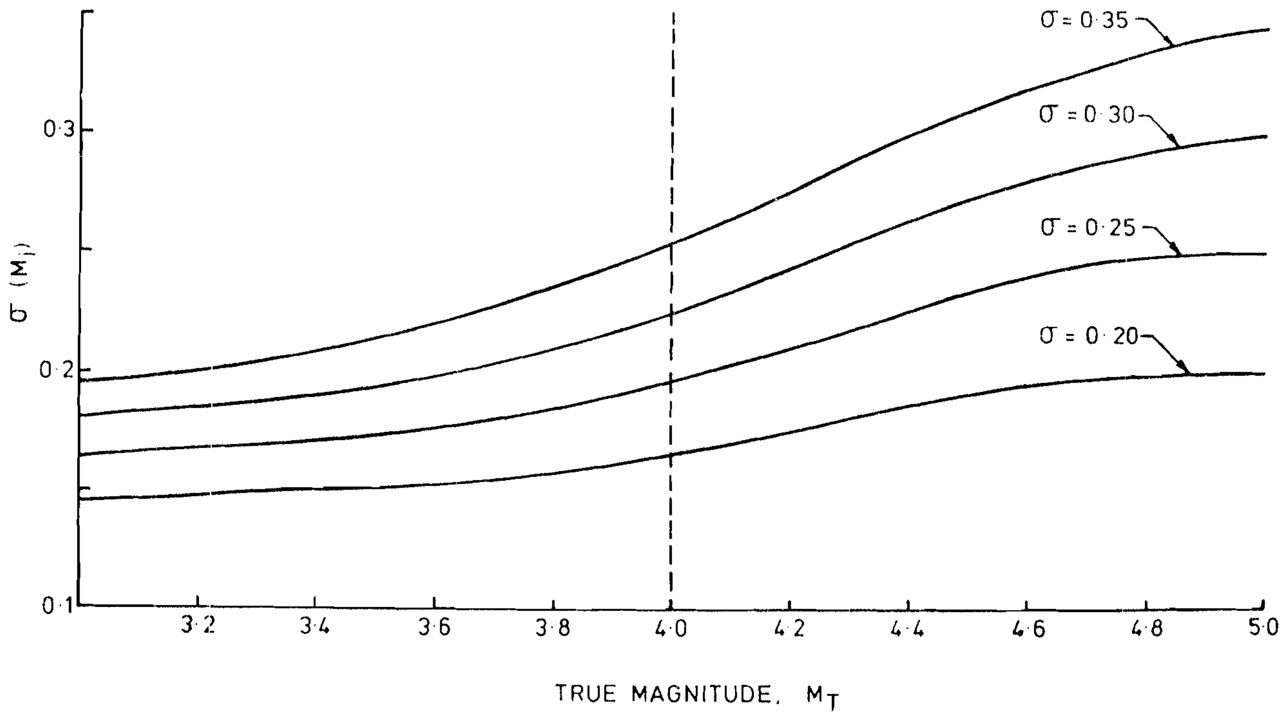


FIGURE 3(b). VARIATION IN SD OF OBSERVED m_i FOR VARIOUS TRUE σ AS FUNCTION OF TRUE MAGNITUDE m_t

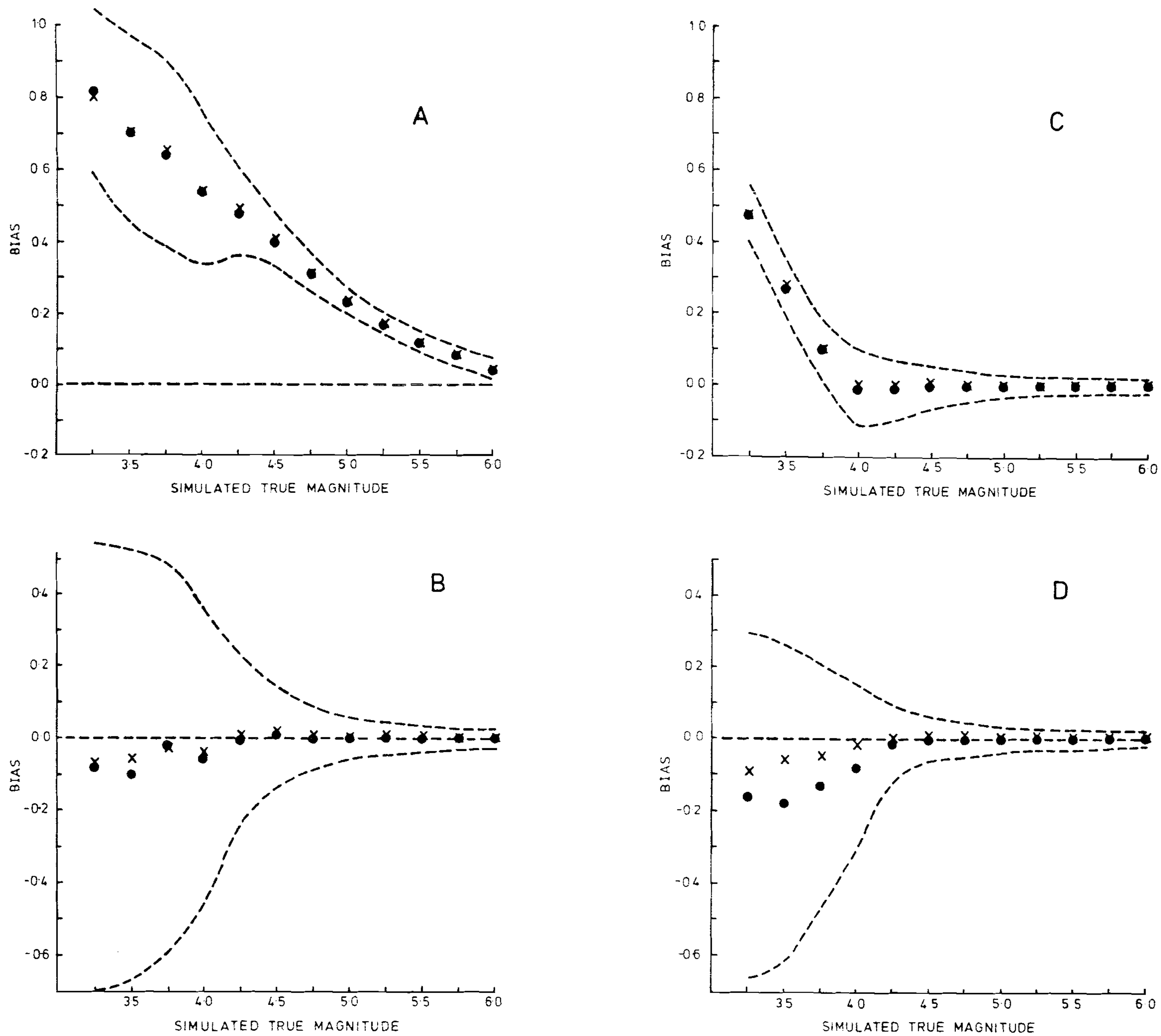


FIGURE 4. SIMULATION RESULTS FOR NETWORK 2.

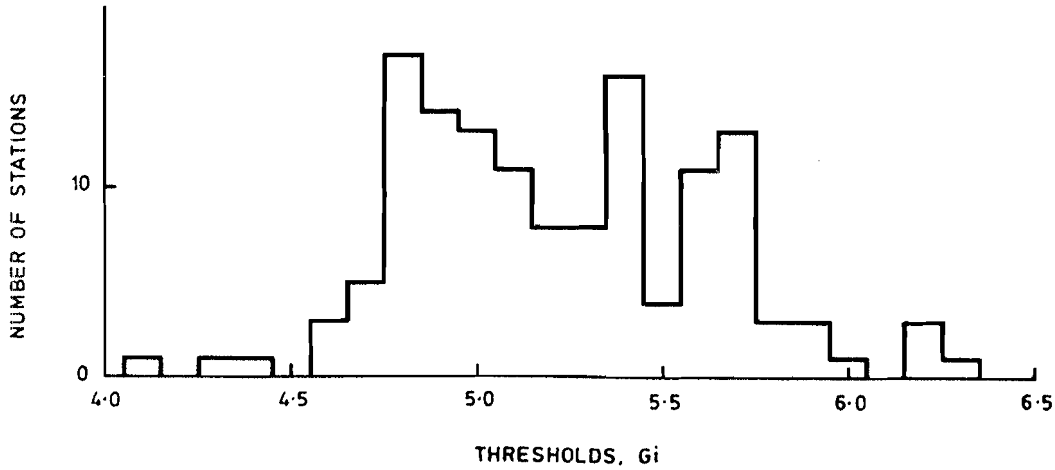


FIGURE 5(a). DISTRIBUTION OF 50% REPORTING THRESHOLDS COMPUTED FOR THE SOVIET SEMIPALATINSK TEST SITE FOR STATIONS IN THE DISTANCE RANGE Δ = 30 to 100°.

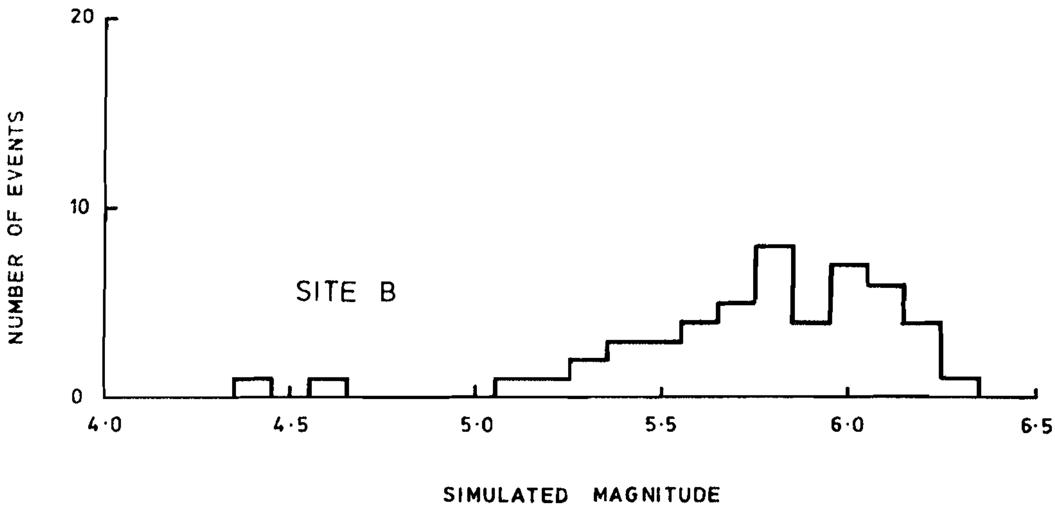


FIGURE 5(b). DISTRIBUTION OF SIMULATED MAGNITUDES FOR SITE B BASED ON THOSE OBSERVED FOR THE SHAGAN RIVER REGION OF THE TEST SITE.

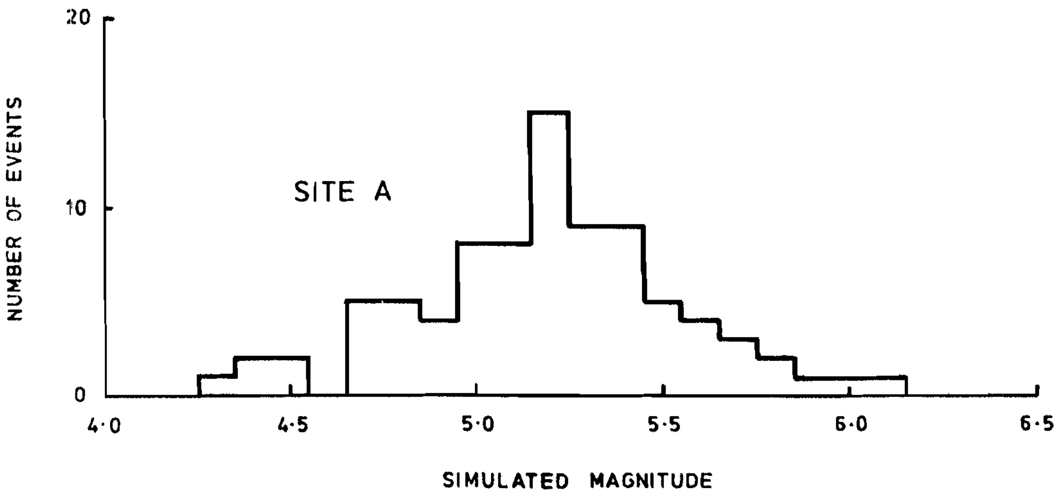
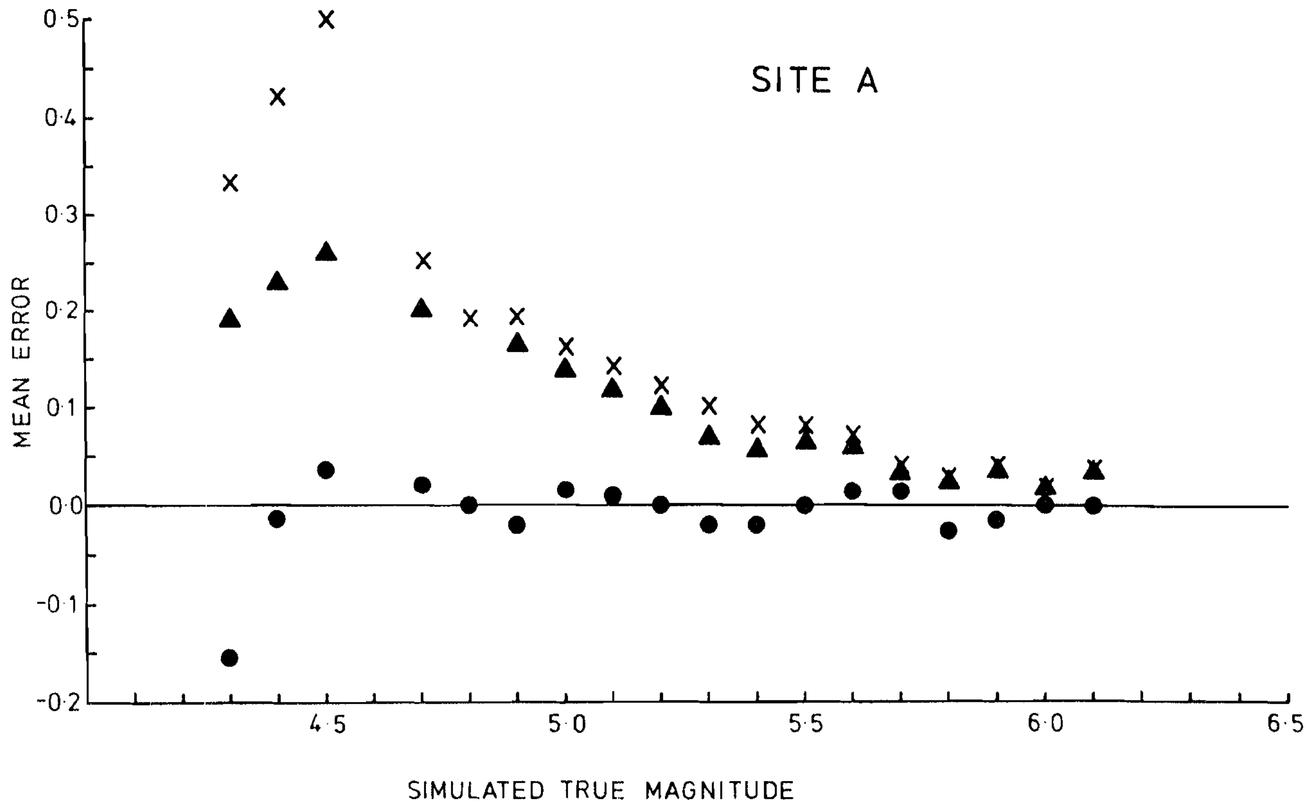


FIGURE 5(c). DISTRIBUTION OF SIMULATED MAGNITUDES FOR SITE A BASED ON THOSE OBSERVED FOR DEGELEN MOUNTAIN REGION OF THE TEST SITE.



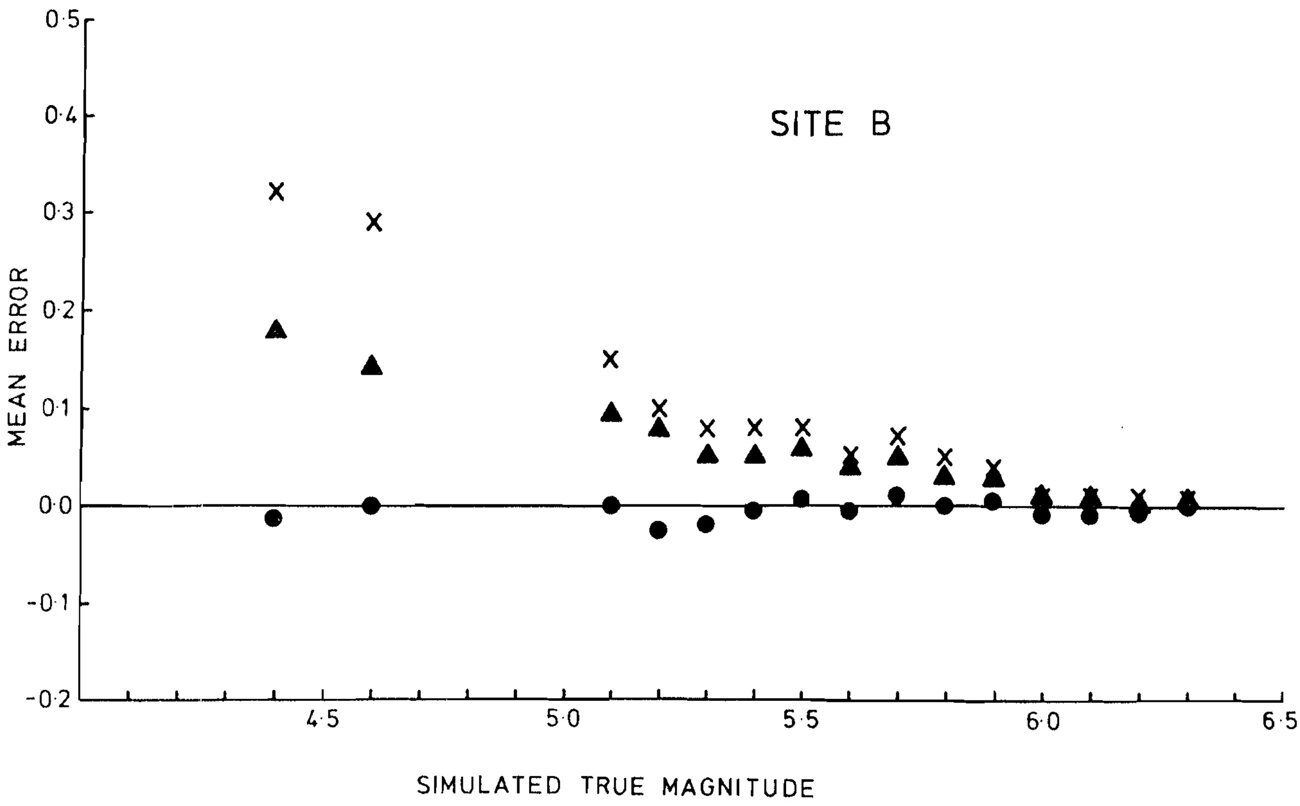
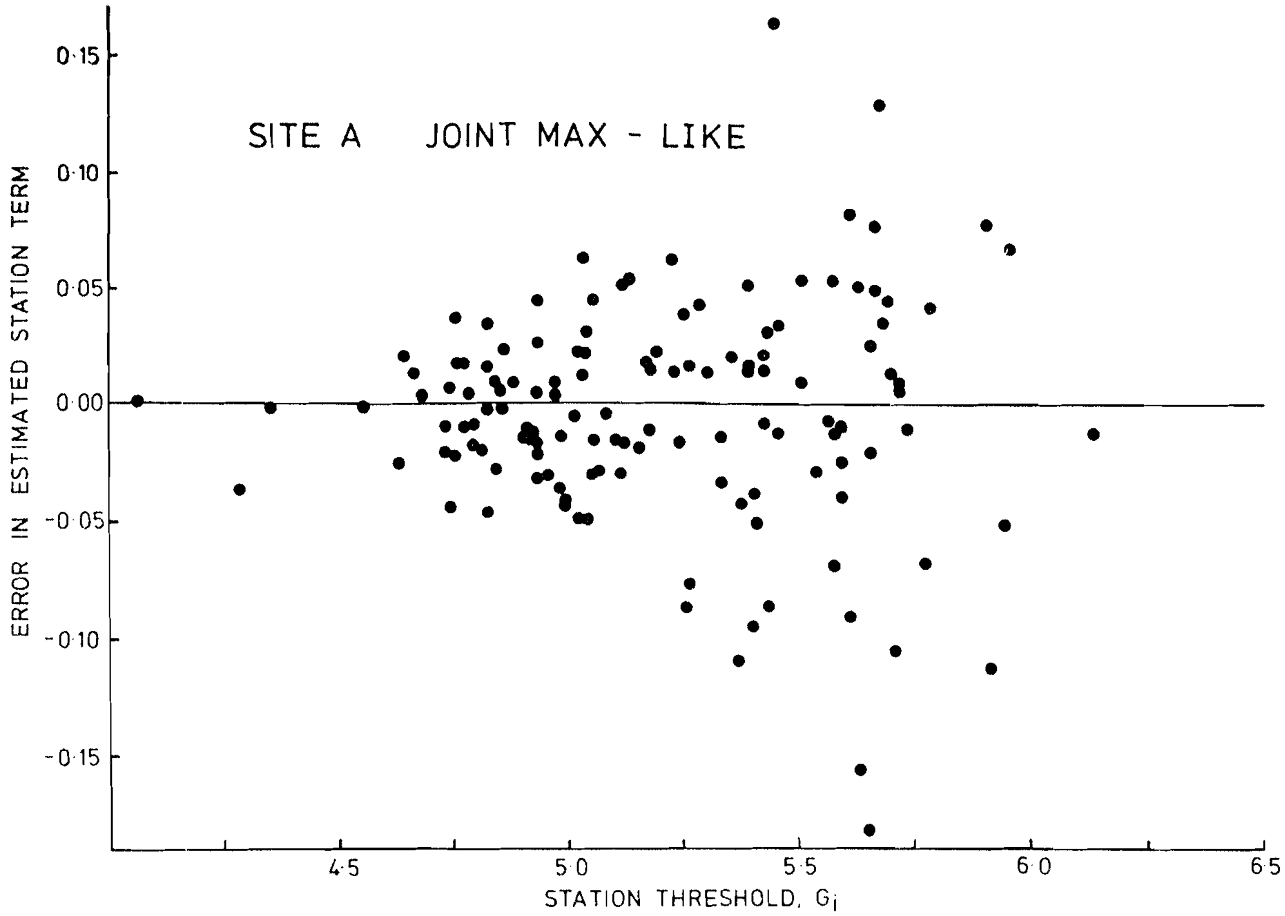


FIGURE 6. MEAN DIFFERENCES FROM THE (SIMULATED) TRUE MAGNITUDES FOR SITES A AND B AS A FUNCTION OF THE TRUE MAGNITUDE.



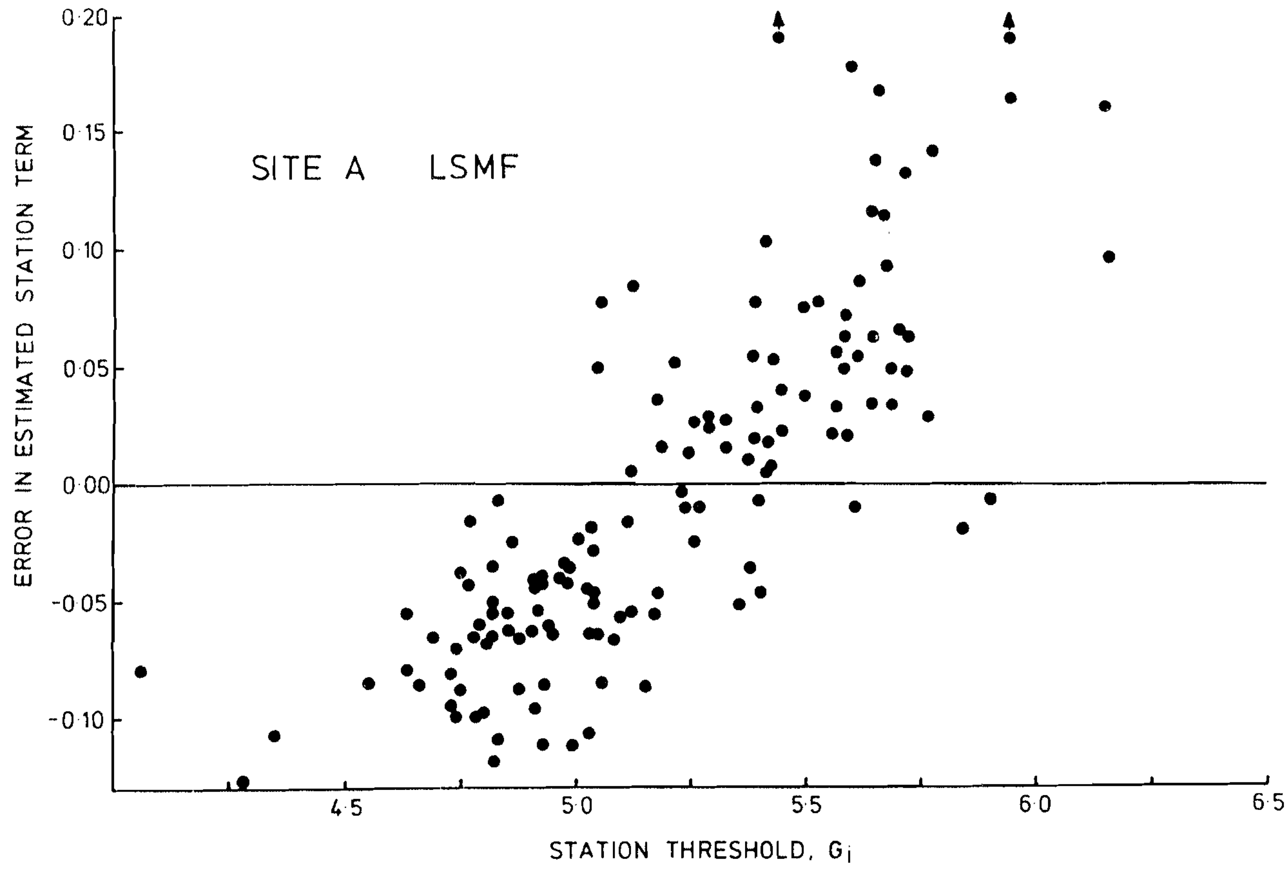


FIGURE 7. ERRORS IN COMPUTED STATION TERMS FOR SITE A (SIMULATED DEGELEN MOUNTAIN) PLOTTED AS A FUNCTION OF STATION THRESHOLD.

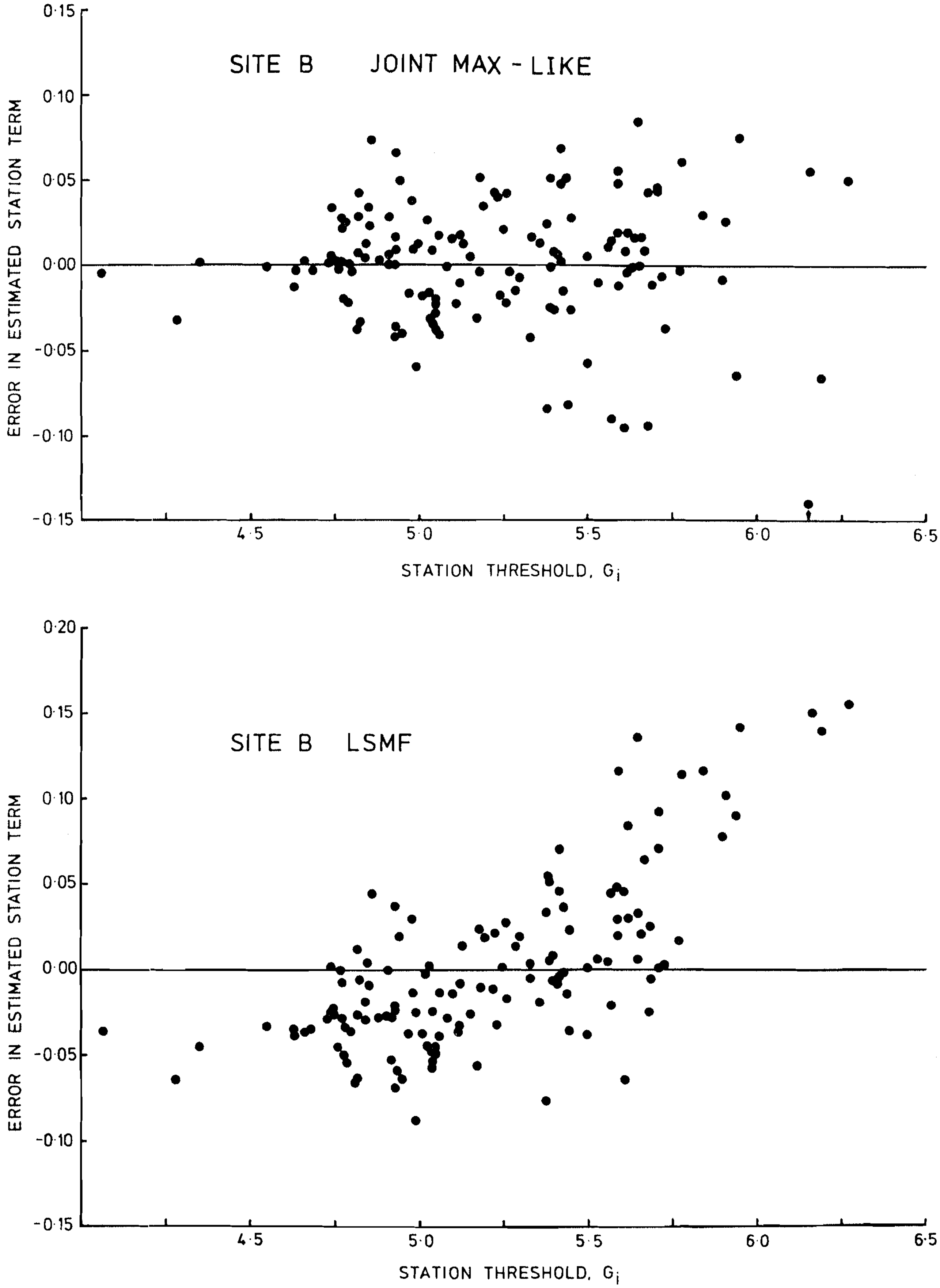


FIGURE 8. ERRORS IN COMPUTED STATION TERMS FOR SITE B (SIMULATED SHAGAN RIVER) PLOTTED AS A FUNCTION OF STATION THRESHOLD.

DOCUMENT CONTROL SHEET

Overall security classification of sheet .... Unclassified ....

(As far as possible this sheet should contain only unclassified information. If it is necessary to enter classified information, the box concerned must be marked to indicate the classification eg (R), (C) or (S)).

1. DRIC Reference (if known)	2. Originator's Reference	3. Agency Reference	4. Report Security Classification
-	AWRE Report No. 022/86	-	UK UNLIMITED

5. Originator's Code (if known)	6. Originator (Corporate Author) Name and Location
-	Atomic Weapons Research Establishment, Aldermaston
5a. Sponsoring Agency's Code (if known)	**6a. Sponsoring Agency (Contract Authority) Name and Location**
-	-

7. Title Some Simulation Studies on Seismic Magnitude Estimators

7a. Title in Foreign Language (in the case of Translation)

-

7b. Presented at (for Conference Papers). Title, Place and Date of Conference

-

8. Author 1.Surname, Initials	9a. Author 2	9b. Authors 3, 4	10. Date	pp	ref
Lilwall, R C	-	-	November 1986	24	17

11. Contract Number	12. Period	13. Project	14. Other References
-	-	-	-

15. Distribution Statement

See front cover

16. Descriptors (or Keywords) (TEST)

SEISMIC WAVES
UNDERGROUND EXPLOSIONS
NUCLEAR EXPLOSIONS
SIMULATION
LEAST SQUARES METHOD

Abstract Conventional seismic magnitude estimates are frequently biased because of the absence of station readings resulting from poor signal-to-noise ratio on the seismograms. Several magnitude estimation techniques, designed to overcome this threshold bias problem, have been published and in this report some of their properties are described using simulation experiments. The presence of bias in the joint least squares analysis method (LSMF), where both magnitudes and station terms for several sources are computed simultaneously, is also investigated. A suggested joint "maximum likelihood" approach provides an attractive alternative.