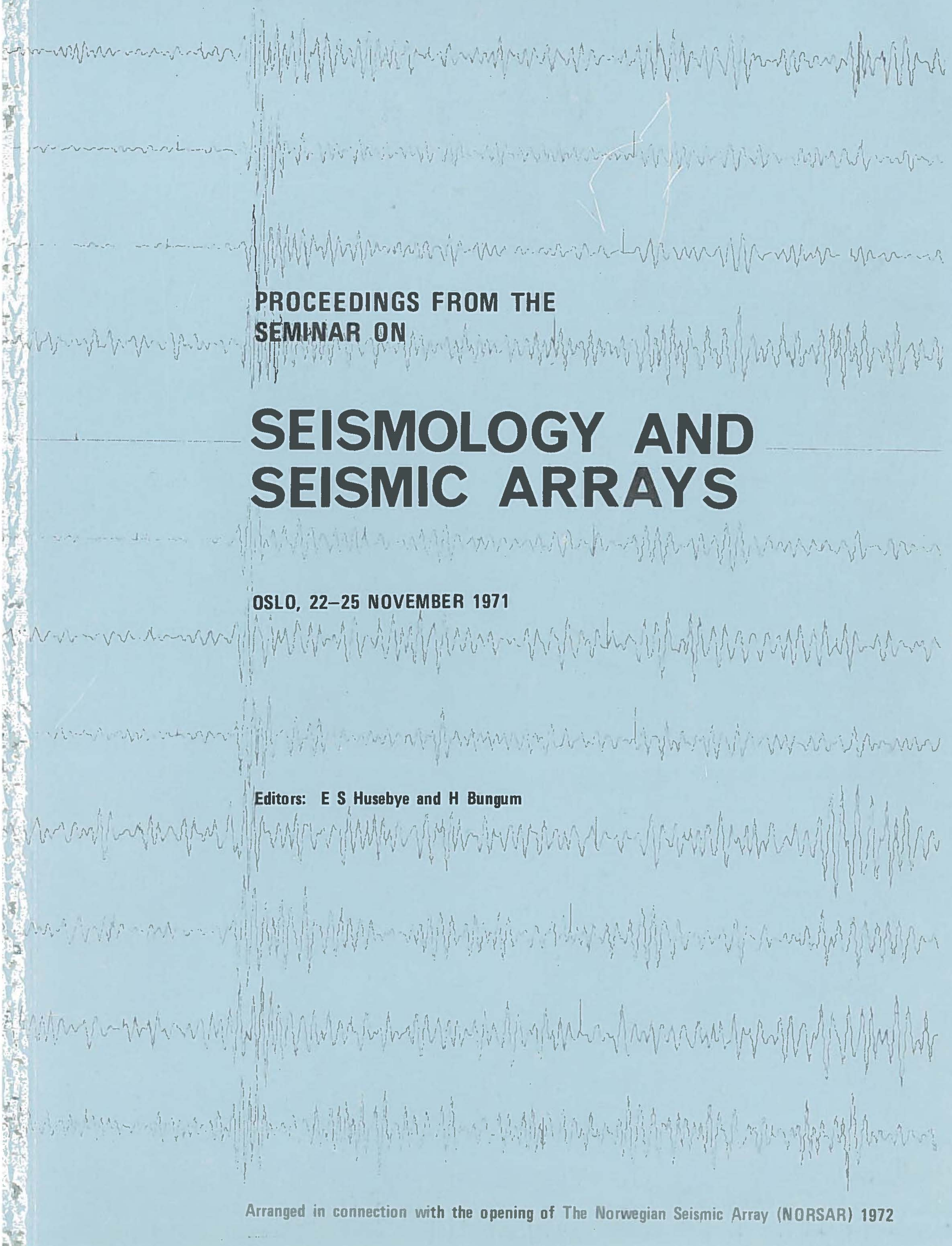
ROYAL NORWEGIAN COUNCIL FOR SCIENTIFIC AND INDUSTRIAL RESEARCH



PROCEEDINGS FROM THE
SEMINAR ON

SEISMOLOGY AND SEISMIC ARRAYS

OSLO, 22–25 NOVEMBER 1971

Editors: E S Husebye and H Bungum

Arranged in connection with the opening of The Norwegian Seismic Array (NORSAR) 1972

AN EVALUATION OF THE ROUTINE PROCESSING OF EVENTS AT NORSAR

H. BUNGUM

NTNF/NORSAR
Kjeller, Norway

ABSTRACT

Data published in the NORSAR seismic bulletin between February and June 1972 has been studied in order to find estimates of the detectability and location accuracy at NORSAR. The detectability is calculated from empirical frequency-magnitude distributions, and the 90% cumulative detectability for the teleseismic zone $(30^\circ<\Delta<90^\circ)$ has been estimated at m_b 4.0, while values for different regions vary from 3.7 to 4.3. These are all NORSAR magnitudes. The magnitude bias between NOAA and NORSAR has been found to be 0.15 ± 0.31 in the teleseismic zone. In this zone, the median location difference between NOAA and NORSAR has been estimated at 160 km, with values for different regions ranging from 130 to 340 km.

INTRODUCTION

The Norwegian Seismic Array came into full operation in the first months of 1971. The array comprises 132 short period vertical seismometers in 22 subarrays, each of which also contains one three-component long period seismometer. The array diameter is about 100 km. The data is digitized at each subarray center, and transmitted to a central recording and data processing center at Kjeller.

The amount of information collected at NORSAR (50 000 baud) is so large that most of it would be useless unless a central, on-line processing facility took care of an initial and automatic data reduction. The research and development needed to build this system took several years, and was by the beginning of 1971 developed to the extent that one then could say that the array was in regular operation.

The array, including the routine data processing, is described in detail by Bungum et al (1971). Fig 1 shows the essential parts of the data processing system at NORSAR. The Detection Processor takes first of all care of the recording of all data on magnetic tapes, including LP data from LASA (Large Aperture Seismic Array) and ALPA (Alaskan Long Period Array), and processed SP data from LASA. Then, the NORSAR SP data is processed in real time in search of seismic events, and a queue of detections, typically 100-300 per day, is created. These detections are nothing more than "event candidates", a few of which, typically 30-50 per day, are selected for further analysis by the off-line Event Processor, which produces an automatic seismic bulletin. This bulletin is reviewed daily by analysts, a step which is equally essential as the automatic processing, and many changes are often made before the bulletin can be distributed. All the professional seismologists at NORSAR have, from time to time, participated here. It has always been a goal in this work that as much as possible of that part of the analyst experience which can be classified and systemized should be fed back into the automatic system.

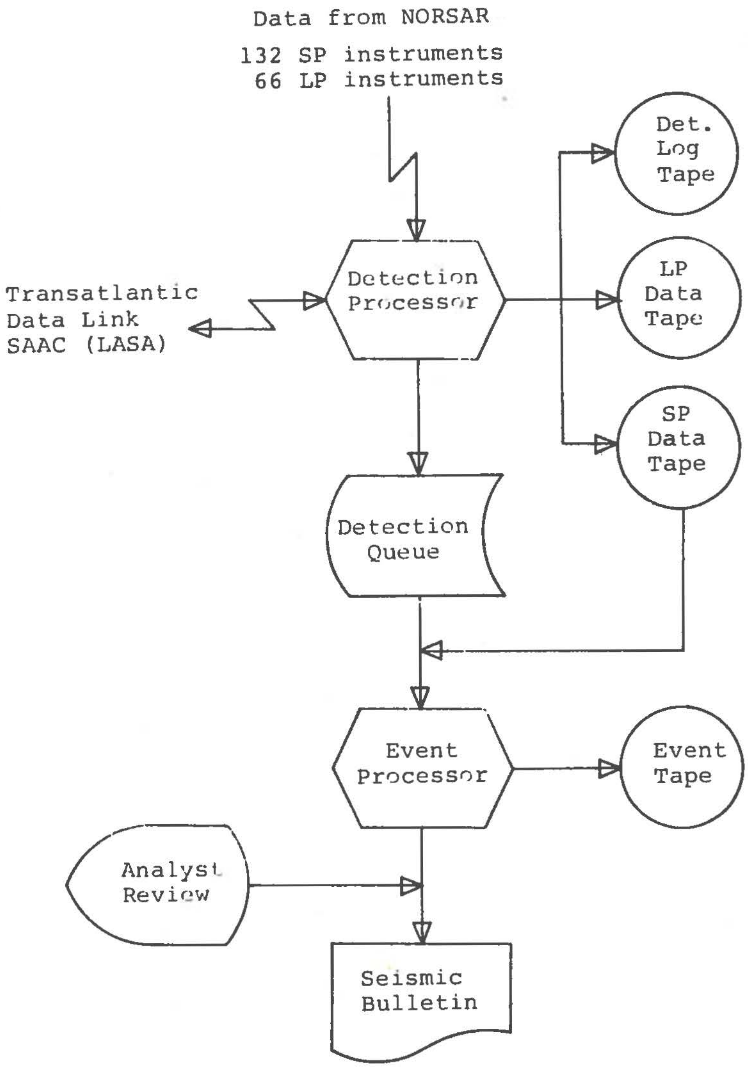


Fig 1. Schematic view of the NORSAR data processing system.

Since 1 May 1971 a reviewed seismic bulletin has been created on a daily basis at NORSAR. This paper is concerned with an evaluation of the data presented in this bulletin. Such evaluation is continuing steadily, and has two main purposes, one of which is to gain experience which in a feedback form can improve the performance of the data processing system. The other purpose, which is the main one in this paper, is to evaluate the performance of the array. That evaluation will here be concentrated on the ability of the array to detect and locate seismic events, estimated through a statistical analysis of the data in the seismic bulletin.

RECENT IMPROVEMENTS

Work continues steadily in order to improve the performance of the array. Some statistical results from the first 6 months of regular operation (May-Oct 1971) were presented by Bungum and Berteussen (1972). These initial results did not quite fulfill the expectations, although it was made clear that the difference could not be explained by any particular error in the system but rather by a combination of possible poor geophysical response and a number of loss sources in the system. Since then, a number of changes have reduced many of these loss factors, leading to significant improvements both in detectability and location accuracy. The main changes are:

- improved analyst performance,
- a new array beam deployment as of 14 Dec 1971,
- on-line filter change from 0.9-3.5 Hz to 1.2-3.2 Hz as of 6 Jan 1972,
- new time delay and location corrections as of 27 Jan 1972.

Also, the average computer time required to process one detection in the Event Processor has been reduced from 9 to 3.5 minutes, which allows more event candidates (detections) to be analyzed with the same computer capacity.

The effect of these changes on the detectability can be seen in Fig 2, which shows the monthly number of events reported by NORSAR for the time period May 1971 - August 1972. Also given is the number of events reported by NOAA and the number reported by both institutions. These curves are also, of course, affected by other factors, first of all time variations in seismicity and long term variations of the noise level (see below). Because of these recent improvements, the results presented here will all be based on data collected after 1 February 1972.

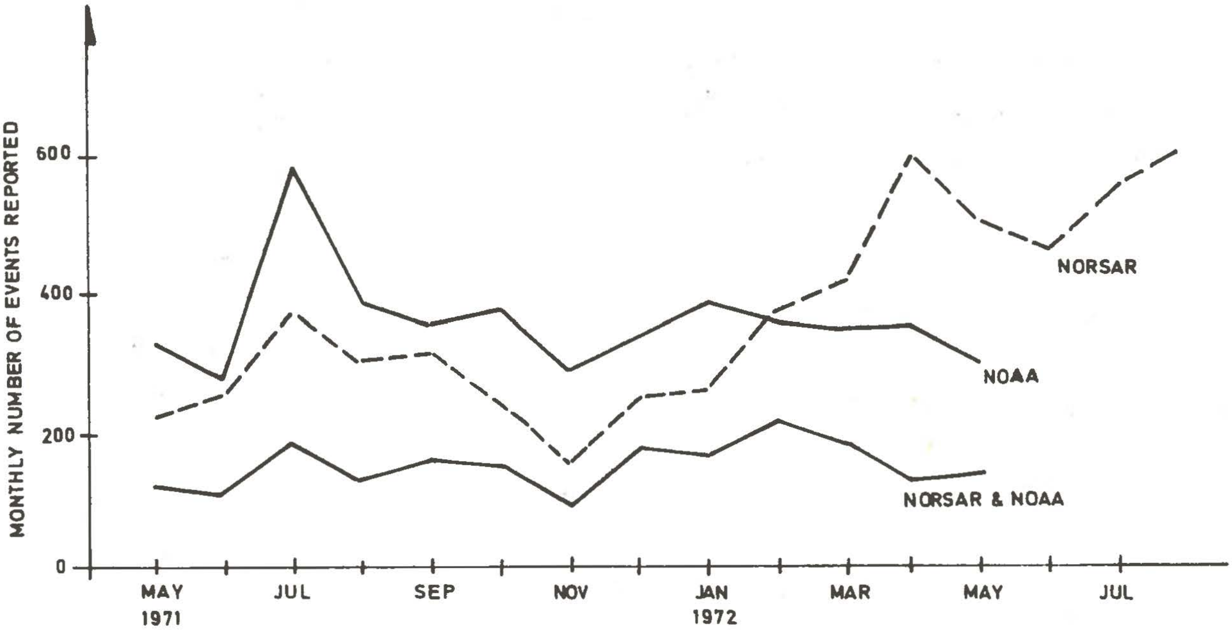


Fig 2. Monthly number of events reported by NORSAR, NOAA and simultaneously by the two institutions.

DETECTABILITY

Detectability as used in this paper can be defined as the long term operational ability to report, with epicentral information, the occurrence of seismic events. The term detectability is therefore used mainly for notational convenience.

One of the aims in the analysis work has been to keep the false alarm rate in the seismic bulletin at a low level. It is not possible to determine this false alarm rate accurately, since there are many regions where no other seismic system reports as many events as NORSAR. However, based on the analysts' experience with the NORSAR receiver operating characteristics we would estimate the false alarm rate under the present operational procedures to be less than 5%. This low false alarm rate cannot be obtained unless a large number of true detections are left unreported. Some of these could have been included by devoting more time and effort to the analysis, and some could probably be confirmed through a study of the bulletins from other networks.

The number of events N above a given magnitude m, within a certain time period, is generally assumed to follow the relationship

$$\mathrm{Log}N = A - b \cdot m \qquad (1)$$

both for the entire world and for more limited geographical regions. On this assumption, empirical frequency-magnitude distributions would then help determine the parameters A and b. Fig 3 shows such a distribution, both incremental and cumulative for the NORSAR teleseismic zone ($30^o<\Delta<90^o$). The slope b is sometimes determined through a maximum-likelihood estimation based on average magnitude (Aki 1965, Shlien and Toksöz 1970). That technique cannot be used on the data in Fig 3 because of the negative bias in the larger magnitudes. This bias is introduced because the 78 dB dynamic range in the digital system cannot give a sufficient resolution for small signals without clipping the larger ones. Therefore, the slope has been estimated by fitting a straight line, in a least squares sense, through the straight part of the cumulative frequency-magnitude curve. Then, the 50% and 90% levels for detectability are computed automatically, based on computation of the assumed number of missed events at any particular magnitude, i.e., the difference between the extended straight line and the empirical curve. As Fig 3 shows, the NORSAR teleseismic data from Feb-June 1972 has a slope of b = 1.00 ± 0.01, and the 50% and 90% detectability levels are m_b 3.6 and 4.0 respectively.

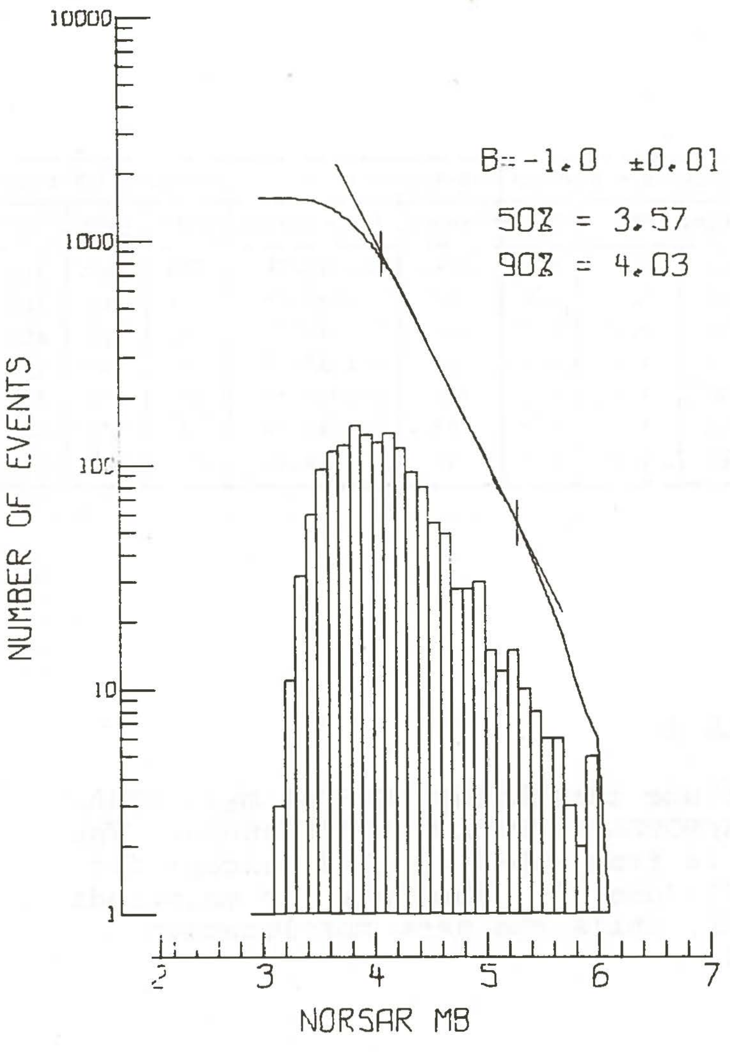


Fig 3. Interval and cumulative frequency-magnitude distribution for data from Feb-Jun 1972, range 30-90 deg. Straight line is a least squares fit through data within bars.

It is important to point out that the magnitudes which are quoted above are all NORSAR estimates. In order to make comparisons with other networks possible, the relation between NORSAR and NOAA magnitudes has therefore been investigated. Another point worth noticing is that by presenting data from the entire teleseismic zone, one is combining data with possible different statistical distributions. Therefore, all the results are also presented regionalized, where the regions are defined in distance and azimuth from NORSAR, as given in Table 1.

Region		Regional limits (deg)		Magnitude threshold			Magnitude bias		Location difference		
No	Name	Azimuth	Distance	Events	50%	90%	Events	NOAA-NORSAR	Events	50%	90%
A		0-360	30-90	1555	3.6	4.0	848	0.15±0.31	509	160	510
A1	Atlantic	180-260	30-90	88	3.6	4.3	13	0.45±0.28	11	340	780
A2	N. America	260-340	40-90	114	3.8	4.2	100	0.20±0.33	61	260	810
A3	Aleutians	340-15	30-90	131	3.4	3.9	119	0.17±0.35	57	150	370
A4	Japan	15-70	50-90	738	3.7	4.1	441	0.10±0.29	271	130	530
A5	C. Asia	40-110	30-60	211	3.2	3.7	89	0.20±0.34	43	130	270
A6	Iran	110-180	30-90	262	3.5	3.8	39	0.23±0.29	53	180	520

TABLE 1

Regionalized results for magnitude thresholds (NORSAR m_b), NOAA/NORSAR magnitude bias and NOAA/NORSAR location differences. The data for magnitude thresholds is from Feb-June 1972, except for region A1 with data from May 71-June 72. The data for magnitude bias is all from May 71-June 72, while the data for location differences is from Feb-May 72.

These regionalized results for the magnitude thresholds are all listed in Table 1. As one can see, the 90% detectability level is m_b 4.0 when all teleseismic data is used, while the values for different regions vary from 3.7 in Central Asia to 4.3 on the Mid-Atlantic Ridge. As mentioned above, this cannot be fully evaluated before the possible magnitude bias between NORSAR and some known reference, say NOAA, has been investigated. Fig 4 shows the NOAA/NORSAR magnitude difference as a function of epicentral distance, where a clear negative bias in the NORSAR data is observed for

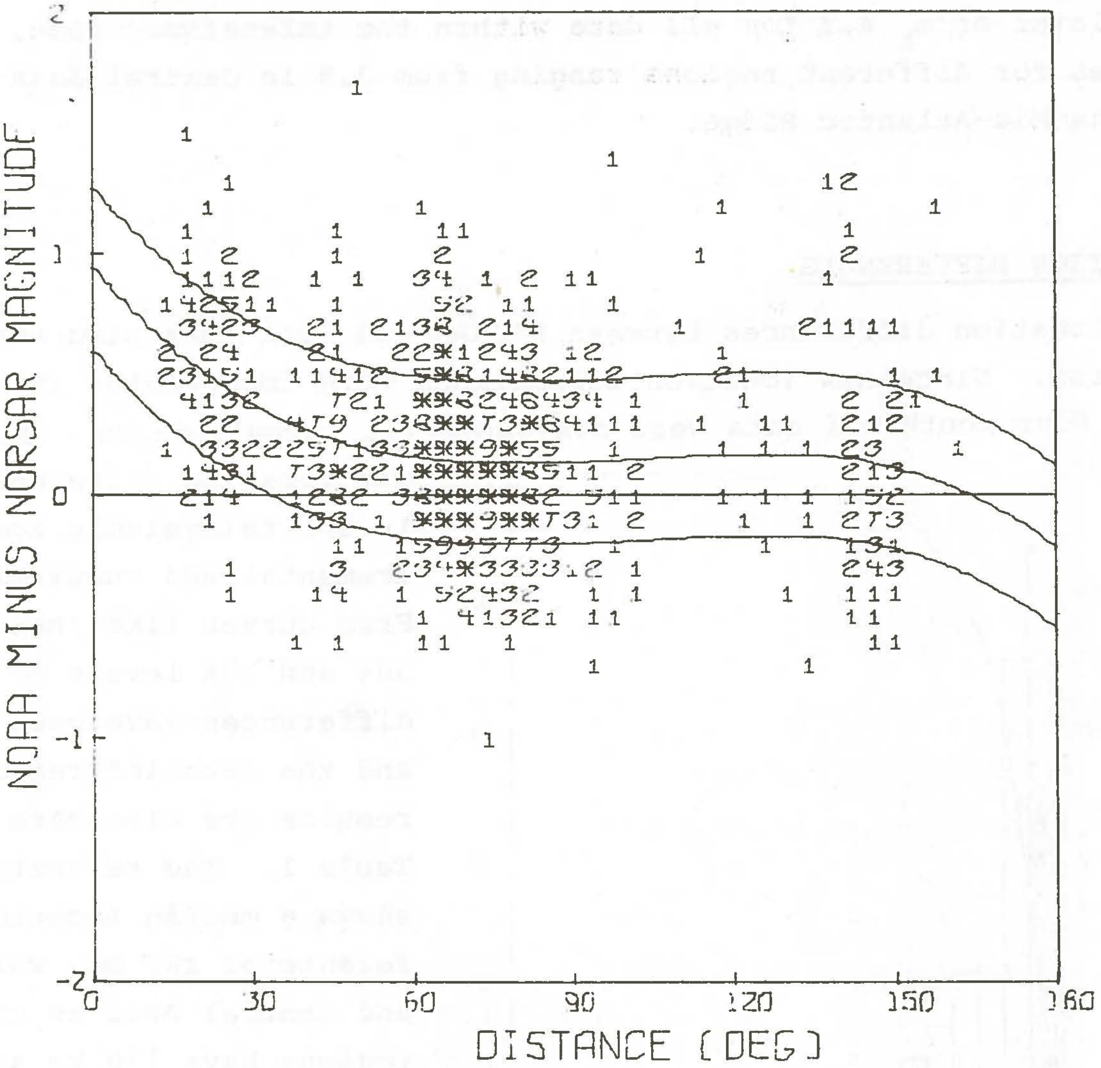


Fig 4. NOAA/NORSAR magnitude differences vs. epicentral distance for data from May 71-March 72. The curves represent a third degree least squares fit through data, with upper and lower bounds (STD).

epicentral distances smaller than 30°. A likely explanation of that bias is the fact that magnitude is measured on the array beam, and more local events have, due to poor coherence across the array, a significant beamforming loss which is not compensated for in the magnitude calculation. As Fig 4 shows, the scatter in the magnitude data is quite large. Table 1, which gives all the detailed results also for the magnitude bias, shows a bias of 0.15 ± 0.31 for all data within 30°-90°, while the different regions have values ranging from 0.10 in Japan to 0.45 on the Mid-Atlantic Ridge. The scatter in the data is approximately of the same size for all regions.

Now, if one should express the NORSAR magnitude thresholds in terms of some "NOAA equivalent magnitude", one would have to add the bias to the threshold values in Table 1. By doing so, one would get a

90% level of m_b 4.2 for all data within the teleseismic zone, with values for different regions ranging from 3.9 in Central Asia to 4.7 on the Mid-Atlantic Ridge.

LOCATION DIFFERENCES

The location differences between NORSAR and NOAA have also been studied. Since new location corrections were implemented in Jan 1972, only four months of data were available for investigation. Fig 5 shows the location differences within the teleseismic zone, incremental and cumulative. From curves like that, the 50% and 90% levels of location differences have been found, and the detailed regionalized results are also here given in Table 1. The teleseismic zone shows a median location difference of 160 km, while Japan and Central Asia as the best regions have 130 km and North America as the poorest has 340 km. The expression location difference (and not location error) has been used because the comparison is made between two estimates which both are affected by uncertainties. Since the NOAA standard error of location is in the range of 10-40 km, this clearly becomes significant for regions where Table 1 shows a location difference of about 120 km, as for Japan.

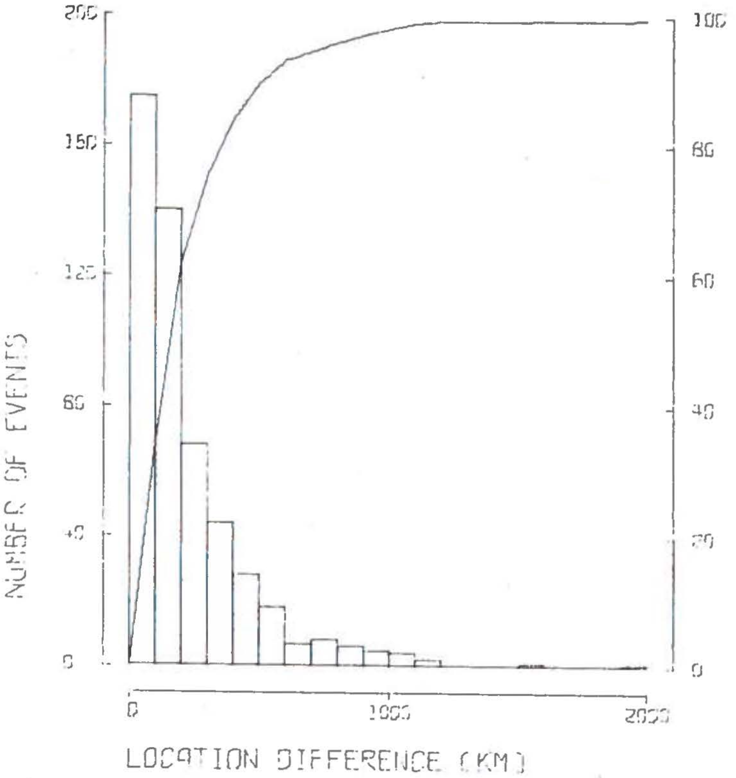


Fig 5. Interval and cumulative distribution of location differences between NOAA and NORSAR for data from Feb-May 1972, range 30-90 deg.

DISCUSSION

It is important to keep in mind when reading Table 1 the special definition of detectability given above. There could easily be a significant difference between the operational and the optimum ability to detect and report events, and both should be investigated.

Another factor of significance is that some of the regions presented in Table 1 so far have not very much data. The regional differences are, however, so large that the trends are quite clear.

Thirdly, the data presented herein is mainly from a time of year when the background noise level is moderate. The noisiest time period so far has been Oct-Dec 71, a period which also shows a minimum in the number of reported events (Fig 2). This has been found from an extensive study of the long term short period noise level within the on-line processing frequency band. For the time period Jan-May 1972, the median noise level within the frequency band 1.2-3.2 Hz has been found to be 1.15 mμ. Some uncertainties apply to the ground motion conversion in this case, while the estimates for the relative variations are more accurate. The 90% level is 2.0 dB above the median and the 10% level is 2.7 dB below. The long term average can be well approximated by a Gaussian distribution, while the short term average of the noise seems to be lognormal (Lacoss 1972).

Starting with this median noise level and a representative short term noise distribution (based on a few hours of data), the detectability has also been calculated after adding the size and distribution of the loss factors in the system. (K.A. Berteussen, personal communication.) This procedure leads to an estimate of $m_b = 4.0$ for the 90% cumulative detectability, in the teleseismic zone, which is the same value as found from the frequency-magnitude distribution.

A third possibility for estimating the detectability is to calculate it from the number of reported events. One must then assume a given frequency-magnitude distribution for the yearly seismicity of the world, say Richter's (1958), who has estimated the parameters A and b in equation (1) to be 8.2 and 1.0, respectively. NORSAR has for the 5 months between Feb and May 1972 reported 1990 events from the teleseismic zone, and if one assumes that 45% of the world seismicity is covered in that zone ($30^o<\Delta<90^o$), it would equal N = 10600 events worldwide for one year. By inserting these values for A, b and N in equation (1), and assuming a 25% uncertainty in N, it would give a (100%) threshold of $m_b = 4.2 \pm 0.1$. This should be compared to the "NOAA equivalent" threshold from Table 1, i.e., NORSAR threshold plus bias, which also is around 4.2. When considering all the assumptions behind this estimate, the only safe conclusion is that there is no discrepancy between the number of reported events and the threshold values given in Table 1.

As for the location differences, the regionalized results also here would have some uncertainties caused by the limited amount of data. Another factor worth mentioning is that the comparison is made only for events above the NOAA reporting threshold, which for some regions is significantly higher than the NORSAR threshold. One should expect that the location error for small events on an average is somewhat larger, but not much, since the error in most cases would still be within the beam radius, typically 2-3 degrees.

A few possible areas of future improvements should also be mentioned. First, there will be implemented a new set of time delay and location corrections, and although most of the possible gain has been extracted from these areas, the updating is expected to improve the results somewhat. Secondly, incoherent beamforming is now under implementation, and this will improve the detectability somewhat, especially for epicentral distances less than 30^{o}. Thirdly, one should also mention the possibility of introducing amplitude weighting in the beamforming, in order to take advantage of the large amplitude anomalies at NORSAR. This is now under investigation.

REFERENCES

1. K. Aki: Maximum likelihood estimate of b in the formula logN = a-bM and its confidence limits, Bull. Earthq. Res. Inst., Vol 43, 237-239, 1965.

2. H. Bungum and K.A. Berteussen: An evaluation of the routine processing of events at NORSAR during the time period May-October 1971, NORSAR Technical Report No. 24, 1 March 1972.

3. H. Bungum, E.S. Husebye and F. Ringdal: The NORSAR array and preliminary results of data analysis, Geophys. J.R.Astr. Soc., Vol 25, 115-126, 1971.

4. R.T. Lacoss: Variations of false-alarm rates at NORSAR, Semiannual Technical Summary, MIT, Lincoln Lab, 30 June 1972.

5. C.F. Richter: Elementary Seismology, W. Freeman & Co., San Francisco, 1958.

6. S. Shlien and M.N. Toksöz: Frequency-Magnitude statistics of earthquake occurrences, Earthquake Notes, Vol 16, No 1, 5-18, 1970.