SEISMOLOGICAL SERIES

of the

DOMINION OBSERVATORY

1968-5

Correlogram Discrimination Parameters from Yellowknife Seismic Array Data

K. Whitham, P. W. Basham
and H. S. Hasegawa

Seismological Service
of Canada

OTTAWA, CANADA

Department of Energy, Mines and Resources

OBSERVATORIES BRANCH

1969

SEISMOLOGICAL SERIES
of the
DOMINION OBSERVATORY

1968-5

CORRELOGRAM DISCRIMINATION PARAMETERS FROM YELLOWKNIFE SEISMIC ARRAY DATA

K. Whitham, P.W. Basham and H.S. Hasegawa

Seismological Service

of Canada

OTTAWA, CANADA

Department of Energy, Mines and Resources

OBSERVATORIES BRANCH

1969

CONTENTS

CORRELOGRAM DISCRIMINATION PARAMETERS

FROM YELLOWKNIFE SEISMIC ARRAY DATA

INTRODUCTION

Recently there has been considerable interest in the application of basic decision theory to the analysis of the efficiency of seismological event identification for purposes of deterrence and control during a ban on underground testing (Ericsson, 1967, 1968).

The purpose of this note is not to discuss or criticize in detail the complex matters and assumptions involved, but to present some relevant numbers derived from a digital computer analysis of events recorded on the Yellowknife array, N.W.T., Canada. The correlogram parameters to be discussed are rudimentary ones, but have been derived from an unbiassed population of world-wide earthquakes and nuclear explosions with epicenters in the third zone (30° - 90°) from Yellowknife.

For a description of the Yellowknife seismic array, the processing methods used, and results obtained, the reader is referred to Manchee and Somers (1966), Somers and Manchee (1966), Weichert, et al. (1967), and Manchee and Weichert (1968).

CORRELOGRAMS AND COMPLEXITIES

As an aid to seismic event identification, the output of the Yellowknife array data processing is usually presented, in part, as a correlogram, the smoothed product of the phased sums of seismic signals along the NS and EW arms of the array. For the Montana Large Aperture Seismic Array (LASA), the product of two separately summed F-ring subarrays is presented as a correlogram. The correlogram is a convenient array output display because it represents in a single time-function those portions of the signals which are strongly coherent across the array, thereby reducing any signal-generated noise. The correlogram also provides a measure of the rate of energy arrival with time after the signal onset, and can be used as a diagnostic aid to the identification of earthquake and underground nuclear explosion sources.

Explosion P-wave seismograms at third zone distances are usually impulsive and relatively simple with nearly all the energy arriving in the first few seconds. The P signal (and therefore the correlogram) of most earthquakes is more complex and includes signal energy arriving after many seconds. These correlogram features can be related to the source properties: an explosion source generates P waves only and is simple in space and time, whereas an earthquake generates more S waves which by S to P conversion near the source extends the P coda. In addition, an earthquake source may be a multiple shock in space or time. Explosions must be comparatively shallow from engineering considerations and therefore have the depth phases pP and sP following P very closely (less than about 1 second), whereas earthquakes can contain later coda energy from these depth phases.

Studies on the diagnostic properties of earthquake and explosion P-wave correlograms and waveforms have been reported by the Atomic Weapons Research Establishment (AWRE) group using U.K. array data (Carpenter, 1964; AWRE Staff, 1965; Douglas, 1967) and by the Lincoln Laboratory, M.I.T. group, using LASA data (Greenfield, 1966; Kelly, 1966, 1968). AWRE Staff (1965), using 150 shallow earthquakes and 13 explosions with magnitude near m5.0 recorded on 3 U.K. arrays, found that 6 per cent only of the earthquakes could not be distinguished from explosions using complexity and symmetry criteria. Their definition of complexity was qualitative: symmetry refers to similar correlograms from different arrays at different azimuths at teleseismic distances. Douglas (1967) classified correlogram complexity broadly into five groups ranging from 'very simple' to 'very complex', but gave no event statistics for these categories. Douglas also introduced the quantitative concept of the correlogram 'energy ratio' or 'complexity coefficient', the ratio of the area under the correlogram in the first 5 seconds after onset to the area between 5 and 35 seconds after onset.

Three years ago an explosion in Novaya Zemlya gave complex records at Yellowknife and elsewhere in North America, although simple records at other U.K.-type arrays. Further study of Novaya Zemlya events by AWRE led to the revised estimate that 60-65 per cent only of all earthquakes of m5.0 and greater could be identified as such (Davies, 1968) using complexity and symmetry.

The problem is obviously one in which studies should be made systematically as a function of magnitude and geographical area or test site location, although this has not yet been done.

The Lincoln Laboratory published preliminary studies on LASA correlogram complexity coefficients (Greenfield, 1966), but have based their recent studies on a beam waveform complexity coefficient (Kelly, 1966, 1968) defined as

$$C = \int_{5}^{35} \left| x(t) \right| dt \Big/ \int_{0}^{5} \left| x(t) \right| dt$$

where $x(t)$ is the bandpass filtered delayed sum from a subarray and the adopted C is the average value from the various subarrays. Kelly (1968) presents results for 85 earthquakes (m3.8 - 6.2) and 19 explosions (m5.0 - 6.4). Extracting statistics from Kelly's cumulative histograms in a form used below for Yellowknife correlogram parameters, we find that 90 per cent of the explosions have $C < 1.2$, compared to only 25 per cent of all earthquakes, 37 per cent of shallow earthquakes, and 54 per cent of shallow continental earthquakes.

YELLOWKNIFE CORRELOGRAM PARAMETERS

The measurements of this study are based on all computer output of third zone earthquakes for the period January-October, 1966, and all available computer processed third zone explosions up to and including 1966. The earthquake output was acquired under a variety of operating conditions, namely, (1) free search in which each tape is processed automatically for all detectable events, (2) selected search in which tapes are processed for all events visible on the helicorder monitor record, and (3) special search in which only events of special interest are processed. For purposes of this study, the earthquakes and explosions have been considered simply as populations of third zone events with no sorting with respect to magnitude, geographic location or focal depth, although it is emphasized that this sorting is essential in any complete application to the test ban problem. It can be mentioned, however, that under free search conditions the cumulative third zone 50 per cent detection level is about m4.0 (Weichert, et al., 1967). The populations of events used here are 742 earthquakes and 35 explosions, a much larger population than that studied by others.

The measurements made on the correlograms are parameters which intuitively should provide discrimination between earthquakes and explosions, but which do not necessarily presuppose geophysical reasons for their choice. The three parameters are: the rise-time (τ) in seconds from the start of the event to the peak energy arrival in the correlogram, the quarter-width (ω) in seconds which is the width of the correlogram at the points where it has fallen by 1/4 from the peak value (analogous to the half-width parameter commonly used in geophysical interpretation), and the coda-peak ratio (γ) which is the ratio of the largest correlogram amplitude occurring after 20 seconds from the start to the amplitude at the correlogram peak. If seismological relationships between these parameters and the P signal are to be given, τ can be considered a measure of the impulsiveness (albeit smoothed by the correlogram averaging time, which is effectively exponential smoothing with a time constant of 1.6 seconds), ω a measure of the duration, and γ a measure of the proportion of late-arriving energy, of the P signal. On the basis of qualitative characteristics of earthquake and explosion P waves discussed in the previous section, each of the parameters, τ , ω , and γ , should, in general, be smaller for explosions than for earthquakes.

Figure 1 shows cumulative percentage distributions of each of the parameters for the earthquakes and explosions. The earthquake curves are denoted by Q and the explosion curves by E . The bracketed numbers are the total numbers of events used for each curve. τ was measured for all earthquakes, but for technical reasons pertaining to the method of correlogram output display, ω and γ measurements could not be made for a number of the earthquakes. The parameters are also combined to form a double-product parameter $\tau\omega$ and a triple-product parameter $\tau\omega\gamma$. Cumulative percentage distributions of the product parameters are shown in Figure 2.

The relative discrimination capabilities of the five parameters are shown more clearly in Figure 3 by what Ericsson (1968) defines as 'identification curves' than by the pairs of cumulative percentage curves themselves. In Figure 3 the values of the earthquake and explosion cumulative percentages for each value of the parameter are plotted against each other; thus each point on a curve corresponds to a particular value of the parameter. The farther an identification curve is to the left in Figure 3, the greater is the efficiency of that parameter in separating earthquakes from explosions. There is not a large difference between the positions of some of the curves, except that ω seems to be the least efficient.

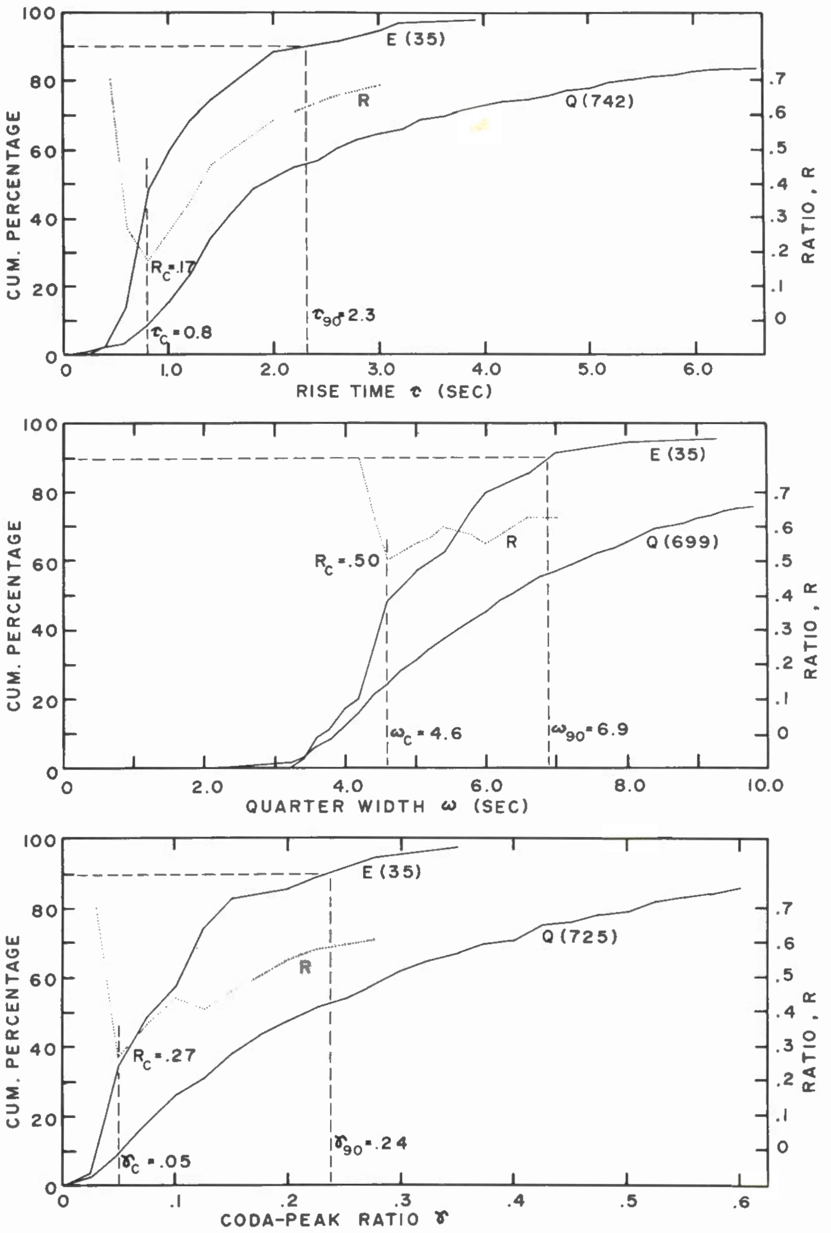

Figure 1. Earthquake and explosion cumulative percentage distributions and ratio curves for single parameters.

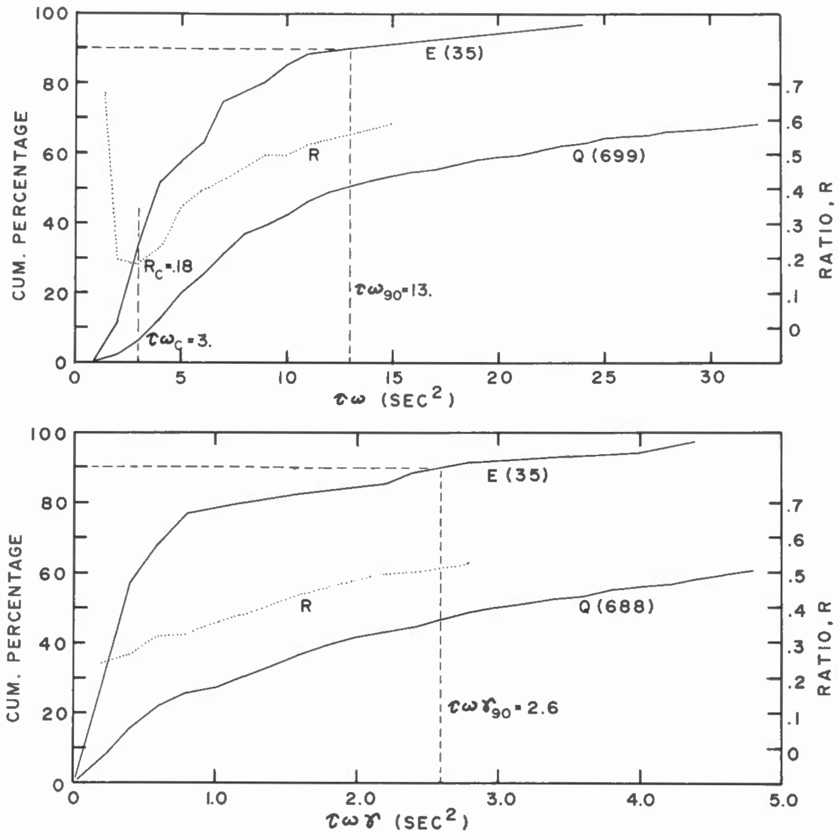

Figure 2. Earthquake and explosion cumulative percentage distributions and ratio curves for product parameters.

PARAMETER STATISTICS

There are a number of ways in which the statistics of these single and product correlogram parameters can be presented to illustrate their value to the earthquake-explosion discrimination problem. Two procedures are discussed here. The first is a tabulation of the percentages of earthquakes which satisfy various single and multiple parameter conditions which, in turn, are satisfied by 90 per cent of the explosions, i.e., the percentages of earthquakes which cannot be distinguished from 90 per cent of the explosions on the basis of the conditions. These are designated as the '90 per cent statistics'. The second procedure is to determine the distribution of earthquakes and explosions relative to parameter decision levels which provide the best separation of explosions from earthquakes; this is a measure of the overlap of the two populations with respect to various discrimination criteria. These are designated as 'decision level statistics'.

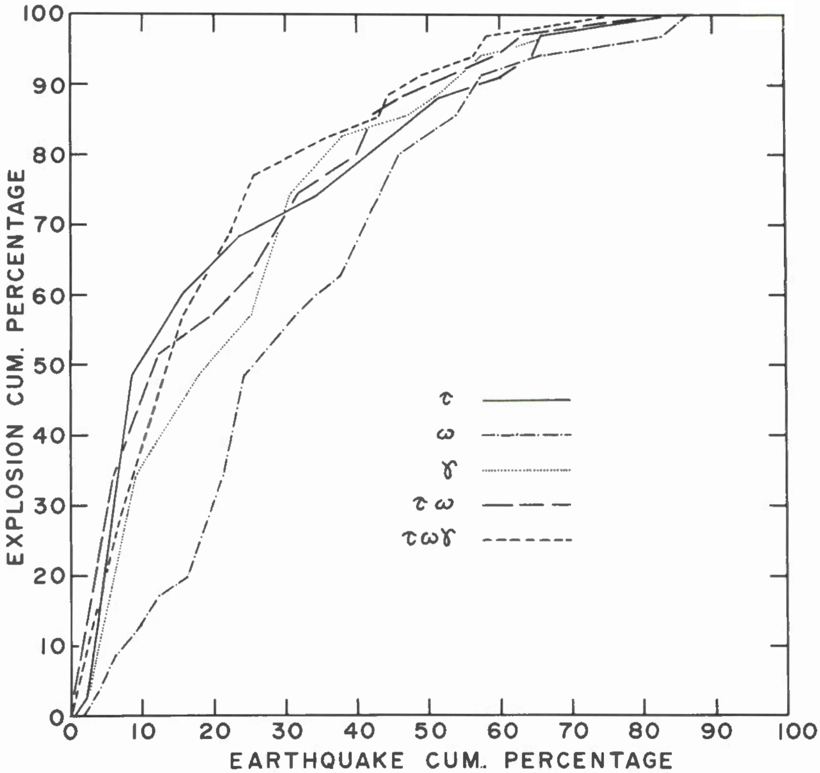

Figure 3. Identification curves for three single parameters and two product parameters.

90 Per Cent Statistics

On each of the cumulative distribution diagrams for the single parameters in Figure 1 and product parameters in Figure 2, the value of the parameter at the 90 per cent explosion level (parameter with subscript 90) is shown by a broken vertical line. The percentages of earthquakes with parameter values satisfying the single conditions, (parameter) $<$ (parameter)$_{90}$, are given in Table I(a). With regard to separation of the greatest number of earthquakes from 90 per cent of the explosions by these single conditions, γ is slightly more effective than the other single parameters, and further slight improvement is gained from the double and triple product parameters.

Table I(b) gives the percentages satisfying the three possible double 'AND' conditions and the triple 'AND' condition, using single parameters. The explosion percentages decrease from the nominal 90 per cent level, but the earthquake percentages also decrease considerably from the values for single conditions. Table I(c) gives the results for the double and triple 'OR' conditions using the single parameters. The explosion percentages increase, two of them to 100 per cent; the earthquake percentages

Table I

PERCENTAGES OF EVENTS SATISFYING '90 PER CENT' CONDITIONS

Condition	Earthquakes	Explosions
a) Single Conditions:		
$\tau < \tau_{90} = 2.3$ sec	56	90
$\omega < \omega_{90} = 6.9$ sec	56	90
$\gamma < \gamma_{90} = 0.24$	52	90
$\tau\omega < \tau\omega_{90} = 13$ sec^2	51	90
$\tau\omega\gamma < \tau\omega\gamma_{90} = 2.6$ sec^2	47	90
b) 'AND' Conditions:		
$\tau < \tau_{90}$ AND $\omega < \omega_{90}$	43	83
$\tau < \tau_{90}$ AND $\gamma < \gamma_{90}$	31	77
$\omega < \omega_{90}$ AND $\gamma < \gamma_{90}$	38	77
$\tau < \tau_{90}$ AND $\omega < \omega_{90}$ AND $\gamma < \gamma_{90}$	29	74
c) 'OR' Conditions:		
$\tau < \tau_{90}$ OR $\omega < \omega_{90}$	68	91
$\tau < \tau_{90}$ OR $\gamma < \gamma_{90}$	77	100
$\omega < \omega_{90}$ OR $\gamma < \gamma_{90}$	75	97
$\tau < \tau_{90}$ OR $\omega < \omega_{90}$ OR $\gamma < \gamma_{90}$	82	100
d) Single Reduced Percentage Conditions:		
$\tau < \tau_{83} = 1.8$ sec	47	83
$\tau < \tau_{77} = 1.5$ sec	38	77
$\tau < \tau_{74} = 1.4$ sec	33	74
$\tau\omega < \tau\omega_{74} = 7$ sec^2	32	74
$\tau\omega\gamma < \tau\omega\gamma_{74} = 0.7$ sec^2	24	74

increase to rather large values. The 'AND' condition improvement is largely illusory, as Table I(d) shows. Thus the percentage of earthquakes satisfying '74 per cent explosion statistics' is 33 per cent applying τ alone, 32 per cent applying $\tau\omega$ alone, 24 per cent applying $\tau\omega\gamma$ alone, but only 29 per cent for the triple 'AND' condition.

To summarize these statistics in statement form, let us assume that the condition, (parameter) $<$ (parameter)$_{90}$, is an explosion property, violated by (nominally) only 10 per cent of the explosions. Table I(a) shows that if only (and any) one of the single or product parameters is considered, roughly half of the earthquakes will appear explosion-like; Table I(b) shows that if an earthquake must have two of, or all three of, the explosion properties to appear explosion-like, then roughly one third of the earthquakes are explosion-like (but the percentage of explosion-like explosions also decreases); Table I(c) shows that if an earthquake is explosion-like when it has any of the explosion properties, then more than three quarters of the earthquakes are explosion-like (but nearly all of the explosions will be explosion-like).

Decision Level Statistics

Routine use of these parameters for event identification (or more realistically, their use as diagnostic aids in the identification process) requires the choice of parameter decision levels such that events with parameters greater than the decision level are taken as earthquakes and events with parameters less than the decision level are taken as explosions. The choice of a decision level depends on the detailed shapes of the event distributions with respect to the parameter and on the degree of overlap of the two populations. In the ideal case of completely separate distributions, the decision level is easily chosen as some parameter value between the two distributions. In the case of overlapping populations, it is shown by Ericsson (1967, 1968) that the best choice of a parameter decision level for minimizing inspections at constant deterrence for test ban control purposes is at the minimum $R = Q/E$ ratio, provided the cumulative percentage of explosions (E) with parameter less than the parameter at the minimum is sufficiently high.

This ratio is shown as a dotted line for each of the parameters in Figures 1 and 2. The decision level is taken as the parameter value (parameter subscript c) at which the ratio is a minimum (R_c). Small R_c values indicate that the explosion cumulative percentage has risen to a large value relative to a small earthquake percentage. Because of skew distribution of the populations and indefinite peak separations, some ratio minima are poorly defined and the R_c values quite large. In the case of $\tau\omega\gamma$ in Figure 2, the ratio has no minimum. A list of the percentages of events separated by the various decision levels is given in Table II. A decision level based on τ appears the best of the parameter decision levels tested--at the ratio minimum, 92 per cent of the earthquakes have $\tau > \tau_c$, whilst leaving approximately one half (49 per cent) of the explosions with $\tau < \tau_c$.

Table II

PERCENTAGES OF EVENTS SEPARATED BY DECISION LEVELS

Decision Level	Q/E Ratio R_c	Condition	Earthquakes	Explosions
$\tau_c = 0.8$	0.17	$\tau < \tau_c$		49
		$\tau \geq \tau_c$	92	
$\omega_c = 4.6$	0.50	$\omega < \omega_c$		49
		$\omega > \omega_c$	76	
$\gamma_c = .05$	0.27	$\gamma < \gamma_c$		34
		$\gamma \geq \gamma_c$	91	
$\tau\omega_c = 3.0$	0.18	$\tau\omega < \tau\omega_c$		34
		$\tau\omega \geq \tau\omega_c$	94	

DISCUSSION

The indiscriminate mixing of events used here with respect to magnitude, focal depth and location has allowed statistics to be compiled for a large population of events. This is important in terms of the reliability of the results; for example, the earthquake percentages in Tables I and II will have statistical variances of about 2 per cent, whereas the statistical variance on Kelly's (1968) results is nearer 10 per cent. The number of explosions used, although about twice as large as used by any previous authors, is still not sufficient to provide reliable estimates of percentages of explosions which satisfy the various criteria. This unreliability becomes critical when using the tails of the explosion probability distributions in game theory extrapolations.

On the other hand, a proper sorting of events with respect to magnitude, depth and location will greatly enhance identification efficiency by allowing criteria in addition to correlogram characteristics to be applied. Reliable focal depths would eliminate all but those earthquakes with $h < 10$ km from the suspicious category; however, determination of h for events shallower than 50 km to within ±10 km is not yet realizable. For the larger magnitude events, say $m > 4.5$ for intracontinental detection (Basham, 1968), and $m > 5.3$ for intercontinental detection (Capon, et al., 1967), the surface-wave magnitude versus body-wave magnitude discriminant can be applied. For purposes of test ban control, a location sort would show many events to be uninteresting; these would include most events with oceanic epicenters and many additional events in politically unimportant locations.

In an evaluation of the efficiency of the correlogram parameter criteria (or any other discrimination criteria), it is essential that explosions and earthquakes from the same general geographical area be compared. This has not been done in the present study because the number of events in any particular area would be too small. However, Yellowknife array data are currently being accumulated in sufficient numbers to test correlogram discrimination criteria separately for each known explosion test site and the surrounding natural seismic region. This will also allow an investigation of the important point of whether the correlogram discrimination efficiency deteriorates at low magnitudes.

It has not yet been determined which method of analyzing the correlogram provides most efficient discrimination. The best of the parameters measured here is the rise-time τ with little or no significant improvement gained from products of two or three parameters or combined single parameter conditions. On the basis of Kelly's (1968) results, the complexity coefficient based on integration of the waveform appears marginally better than the single or product parameter discriminants. An evaluation of the relative merits of these methods will be made from current Yellowknife array output, but it appears no correlogram criterion will be devised which will separate more than about 70 per cent of shallow ($h < 50$ km) earthquakes from 90 per cent of explosions.

REFERENCES

Atomic Weapons Research Establishment (Staff), 1965. The detection and recognition of underground explosions. United Kingdom Atomic Energy Authority, special rept.

Basham, P.W., 1968. Geophys. J.R. Astr. Soc., 17, in press.

Capon, J., R.J. Greenfield, and R.T. Lacoss, 1967. Semiannual Technical Summary, Lincoln Laboratory, M.I.T., 30 June, 3.

Carpenter, E.W., 1964. Institute of Science and Technology, University of Michigan, VESIAC Rept. 4410-67-X.

Davies, D., 1968. Seismic methods for monitoring underground explosions, SIPRI, Stockholm.

Douglas, A., 1967. Willow Run Laboratories, University of Michigan, VESIAC Rept. 7885-1-X, 189.

Ericsson, U., 1967. Research Institute of National Defence, Stockholm, Rept. FOA 4 c 4286-20(23).

Ericsson, U., 1968. Research Institute of National Defence, Stockholm, Rept. FOA 4 c 4338-20.

Greenfield, R.J., 1966. Semiannual Technical Summary, Lincoln Laboratory, M.I.T., 30 June, 15.

Kelly, E.J., 1966. Semiannual Technical Summary, Lincoln Laboratory, M.I.T., 31 Dec., 7.

Kelly, E.J., 1968. Technical Note 1968-8, Lincoln Laboratory, M.I.T.

Manchee, E.B., and H. Somers, 1966. Pub. Dom. Obs., 32, 69.

Manchee, E.B., and D.H. Weichert, 1968. Bull. Seismo. Soc. Am., 58, 1359.

Somers, H., and E.B. Manchee, 1966. Geophys. J. R. Astr. Soc., 10, 401.

Weichert, D.H., E.B. Manchee, and K. Whitham, 1967. Geophys. J. R. Astr. Soc., 13, 277.

THE QUEEN'S PRINTER
OTTAWA, 1969