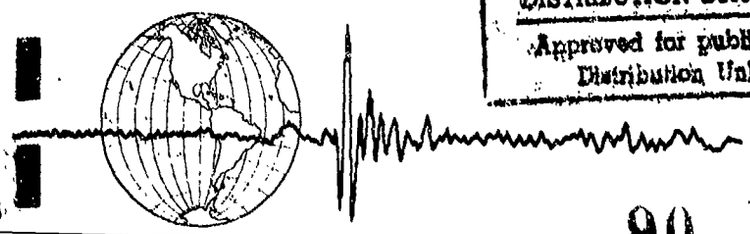
Technical Report C86-05
March 1986

QUARTERLY TECHNICAL REPORTS FOR 1 JULY – 30 SEPTEMBER 1985 AND 1 OCTOBER – 31 DECEMBER 1985

Center Staff

SPONSORED BY:
DEFENSE ADVANCED RESEARCH PROJECTS AGENCY



Center for Seismic Studies
1300 N. 17th Street, Suite 1450
Arlington, Virginia 22209-3871
Telephone: (703) 276-7900

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188
Exp. Date: Jun 30, 1986

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Unlimited
2b DECLASSIFICATION / DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S) SAIC Quarterly Technical Report C86-05	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION Science Applications International Corp. 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Defense Supply Service - Washington
6c. ADDRESS (City, State, and ZIP Code) 1710 Goodridge Drive McLean, Virginia 22102	7b. ADDRESS (City, State, and ZIP Code) Room 1D245, The Pentagon Washington, D. C. 20310
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Defense Advanced Research Projects Agency 8b. OFFICE SYMBOL (If applicable) DSO/GSD	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER MDA903-84-C-0020
8c. ADDRESS (City, State, and ZIP Code) 1400 Wilson Boulevard Arlington, Virginia 22209	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
Research and Development at the Center for Seismic Studies (Unclassified)

12. PERSONAL AUTHOR(S) R. Alewine, C. Archambeau, R. Baumstark, K. Bostic, G. Bulin, P. Dysart, J. Frazier, P. Glover, D. Harkrider, L. Huszar, H. Israelson, A. Jurkevics, J. Pulli, (cont.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Quaterly Technical Report	FROM 7-1-85 TO 12-31-85	28 April 1986	108

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Seismology, Monitoring Nuclear Tests, Data Center, SNAP/D, Seismic Data Exchange, Detection (cont.)

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

Evaluation of the Group of Scientific Experts Technical Test (GSETT), and the exploitation of the GSETT database for more general investigations of network assessment continued during this period. These analyses provide material for subsequent United States reports to the GSE in addition to increasing our understanding of seismic network design and capabilities. Research with the associated objectives of providing analysis tools for the Center while improving understanding of discrimination between events at regional distances was initiated, together with work on yield estimation, primarily in support of the DARPA-managed DoD Seismic Review Panel. The report also covers work related to on-line data acquisition and other programming efforts, and on database development.

Keywords:
#12 - M. Tiberio
#18 - Regional Seismic Data, Surface Waves, Yield, NORESS, Seismic Coupling

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Ann U. Kerr	22b. TELEPHONE (Include Area Code) (202) 694-3145 / 22c. OFFICE SYMBOL DSO/GSD

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted
All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

Technical Report C86-05
March 1986

QUARTERLY TECHNICAL REPORTS FOR 1 JULY – 30 SEPTEMBER 1985 AND 1 OCTOBER – 31 DECEMBER 1985

Center Staff

Sponsored by:
DEFENSE ADVANCED RESEARCH PROJECTS AGENCY
Monitored by:
Defense Supply Service – Washington
Under Contract No. MDA 903-84-C-0020

Science Applications International Corp.
1735 Jefferson Davis Highway, Suite 907
Arlington, VA 22202

Table of Contents

1. INTRODUCTION

With the completion of the Group of Scientific Experts Technical Test (GSETT) during the first quarter of CY 1985, and the completion of most of the documentation during the second and early part of the third quarters of CY 1985, the Center for Seismic Studies' scientific and technical staff was able to devote substantially more effort to research. A revised and updated research plan and a companion database development plan were prepared during the August-September time period. These plans were subsequently approved by DARPA on 25 October, although research on several tasks commenced prior to that date as a result of earlier informal concurrence by DARPA. This report summarizes internal research and development by the Center's staff in consonance with the revised plan and additional requests from DARPA during the fourth quarter and part of the third quarter of CY 1985.

Continuing with work initiated during the previous quarter, one important line of research was the evaluation of the GSETT (as contrasted with previous work on documentation of the test), and the exploitation of the GSETT database for more general investigations of network assessment. These analyses, outlined in part two of this report, provide material for subsequent United States reports to the GSE in addition to increasing our understanding of seismic network design and capabilities. Parts three and four of this report outline research with the associated objectives of providing analysis tools for the Center while improving understanding of discrimination between events at regional distances. Part five describes work on yield estimation, primarily associated with work in support of the DARPA-managed DoD Seismic Review Panel. Part six of the report covers work related to on-line data acquisition and other programming efforts, and on database development.

Two of the technical contributions were prepared by visiting scientists at the Center, whose work was supported in whole or in part by this contract.

2. EVALUATION OF THE GSE TECHNICAL TEST (GSETT)

2.1. SIMULATIONS WITH A PROGRAM FOR SEISMIC NETWORK ASSESSMENT

A computer program, SNAP/D, is being used for various studies in seismic network detection capabilities at the Center for Seismic Studies. SNAP/D, the *Seismic Network Assessment Program for Detection*, is a computer model designed by Pacific-Sierra Research Corporation for assessing the capability of a network to detect and locate seismic events. It is intended to replace *NETWORTH*, the computer model that has been widely used in the verification community for evaluating network performance. Among the attributes of the SNAP/D program listed in the manual are the following:

- Propagates up to 10 seismic waves in a single run.
- Calculates attenuation and travel time as functions of focal depth, regional structure, and event type.
- Allows user specification of multiwave network detection criteria.
- Accommodates advanced techniques for determining the uncertainty in focal depth and location.
- Calculates the relative importance of individual waves and stations in hypocenter estimates.

This report presents examples run with this program to compare capabilities of the GSETT and other networks. SNAP/D required several parameters to be specified, including seismic network event detection criterion and station signal detection characteristics.

The term seismic *event* detection refers to the procedure of defining and locating a seismic event on the basis of *signal* detections at individual stations. In most simulation studies carried out in the past, the criterion for seismic *event* detection only requires that a minimum of *four* stations within 100° from the event epicenter detect *P-signals* of the event. The event detection criterion employed by the WASHIDC[1] during GSETT is somewhat more complex, since it includes not only several kinds of signal phases but also deals with both single and array stations. The following criterion was used here to simulate, in a simple manner and as closely as possible, the GSETT criterion: detections of P- or PKP-signals at 4 or more stations with the additional constraint that not all detections

[1] The Center for Seismic Studies functioned as an "International Data Center", designated as WASHIDC, during the GSETT.

be PKP-signals.

This criterion was incorporated in the SNAP/D program by using the distance interval 0 to 100° for P-signals and 100-180° for the PKP-phase. The amplitude distance curve of Veith and Clawson (1972) was used for the distance interval 0-100° and was extended to 180° with a smoothed "NORSAR" curve (Ringdal 1984). The WASHIDC also used the Veith-Clawson curve for defining events and calculating magnitudes in the interval up to 100°. The extension of the curve beyond 100° used here is believed to model wave propagation more appropriately than the procedure actually used for the definition and location of events during the GSETT. During GSETT, data for stations in the interval 100-110° were not used and for data at distances beyond 110°, a constant amplitude correction equal to that of the Veith-Clawson curve at 90° was applied. The standard deviation of the signal amplitudes, often denoted σ_s, which is equal to the standard deviation of station magnitudes for a given event, was assumed to be equal to 0.36 (Veith and Clawson, 1972).

Overall, 71 of the participating GSETT stations were used in the network. Mean values of station noise amplitudes obtained from a revision based on the GSETT data were used. These revised values, which to some extent are based on station amplitude detection thresholds, have been normalized to a minimum signal-to-noise ratio of 1.5 for signal detection.

Values for station *reliabilities* were obtained as the ratio of the number of days for which arrival time data was received at the WASHIDC and the total number of days (61) for the GSETT. Twelve, or about 15%, of the stations have by this definition a reliability of 100% and 33, or almost half, of the stations have 55 or more data days which correspond to reliabilities of 90% or more. The numbers of data days were obtained from a data base that was upgraded since the GSETT was carried out and for some stations the numbers may therefore be too high.

The 90% magnitude detection threshold contours for the GSETT network are shown in *Figure 2-1*. The contours have been drawn by hand from computed threshold values at grid points every 15° in latitude and longitude. The threshold varies between 3.8 and 4.7 around the world. The range for the Northern and Southern Hemispheres is 3.8-4.4 and 4.1-4.7, respectively.

Because of the significant geographical variation of the threshold, a comparison between simulated and empirical estimates must be based on regional data. The geographical distribution of the earthquakes reported by the WASHIDC during the GSETT, which follows the well known pattern of world-wide seismicity, is also indicated in *Figure 2-1*. More than 75% of the events occur in a limited area within which the threshold varies between 3.8 to 4.4. Weighted average thresholds were computed, taking the geographical distribution of the events into account. The GSETT events were grouped according to location, to the nearest point of the 15° latitude and longitude grid net. The number of events at each grid point was used as a weighting factor. Announced and presumed underground nuclear explosions, as well as small-magnitude events in Northern and Central Europe, were not included among these events.

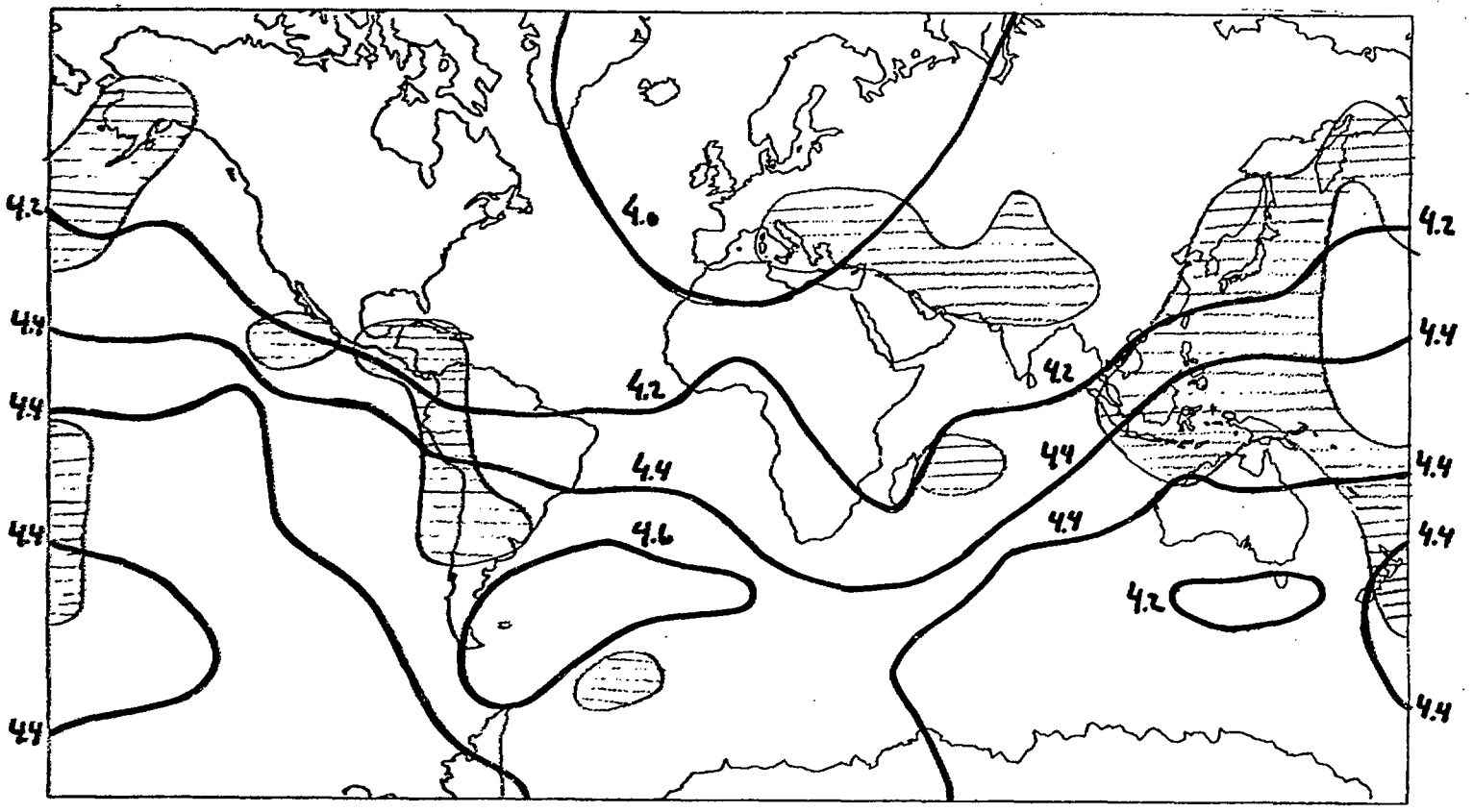


Figure 2-1. **Simulated contours of the 90% m_m-thresholds for the GSETT network. The shaded areas outline the locations of 75% of the earthquakes defined by the GSETT network.**

The following weighted averages were obtained:

Table 2-1		
Region	Threshold	
-	Simulated	Empirical
Northern Hemisphere	4.14	4.15
Southern Hemisphere	4.38	4.55
All events	4.26	4.35

The simulated average values are compared in the table with empirical estimates obtained from the magnitude distribution of the seismic events. The difference between the two kinds of estimates varies between 0.01 and 0.17, depending on region. The simulated thresholds are slightly lower than the empirical ones. Use of the reliabilities in Table 2-1 gave thresholds between 0.03 and 0.07 magnitude units larger than using a 100% reliability for all stations.

The detection capability of the individual stations of the GSETT network varies significantly. Most of the events defined and located by the WASHIDC were based on observations from about one-third of the stations. The detection thresholds for a network of 26 stations with noise values below 6.6 nm was compared with those for the entire network. The difference varies between 0.0 and 0.2 and the mean difference is 0.07 for the world. The 26 stations in the network may not constitute the optimal subset of GSETT stations. They were simply chosen on the basis of their noise level to illustrate that the best third of the 71 GSETT stations, in terms of noise level, perform almost as well as the entire network.

Simulated and empirical estimates of the detection thresholds of a subset of 115 stations reporting to the ISC were obtained earlier (Ringdal, 1984). The capability of this 115-station ISC network represents the level currently available from a public international network based on *teleseismic* signal detections. The simulated results for the ISC network were reproduced using an extended version of the SNAP/D code and parameter settings given by Ringdal (1984). The computational parameters for these simulated estimates were different from the ones used above for the GSETT network. This is particularly true for the amplitude-distance curve.

To make a direct comparison with the GSETT stations, estimates of the detection thresholds of the 115-station ISC network were also obtained. These estimates used the same event detection criterion and parameters for station signal detection as those used for the GSETT stations above, except for the station noise amplitudes. The difference between the two networks is not larger than 0.2 magnitude units. The threshold for the ISC network is, on average, 0.08 magnitude unit lower than the 71-station GSETT network and 0.19 lower than the 26-station GSETT network.

Hans Israelsson

2.2. MINIMUM SIGNAL-TO-NOISE RATIO FOR SIGNAL DETECTION ESTIMATED FROM GSETT DATA

Models for simulating the detection capability of a seismic network often use a simple relation to describe a signal detection to be declared at an individual station. The signal detection criterion is sometimes written as:

$$A > SNR_{\min} N,$$

where A is the signal amplitude, $SNR_{\min}$ is the detection threshold or minimum signal-to-noise ratio at which a signal can be detected, and N is the seismic noise level or amplitude.

It is assumed here that recorded noise and signal can be treated as a narrow bandpass system (Thomas, 1969) represented by the teleseismic frequency band. The noise, $n(t)$, is then assumed to be a normal stochastic process, which can be written as a function of time, t, in the form:

$$n(t) = E_n(t)\cos\left(2\pi f_c + \phi_n(t)\right).$$

Here, E_n and ϕ_n are the envelope and phase functions, respectively, and f_c is the center frequency of a narrow band between 1 and 2 Hz. An added sinusoidal signal with amplitude A results in the process:

$$s(t) = n(t) + A\cos(2\pi f_c + \phi_s).$$

Here, A is a constant and corresponds to the true signal amplitude and ϕ_s is an arbitrary phase angle related to the signal.

The short period noise measurements *(NSZ)* during the GSETT consisted, according to the test procedures, of the maximum trace amplitude within 30 s before signal onset at frequencies close to that of the signal.

If the power of the noise is denoted by σ_n, that is, the expected value $E(n^2(t)) = \sigma_n^2$ then the noise envelope follows a Rayleigh distribution with frequency function:

$$f_{E_n}(x) = \frac{x}{\sigma_n^2} exp(-x^2/2\sigma_n^2).$$

The mean value of this distribution is $\sigma_n\sqrt{\pi/2}$ and the maximum of the frequency distribution function occurs for $x = \sigma_n$. These values correspond, however, not to the measured noise amplitude, NSZ. In order to relate the measured NSZ values to the noise process here some simplifying assumptions are necessary. First, the noise envelopes are

assumed to be stationary over a time interval of T seconds. The maximum value of the noise amplitudes is also assumed to be represented by the maximum of $N=30/T$ independent envelope values which follow the same Rayleigh distribution introduced above. The median value of the maximum of the N independent envelope values becomes the solution to:

$$1-exp(-x^2/2\sigma_n^2)=\frac{1}{2^N},$$

which gives a value representing the noise as the *NSZ*-amplitude in a 30 s time interval, $\max(E_n(N))$. For $T=1$ the NSZ-value $\max(E_n(30)) \approx 2.75\,\sigma_n$ which corresponds to about the 95% probability of the envelope distribution function; in other words, 95% of the envelopes are smaller than this value.

False alarm probabilities and the number of false alarms per day can be roughly calculated from the model above. The probability of a false alarm can be written as:

$$P(FA)=P(E_n > SNR_{min}\max(E_n(N)))$$

and the average number of false alarms per day becomes:

$$N(FA)=2N\cdot 60\cdot 24e^{-x^2/2\sigma_n^2}$$

with

$$x=SNR_{min}\max(E_n(N)).$$

The number of false alarms per day computed from the above formulas for different T and SNR_{min}-values are listed in Table 2-2. The number of false alarms for the GSETT stations is not known, but it is reasonable to assume that it is quite small, of the order of

Table 2-2

T	SNR		
(s)	1.5	1.75	2.0
1	17.5	0.8	0.02
2	40.6	3.3	0.17
3	65.7	7.3	0.58
5	118.2	19.5	2.44

one or less. On this assumption, the table above suggests that the time interval over which the noise process is stationary is also small, approximately one second if the detection threshold is close to 1.5. The values in the table suggest that a threshold of 1.5 would be difficult to operate on with a small number of false alarms, but that a value of 1.75 would give an acceptable number of false alarms. A low threshold value implies a short period of time over which the noise envelopes are stationary.

The signal amplitudes during GSETT were measured as the maximum trace excursion in four time windows (*M1X–M4X* in GSE notation) after signal onset. The amplitude measurements are represented here by the envelope of the signal model E_R with amplitude A which has a frequency distribution equal to:

$$f_{E_R}(x) = \frac{x}{\sigma_n^2} exp(\frac{-(x^2+A^2)}{2\sigma_n^2}) I_0(\frac{Ax}{\sigma_n^2})$$

where I_0 is the modified Bessel function of the first kind and order zero (Thomas, 1969). The detection probability for a given minimum detection threshold, SNR_{min} and signal amplitude, A, can be related to the number of false alarms:

$$P(Detection|\, A, SNR_{min}) = P(E_R > SNR_{min} \max E_n) = DP$$

$$n(FA) = 30 \cdot 60 \cdot 24 \cdot P(E_n > SNR_{min} \max E_n).$$

The number of false alarms per day is obtained from the formula above with $T = 2s$. By eliminating the parameter $SNR_{min}\max(E_n)$, one obtains the so called receiver operating characteristics, or ROC-curve, in terms of the cumulative distribution functions, F, for E_R and E_n :

$$n(FA) = 30 \cdot 60 \cdot 24 \cdot (1 - F_{E_n}(F_{E_R}^{-1}(1 - DP))$$

The ROC curves in *Figure 2-2* show the trade off between detection probability and number of false alarms. The curves show that if a false alarm rate of 1 per day is acceptable, the detection probability becomes about 25% for signals with amplitudes as small as 1.75 times the NSZ values. Station detection probabilities as low as 25% are relevant in the context of network *event* detection, since at least four stations are usually required to detect signals from an event with a 90% probability. The 90% detection probability is thus not required at each individual station, for which it is usually sufficient to detect a signal with much lower probability. This is particularly true if the number of stations is large.

Amplitude measurements of noise (*NSZ*) and signals (*MjX*) made at the same frequency can be used to estimate upper bounds for the threshold parameter SNR_{min}. This means that the actual SNR_{min} should be equal to or smaller than the minimum of the

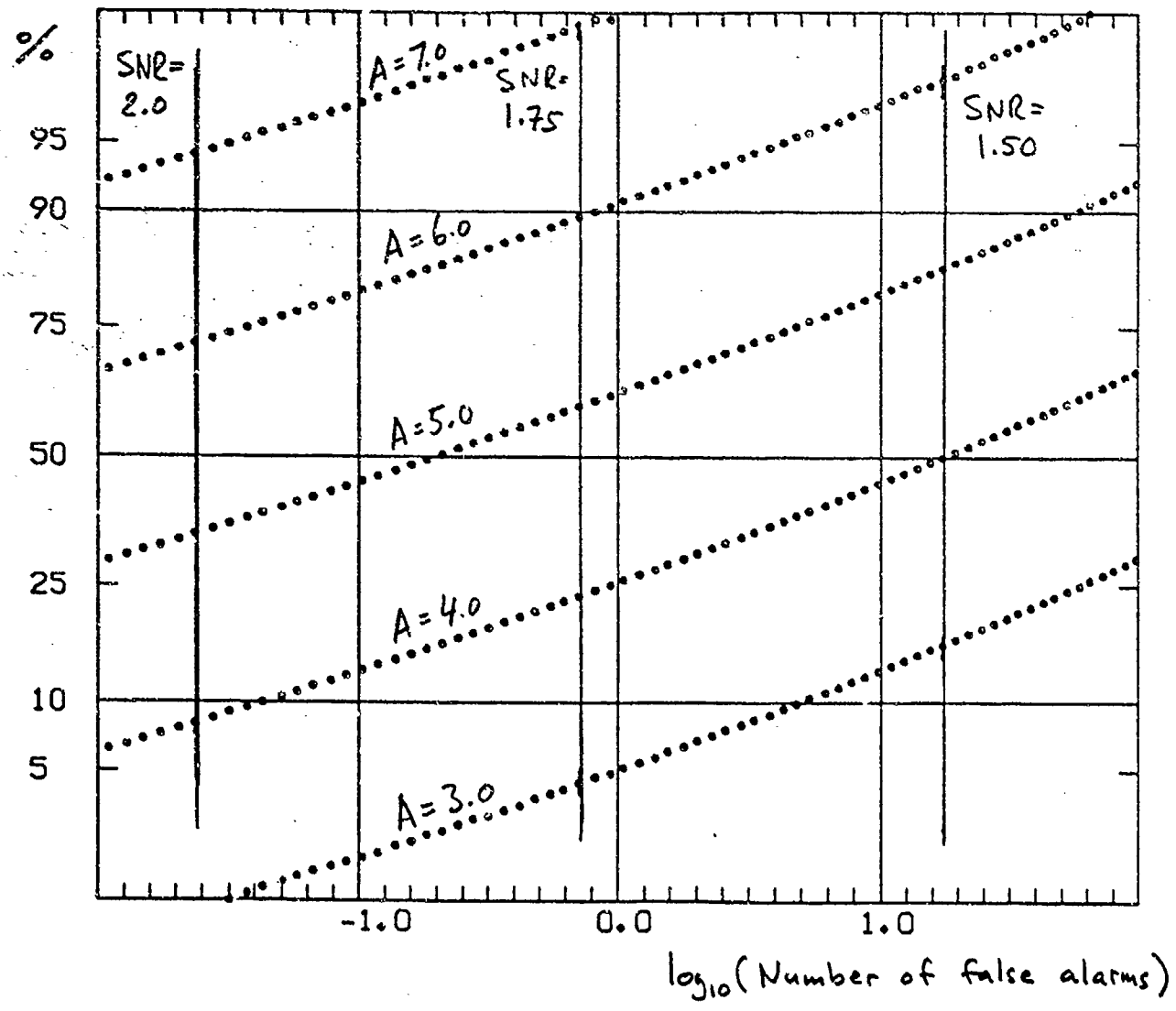


Figure 2-2. Receiver operating characteristics or ROC-curves showing the detection probability as a fucntion of number of false alarms for different values of the true signal amplitude, A (in σ_n-units) and the threshold, SNR_{min}.

amplitude ratio *MjX/NSZ*. For a large number of observations this minimum would be close to the threshold. Such upper bound estimates obtained for the United States stations contributing to the GSETT have a median value of 2.2.

The ratio of the amplitude detection threshold and mean noise amplitude in the teleseismic frequency band is also assumed to be approximating the threshold value as defined here. The empirical cumulative distribution function for ratios at 14 of the GSETT stations is shown in *Figure 2-3*. The median value is 2.2, which is in agreement with the value obtained independently from *MjX/NSZ* amplitude ratios.

The data from the GSETT defined in this way suggest that a SNR_{min} of approximately 2 is more likely to be representative of the average station than a SNR_{min} of 1.5. This is supported by both kinds of data, but there is variation among stations. The distribution in *Figure 2-3* is close to a normal distribution and shows comparatively small variation. However, even if the mean value is around 2.0, there are stations with clearly lower as well as higher thresholds.

Hans Israelsson

2.3. DETECTION OF UNDERGROUND NUCLEAR EXPLOSIONS BY THE GSETT SEISMOLOGICAL STATION NETWORK

At least 12 announced and presumed underground nuclear explosions were reported for the period of the GSETT, 15 October to 14 December 1984. The locations and times of the events are given in Table 2-3. The listing of events in Table 2-3 should not be construed as complete with regard to this particular kind of source for the GSETT period. It merely includes events that were actually reported during the GSETT.

Measured amplitude and period values (*M*1*X*-parameter) recorded at the GSETT stations from the events have been used in the analysis here. Table 2-3 lists body wave magnitudes m_b estimated by the USGS and the WASHIDC during the GSETT, and revised maximum likelihood values $m_b(max)$ subsequently computed by the WASHIDC.

Figure 2-4 shows the average period for each event as a function of magnitude $m_b(max)$. The expected increase of the period with increasing magnitude due to source scaling can be seen for events in E. Kazakh, S. Nevada, and the Tuamotu Archipelago. However, it is less pronounced than the typical values sometimes obtained by the derivative $\frac{dT}{dm} \approx 0.2$ (Marshall *et al.*, 1979). The difference may in part be due to the maximum likelihood technique giving smaller values for small magnitude events than a straight forward average procedure. The mean values for the different events in *Figure 2-4* are based on different stations. Data is available only for a few stations for each of the sites, for small as well as large events. A comparison of data for stations that reported amplitudes for both small and large events showed an even less pronounced scaling effect. Only 5 of 15 observations showed an increase in period with increasing

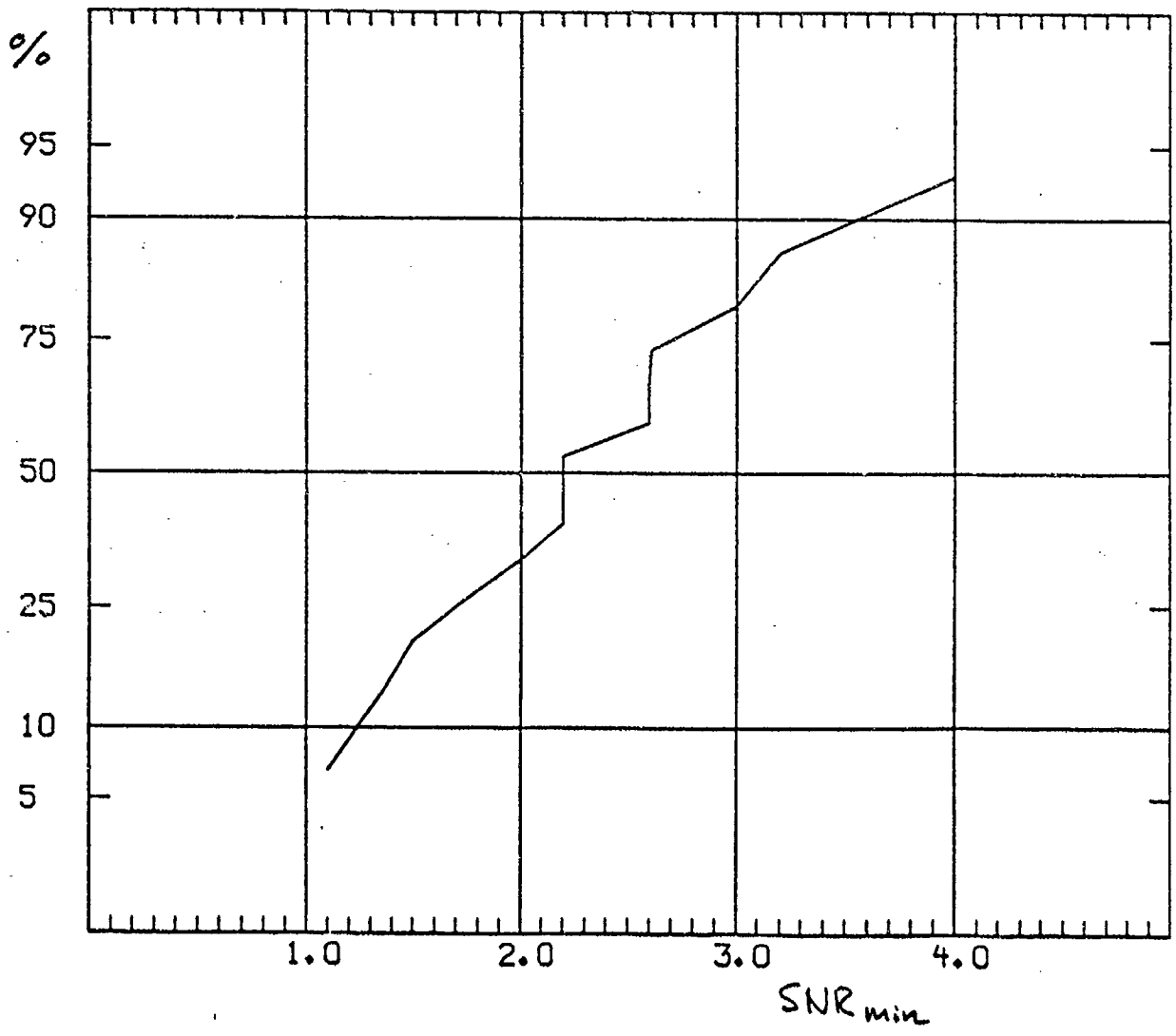


Figure 2-3. Empirical distribution function of SNR_{min}-values for 14 GSETT stations.

Table 2-3								
ANNOUNCED AND PRESUMED UNDERGROUND NUCLEAR EXPLOSIONS*								
Date	Time	Lat	Long	mb			Ms	
				USGS	WASHIDC	max	USGS	max
84/10/18	04:57:05.7	49.8N	78.1E	4.5	4.6	4.07		
84/10/25	06:29:57.7	73.4N	55.0E	5.9	5.4	5.66	4.7	
84/10/27	01:50:10.6	49.9N	78.8E	6.2	6.1	6.07	4.4	
84/10/27	05:59:57.1	46.8N	48.1E	5.0	4.7	4.95		
84/10/27	06:04:56.7	46.8N	48.1E	5.0	4.4	4.91		
84/10/27	17:16:00.0	21.1S	138.9W		4.1	4.21		
84/11/02	20:44:58.5	21.1S	138.9W	5.7	5.3	5.45		
84/11/10	16:40:00.0	37.0N	116.0W	4.5	3.8	4.05		
84/11/23	03:55:04.8	49.9N	78.1E	4.7	4.5	4.28		
84/12/02	03:19:00.0	49.9N	79.1E	5.8	5.2	5.77	4.6	
84/12/06	17:29:00.0	21.1S	138.9W	5.6	5.3	5.38		
84/12/09	19:40:00.0	37.3N	116.5W	5.5	5.4	5.29	4.2	4.35

* Times and locations are those given in the Monthly Bulletins of the USGS except for the event on 27 October which was not included in the USGS Monthly listing. The data for this event are those obtained at the WASHIDC during the GSETT.

magnitude.

The variation in station magnitude bias is sometimes described by a signal amplitude standard deviation which is often assumed to be 0.36 magnitude units (Veith and Clawson, 1972). The standard deviation of the station magnitudes for a given event, varies significantly, however, and is correlated with the mean wave period. In the scatter diagram in *Figure 2-5* this standard deviation ranges between 0.2 to almost 0.5 for the GSETT explosions.

The following source corrections can be obtained from earlier published results (Marshall *et al.*, 1979):

Table 2-4			
Site	Period(s)	Pn-velocity	Correction
E. Kazakh	0.79	8.3	0.00
Novaya Zemlya	0.68		
SW.Russia	0.44	8.15	0.22
Tuamotu A.	1.05		
S.Nevada	1.07	7.8	0.23

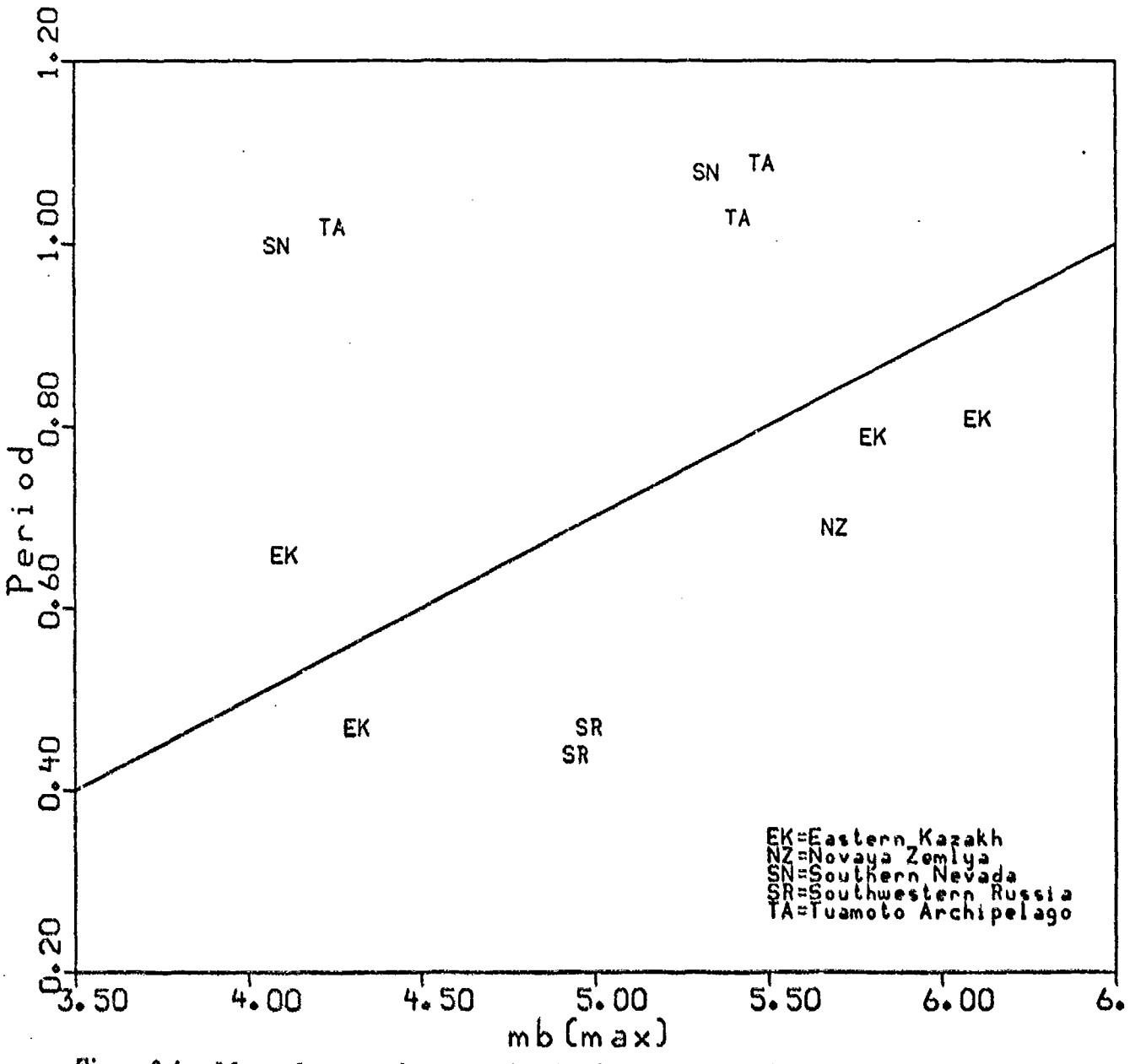


Figure 2-4. Mean of measured wave period for the events plotted agianst the maximum liklihood estimate of m_b. The straight line has a slope of $\frac{dT}{dm}$=0.2.

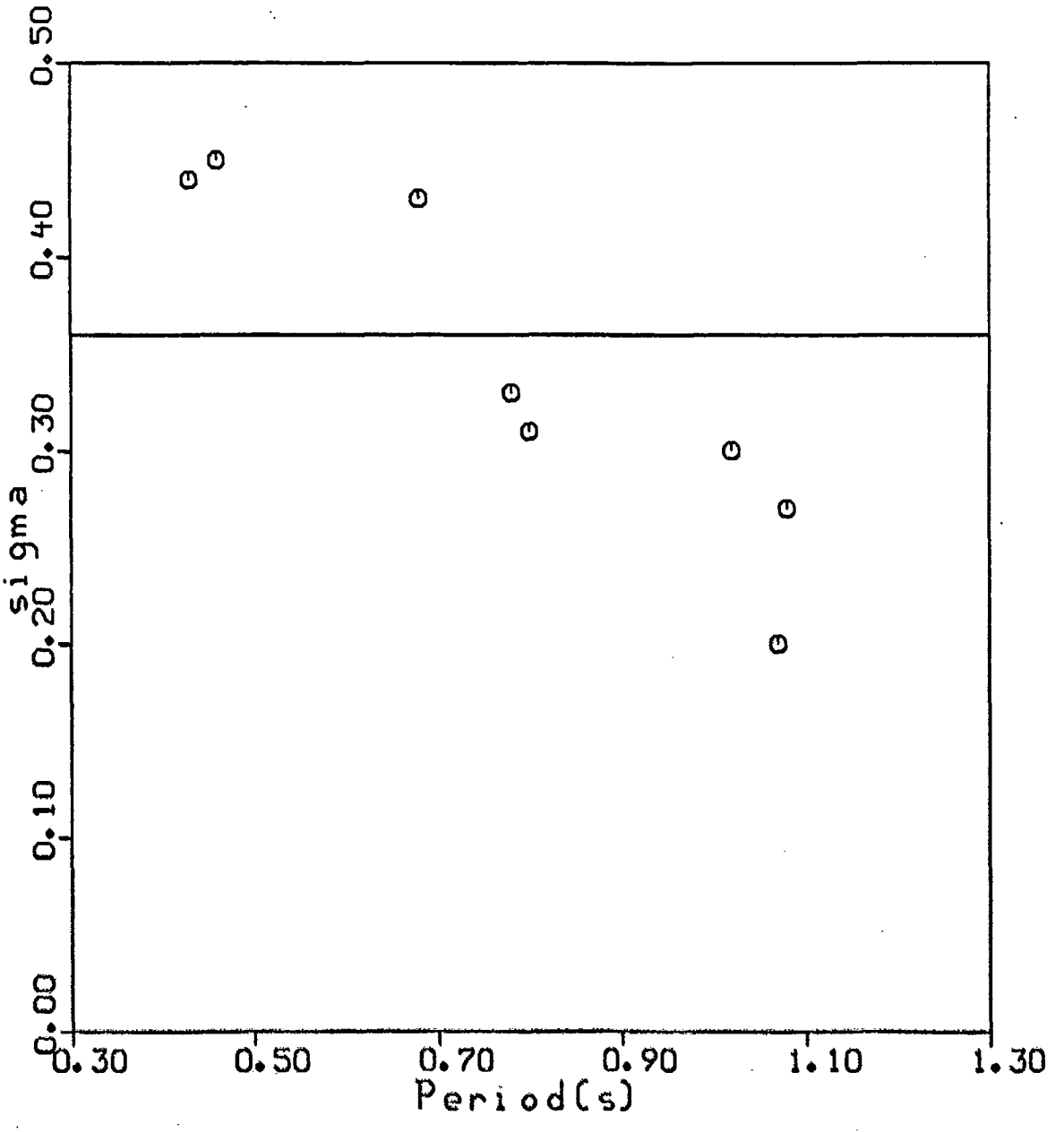


Figure 2-5. Standard deviation of station magnitudes for the events as a function of measured mean wave period. The horizontal line correspond to the often used value of 0.36 of the standard deviation of the standard deviation.

The difference between the $m_b(max)$ of the large E. Kazakh and the S. Nevada events is about 0.8 magnitude units, as shown in Table 2-4. If the source bias corrections above were applied, the difference would still be 0.5 to 0.6. This difference is quite large considering the reported M_s values for these events (Table 2-3).

The amplitude/period data from explosions during GSETT illustrate that some parameters of significance for the network detection capability deviate in a systematic way from a standard, average earth model. For instance, the mean signal period observed from explosions differs among the source sites (different instrumentation at the particular stations reporting events from each test site also has an effect on the recorded amplitudes and detectability, making it difficult to isolate a pure source effect). In approximate terms, the *characteristic* dominant period for S. Nevada and Tuamotu Archipelago is around 1 s, whereas it is around 0.5 s for E. Kazakh, SW Russia and Novaya Zemlya.

The station noise level at 1 and 2 Hz differs significantly at many stations and in most cases it is lower at 2 Hz. Noise data at both periods were available for 39 stations at the WASHIDC. About two-thirds of these stations have lower noise thresholds at 2 than at 1 Hz.

As illustrated in *Figure 2-5* the variation of signal amplitudes is also a function of period and it is assumed here that the standard deviations for sites with 1.0 s and 0.5 s periods are 0.2 and 0.45 respectively. This assumption is based on the scatter of station magnitudes of large events, and it may not be applicable to smaller events.

The detection capability for the 39 station network was calculated for the four cases of noise frequency and signal standard deviations. Differences in the 90% detection thresholds between the extreme cases (1 Hz, σ_s=0.20 and 2 Hz, σ_s=0.45) vary with geographical location, and are around 0.2 or more. The lowering of the thresholds, due to changes in noise amplitudes by going from 1 to 2 Hz, is about the same as by increasing the σ_s-value from 0.2 to 0.45. Magnitude detection thresholds at 90% probability for the five event sites are compared below with the smallest observed magnitudes during the GSETT.

The smallest magnitude of the events for three of the sites in the table below was obtained by using σ_s-values of 0.2 (for S. Nevada and Tuamotu Arch.) and 0.45 (E. Kazakh) in the maximum likelihood estimation, instead of the value 0.36. The maximum likelihood estimates are sensitive to the standard deviation, σ_s only for smaller events,

Table 2-5

Noise Frequency (Hz)	Signal Std. Deviation
1	0.20
1	0.36
2	0.36
2	0.45

Table 2-6		
Site -	Threshold at 90%	Smallest Magnitude
E.Kazakh	3.89	3.95
N.Zemlya	3.69	
SW.Russia	3.74	
S.Nevada	4.24	4.26
Tuamotu Arch.	4.28	4.41

i.e., below $m_b < 5.0$.

As noted above, the magnitude source bias is significantly different (at least 0.2 units) between sites like E. Kazakh and S. Nevada and, if it were to be included here, the differences in magnitude thresholds would be even more pronounced.

In conclusion, the data recorded by GSETT stations from announced and presumed underground nuclear explosions illustrate the frequency dependence of network magnitude detection thresholds. The variation of the detection threshold with frequency is caused both by variation of the station noise amplitude, receiver and source upper mantle attenuation. Some variation and scatter in the network magnitude data is also caused by differences in general among the GSETT stations with regard to equipment and procedures for detection and analysis.

Hans Israelsson

2.4. UNASSOCIATED SIGNAL DETECTIONS AT GSETT STATIONS

Assessments of capabilities of seismic station networks are often focused on estimates of magnitude detection thresholds. However, global station networks with low magnitude thresholds produce a large number of individual station detections that cannot be associated with any seismic event defined and located by the network. Association of detected signals is one important function for a workable global test ban monitoring system.

In order to conform with the standard "Networth" approach for detection capability computations, it is assumed that station j, in a network of n stations, detects a *signal* from an event with magnitude m, with probability p_j:

$$p_j(m) = \Phi\left(\frac{m - T_j}{\sigma_j}\right)$$

where Φ is the standard normal distribution function and T_j is the station magnitude detection threshold (50%) with its associated standard deviation σ_j.

Another assumption is that a seismic *event* is detected and located by the network if more than n_{min} stations detect *signals* from the event.

The ratio of associated-to-detected arrivals at station j can be calculated from the probability that it detects signals from an event with magnitude m and that at least n_{min} other stations of the network also detect the event. If the number of detecting stations, apart from station j, is denoted $\xi_{\neq j}(m)$, then the probability that a detected signal at station j is associated can be written as:

$$a_j(m)=p_j(m)P(\xi_{\neq j}(m)\geq n_{min})$$

If the magnitudes of the natural seismicity have an exponential distribution, *i.e.*, $\approx e^{\alpha-\beta m}$ for $m_{min}\leq m\leq m_{max}$ then the total number of associated detections becomes:

$$N_{assoc}(j)=\int_{m_{min}}^{m_{max}} a_j(m)e^{\alpha-\beta m}dm$$

and the total number of detected signals:

$$N_{detect}(j)=\int_{m_{min}}^{m_{max}} p_j(m)e^{\alpha-\beta m}dm$$

and the ratio of associated to detected signals:

$$R(j)=\frac{N_{assoc}}{N_{detect}}.$$

The dependence of α disappears in the ratio $R(j)$, but if associated detections are considered from many different seismic regions, n_r, with approximately the same β-value but different α-values, the formula for the ratio has to be extended to:

$$R(j)=\sum_{i=1}^{n_r} N_{assoc}(ij)/\sum_{i=1}^{n_r} N_{detect}(ij).$$

The number of associated signals and the ratio of associated-to-detected signals for the network of stations can be calculated from a combination of the formulas above for individual stations. In some instances, it becomes simpler to calculate these values

independently, as outlined below.

The number of detecting stations for an event with magnitude m is a stochastic variable, ξ_m. The expected number of detected signals can then be written as:

$$E\xi_m = \sum_{k=1}^{n} kP(\xi_m = k) = \sum_{j=1}^{n} p_j(m)$$

It is assumed that an event is detected if signals from the event are detected at more than n_{min} of the stations. Also assumed is that unassociated detections occur only in cases where n_{min} or less stations have detected signals from the event. The expected number of unassociated signals for an event is also a stochastic variable ζ_m and its mean value can be written as:

$$E\zeta_m = \sum_{j=1}^{n_{min}} jP(\xi_m = j).$$

The ratio for the number of unassociated and total number of detected signals can be written as:

$$r_m = \frac{E(\zeta_m)}{E(\xi_m)},$$

and is a function of magnitude, m. If the number of events as a function of magnitude is proportional to $e^{\alpha-\beta m}$, then the number of unassociated and total number of detected phases can be written as:

$$N_{unass} = \int E(\zeta_m) e^{\alpha-\beta m} dm$$

$$N_{detect} = \int E(\xi_m) e^{\alpha-\beta m} dm$$

and the ratio of the number of unassociated and total number of detected signals becomes:

$$R = \frac{N_{unass}}{N_{total}}.$$

For regional variation of the parameters α and β, this ratio has to be extended as for the ratio for individual stations above.

Data sets from the Experimental International Data Centers at Stockholm (*STOCIDC*) and at Washington (*WASHIDC*) have been used for the analysis of signal association, which is limited to teleseismic first arrivals. The STOCIDC data, which covers the main phase of the test, *i.e.*, 47 days, was presented in a working paper at the *GSE*. The WASHIDC data set, which covers the entire 61-day period, was compiled immediately after the GSETT. The data available at the two IDC's during GSETT also differed. The differences between them is used to indicate the uncertainty in the association percentages.

The differences between the STOCIDC and WASHIDC data sets are shown in *Figure 2-6* which compares associated detections (normalized to percent). In *Figure 2-6* two stations, EKA and GRA1, have significantly different association percentages. The difference for GRA1 may be due to data reported for other subarrays instead of this particular subarray. The standard deviation of the difference in association percentages is about 10%, which gives an idea of the degree of uncertainty in these types of numbers.

With the formulas above, one can compute an overall value for the unassociation percentage. The geographical effect of seismicity is taken into account by weighting each region (*i.e.*, cell in the 15 by 15 ° grid net used) with the number of earthquakes defined for that cell by the GSETT. The resulting values are presented as association rather than unassociation percentages and are compared in *Figure 2-7*, with the observed values for the GSETT, represented by the maximum per cent of the STOCIDC and WASHIDC data sets.

The calculations are limited to 20 stations. For 13 of the stations, the calculated and observed values show reasonable agreement, considering the uncertainty in the calculations, as illustrated by the 10% standard deviation of the differences between STOCIDC and WASHIDC. For seven stations, the calculated percentages are clearly higher than the observed ones: EKA, LAC, MAT, MAW, PRU, RSNY, and RSON. This can partly be due to the use of excessively high noise values based on amplitude detection thresholds in the calculations for these stations. A rerun using the actual noise values did reduce the percentages somewhat, but not enough to remove the outlying character of these stations, at least for EKA, MAT, and RSNY. Local signals among the teleseismic ones could explain the persistent discrepancy for these stations. A significant reduction in association percent is obtained for both MAW and LAC by using their actual noise values rather than the amplitude thresholds (converted into equivalent noise amplitudes). The actual noise values for these stations are, however, hardly consistent with their numbers of detected signals.

Hans Israelsson

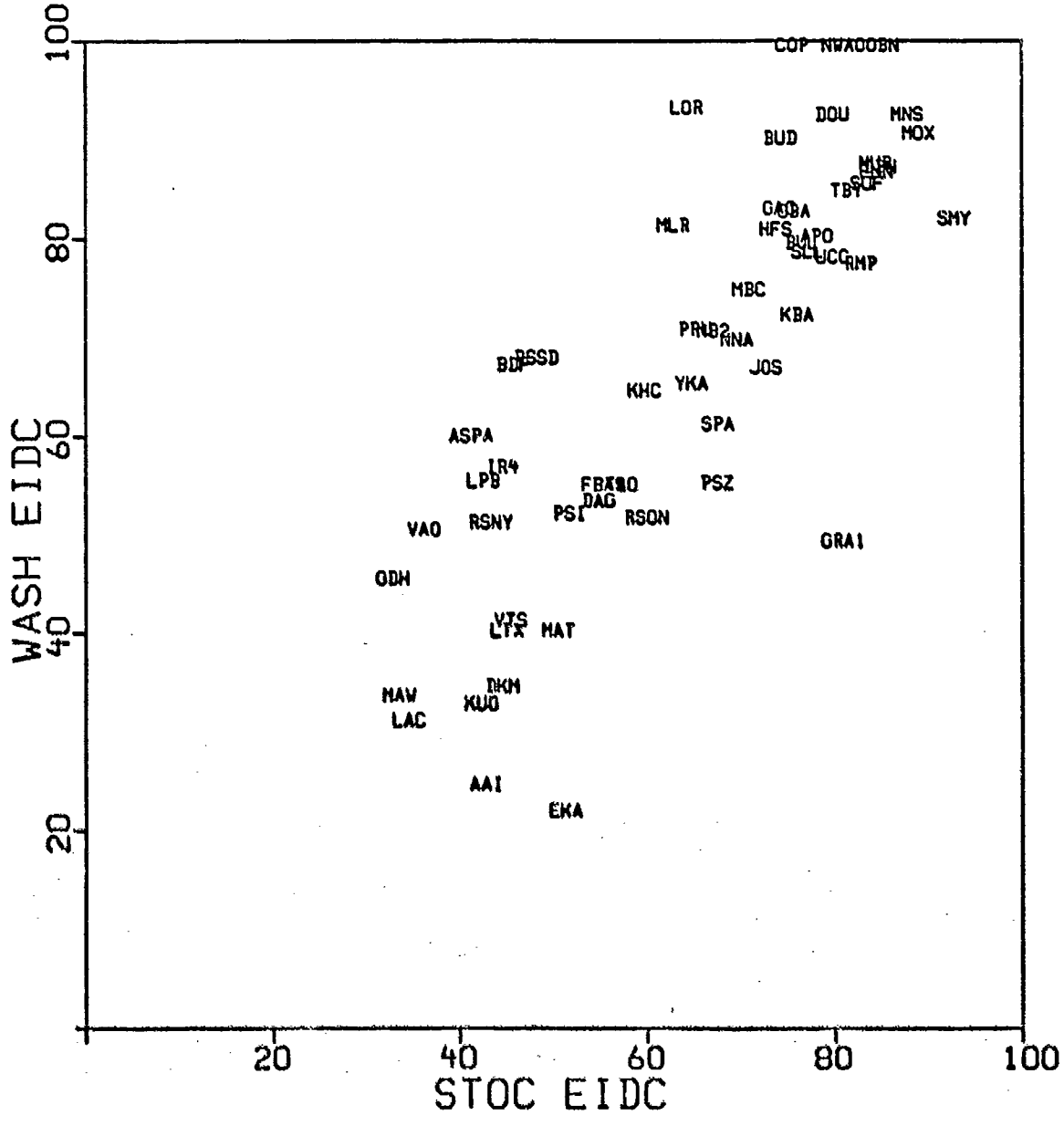


Figure 2-6. | Comparison of association percentages of teleseismic detections at the GSETT stations as obtained from the data sets at STOCIDC and WASHIDC studied here.

2.5. TRANSMISSION EFFICIENCIES OF THE WMO/GTS

This study compares statistics for transmission of seismological messages via the GTS during the GSETT with those of the WMO monitoring for transmission of meteorological messages in 1984. The WMO monitoring occurred during the period 1 October to 15 October, just prior to the start of the GSETT.

In the 1984 monitoring period, the WMO checked the availability of weather messages over globally distributed WMO/GTS circuits. Approximately 660 messages were sent daily from the test participants. The message arrivals were monitored at several regional centers. Detailed data from this worldwide test were made available by the WMO headquarters in Geneva. Data from 38 national reporting centers were chosen to compare with the circuits used in the GSETT. Message arrival statistics from these 38 sites were collected at the three WMO regional centers (Bracknell, Moscow, and Washington), corresponding closely to the locations of the EIDC's in the GSETT.

Preliminary transmission efficiency percentages for *meteorological* messages collected during the WMO monitoring test are listed in Table 2-7, which include data for transmission channels from countries participating in the GSETT to the three WMO regional centers: Bracknell, Moscow, and Washington. The EIDC in Stockholm is not directly linked to a WMO Center on the Main Telecommunication Network (MTN) and Bracknell was chosen here as a closely located WMO Center.

For each country, Table 2-7 also gives the nearest WMO center (Bracknell, Moscow, or Washington) and indicates whether the connection to this center is direct. A direct connection has no node between the national facility and the nearest WMO center. Information on WMO centers and connections in Table 2-7 was obtained from the chart of the implementation of the GTS, included in the Conference Room Paper 134/Rev.1. It is seen from the table that the average network transmission efficiency ranges from 78% to 87%, with a regional center average of 82.3%.

The percentages in Table 2-7 show clear regional variation. Perhaps the most conspicuous regional effect is the difference between European and non-European countries. In Table 2-7, 51 out of 54 communication channels between European countries and the three regional WMO centers have values above 95%, and 37 channels have percentages at 99% or above. This is in sharp contrast to the 60 channels from non-European countries, which exceed 95% only in 6 cases.

The difference in transmission efficiency between European and non-European countries could be partly due to most European countries having a direct circuit to one or more of the WMO regional centers. There are direct circuits only in three cases for non-European countries. Another factor accounting for the difference between European and non-European countries could be the lack of interest in meteorological data outside Europe, where two of the WMO centers (Bracknell and Moscow) considered here are located. The data outside Europe could be less relevant at these centers, and may therefore not be accumulated to the same extent.

The WMO monitoring results are compared here with the transmission of *seismological* messages during the GSETT from the national seismological centers to the EIDCs in Moscow, Stockholm, and Washington.

TABLE 2-7 WMO MONITORING TRANSMISSION PERCENTAGES										
Region	Country	Channel		Percentages			Regional Averages			
-	-	-	-	-	-	-	-	-	-	Center
-	-	Nearest Center	Type	Mosc	Wash	Brack	Mosc	Wash	Brack	Average
Africa	Egypt*	Brack(Mosc)	(direct)	54.2	57.9	74.9	61.0	74.4	83.5	73.0
	Kenya	Brack		69.7	86.8	97.3				
	Zambia*	Brack		43.7	64.3	68.8				
	Zimbabwe	Brack		76.3	88.4	92.8				
Antarctica	Antarctica*	Brack		61.1	63.9	70.6	61.1	63.9	70.6	65.2
Asia	India*	Mosc	direct	68.9	71.5	81.0	61.3	79.3	77.1	72.6
	Indonesia*	Wash		23.3	52.9	48.7				
	Japan*	Wash	direct	96.6	100.0	94.7				
	Pakistan	Wash		61.3	77.6	70.7				
	Thailand	Wash		56.4	94.5	90.4				
Europe	Austria*	Mosc		100.0	100.0	100.0	99.8	98.8	97.2	98.6
	Belgium*	Brack	direct	100.0	96.7	100.0				
	Bulgaria*	Mosc	direct	100.0	98.4	82.8				
	Czechoslo*	Mosc	direct	100.0	100.0	87.8				
	Denmark*	Brack	direct	100.0	96.3	100.0				
	Finland*	Mosc	direct	100.0	100.0	100.0				
	France	Brack	direct	100.0	99.6	99.4				
	GDR*	Mosc		100.0	100.0	96.7				
	Germany,F.R.*	Brack	direct	100.0	96.6	100.0				
	Hungary*	Mosc		100.0	100.0	98.9				
	Ireland	Brack	direct	100.0	98.4	100.0				
	Italy*	Brack		97.8	97.8	97.0				
	Netherlands*	Brack	direct	100.0	100.0	99.2				
	Norway*	Brack	direct	100.0	98.0	100.0				
	Romania	Mosc		99.2	99.6	94.6				
	Sweden*	Mosc	direct	99.9	99.3	100.0				
	UK*	Brack	direct	100.0	100.0	97.9				
	USSR*	Mosc	direct	99.9	98.5	95.4				
N.America	Canada*	Wash	direct	81.3	95.0	81.2	84.0	92.5	84.9	87.1
	USA*	Wash	direct	86.6	89.7	88.6				
S.America	Argentina	Wash	direct	53.7	76.9	39.3	40.1	62.7	27.5	43.4
	Bolivia	Wash		27.0	66.6	15.2				
	Brazil	Wash		42.9	60.3	32.0				
	Colombia	Wash		27.6	43.1	28.5				
	Peru*	Wash		48.9	64.7	22.6				
S.Pacific	Australia*	Wash		81.1	87.6	87.6	63.0	90.1	89.4	80.8
	Fr.Polynesia	Wash		69.5	97.9	96.5				
	New Zealand*	Wash		58.3	84.7	84.1				
Average	All Countries			78.0	87.0	82.0				82.3
	with * only			83.9	89.0	86.9				86.6

Countries marked with * in the table were listed as WMO/GTS users in the procedures for the GSETT (conference Room Paper 134/Rev.1) prior to the test.

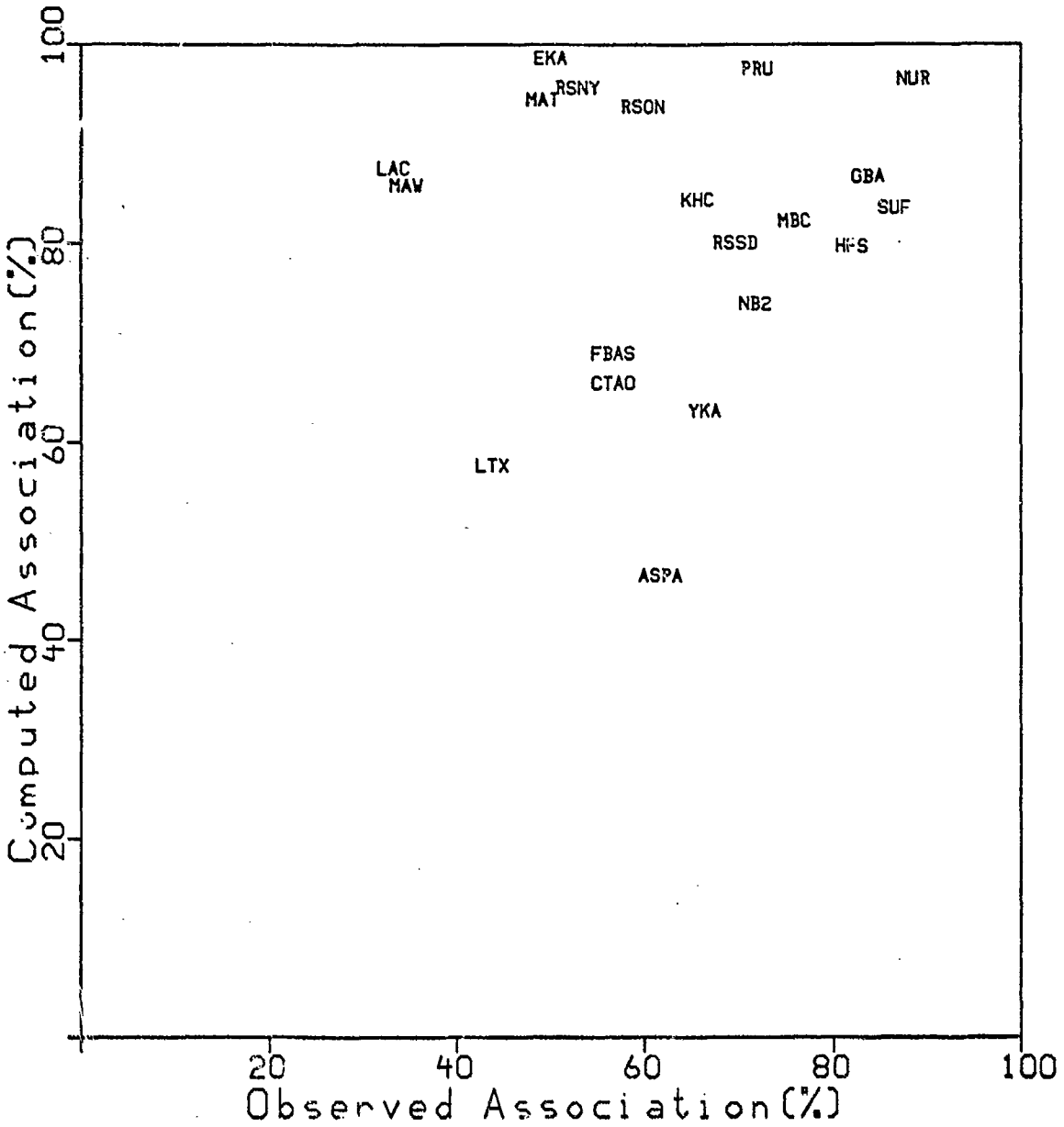


Figure 2-7. | Comparison between calculated and observed association percentages for 20 GSETT stations. The observed values are represented by the maximum value of the two data sets from STOCIDC and WASHIDC.

Transmission percentages for GSETT have been summarized in Table 2-8 and include values for about 120 transmission channels to the three EIDC's. The data, which cover the period 27 October 1984 to 14 January 1985, have been obtained as the ratio of the number of unique Level I data messages received and the number of unique messages sent. Values representing transmission with and without re-transmission are listed in the table and the countries are grouped by geographical region. In the table, the average transmission efficiency ranges from 56.8% to 70.2%, with an EIDC average of 64.9% before retransmission and 71.6% using the retransmission procedures.

Figure 2-8 compares the transmission percentages for GSETT and the WMO monitoring to the EIDCs and WMO Centers in Moscow and in Washington. Data for Stockholm have not been included in the figure since the EIDC there is not directly connected with a WMO Center on the MTN. The GSETT data in *Figure 2-8* are, except for a few cases, consistently on or below the line in the figure which represents equal transmission percentages for GSETT and WMO. The GSETT values are consistently equal to or smaller than those for the WMO monitoring. This is not surprising, since the transmission of seismic data during the GSETT also involved some transmission outside the WMO/GTS itself, *i.e.*, between national seismic facilities and national WMO centers. Most of the channels with clear differences between GSETT and WMO percentages relate to countries that entered the GSETT at a late stage and for which sufficient time for the necessary preparatory arrangements with WMO/GTS may not have been available.

The network of stations identified with an asterisk in Table 2-8 are those which announced early participation and provided notification of participation to the WMO. The average transmission percentage for the WMO monitoring is 87% for these same channels (not including participants that joined the GSETT at a late stage). The corresponding figure for the GSETT is 74%. If retransmissions are included it becomes 82%. Even if the WMO monitoring values are slightly higher, the distribution of the values for the individual countries show clear similarities for the GSETT and for the WMO monitoring. Channels with low values for WMO also have low values for GSETT. The differences between GSETT and WMO values are, for most channels, quite small.

The percentage of successful retransmissions to the EIDC was calculated as the number of unique Level I data messages received after the first transmission, divided by the total number of unique Level I data messages sent minus the number received during the first transmission. In many instances, the number of retransmission messages sent is quite small and the calculated retransmission percentages are associated with large uncertainities. The percentages in *Figure 2-9* have been plotted against the transmission percentages (first transmission) and the data are compared with a straight-line and a quadratic curve. If the transmission probability between two places, A and B on the WMO/GTS is symmetrical, *i.e.*, the transmission percentages from A to B and from B to A are equal, the data in *Figure 2-9* would follow a quadratic curve. In spite of the scatter, the data suggest that a retransmission is successful only for channels that have fairly high transmission percentages.

Hans Israelsson
Ralph Alewine

TABLE 2-8 GSETT TRANSMISSION PERCENTAGES

Region	Country	No. of Messages Sent	Without Re-Trans			With Re-Trans			Regional Averages			EII
-	-		MOSC	STOC	WASH	MOSC	STOC	WASH	MOSC	STOC	WASH	Aver
Africa	Egypt*	15	6.7	0.0	0.0	6.7	0.0	0.0	3.3	38.3	27.6	23.
	Zambia*	47	0.0	6.4	55.3	0.0	76.6	55.3				
Antarctica	Antarctica*	56	69.6	82.1	66.1	92.9	100.0	92.9	92.9	100.0	92.9	95.
Asia	India*	51	98.0	88.2	58.8	98.0	88.2	58.8	36.3	87.1	72.6	65.
	Indonesia*	41	4.9	73.2	61.0	4.9	73.2	61.0				
	Japan*	50	6.0	96.0	92.0	6.0	100.0	98.0				
Europe	Austria*	22	95.5	100.0	95.5	100.0	100.0	100.0	86.5	92.3	91.1	90.
	Belgium*	26	76.9	88.5	88.5	84.6	96.2	96.2				
	Bulgaria*	35	68.6	80.0	74.3	85.7	88.6	82.9				
	Czechoslo*	46	89.1	97.8	91.3	97.8	100.0	100.0				
	Denmark*	28	78.6	75.0	75.0	82.1	85.7	78.6				
	Finland*	36	100.0	100.0	97.2	100.0	100.0	100.0				
	France	51	29.4	66.7	66.7	60.8	88.2	88.2				
	GDR*	33	84.8	84.8	81.8	97.0	93.9	97.0				
	Germany,F.R.*	50	94.0	100.0	92.0	98.0	100.0	100.0				
	Hungary*	50	86.0	82.0	82.0	90.0	92.0	88.0				
	Italy*	31	80.6	87.1	87.1	93.5	96.8	96.8				
	Ireland	35	51.4	82.9	80.0	65.7	94.3	91.4				
	Netherlands*	52	90.4	96.2	98.1	96.2	96.2	100.0				
	Norway*	51	94.1	100.0	92.2	96.1	100.0	100.0				
	Romania	36	16.7	33.3	27.8	16.7	33.3	27.8				
	Sweden*	144	90.3	100.0	99.3	96.5	100.0	100.0				
	UK*	51	82.4	88.2	86.3	96.1	98.0	96.1				
	USSR*	39	100.0	94.9	82.1	100.0	97.4	97.4				
N.America	Canada*	143	90.2	91.6	86.0	100.0	100.0	99.3	98.3	100.0	99.7	99.
	USA*	422	48.6	99.8	100.0	96.7	100.0	100.0				
S.America	Argentina	13	7.7	38.5	0.0	7.7	38.5	0.0	12.5	27.9	24.6	21.
	Brazil	49	0.0	30.6	28.6	0.0	30.6	28.6				
	Colombia	3	0.0	0.0	33.3	0.0	0.0	33.3				
	Peru*	33	42.4	42.4	36.4	42.4	42.4	36.4				
S.Pacific	Australia*	193	90.2	94.3	91.2	97.4	100.0	97.4	36.6	41.3	43.4	40.
	Fr.Polynesia	31	0.0	12.9	16.1	0.0	12.9	16.1				
	New Zealand*	48	0.0	4.2	10.4	12.5	12.5	16.7				
Averages	All Countries		56.8	70.2	67.6	64.3	76.8	73.8				71.
	with * only		69.0	76.9	76.2	75.8	86.1	82.6				81.

Countries marked with * in the table were listed as WMO/GTS users in the procedures for the GSETT (conference Room Paper 134/Rev.1) prior to the test. The regional averages relate to percentages with re-transmission.

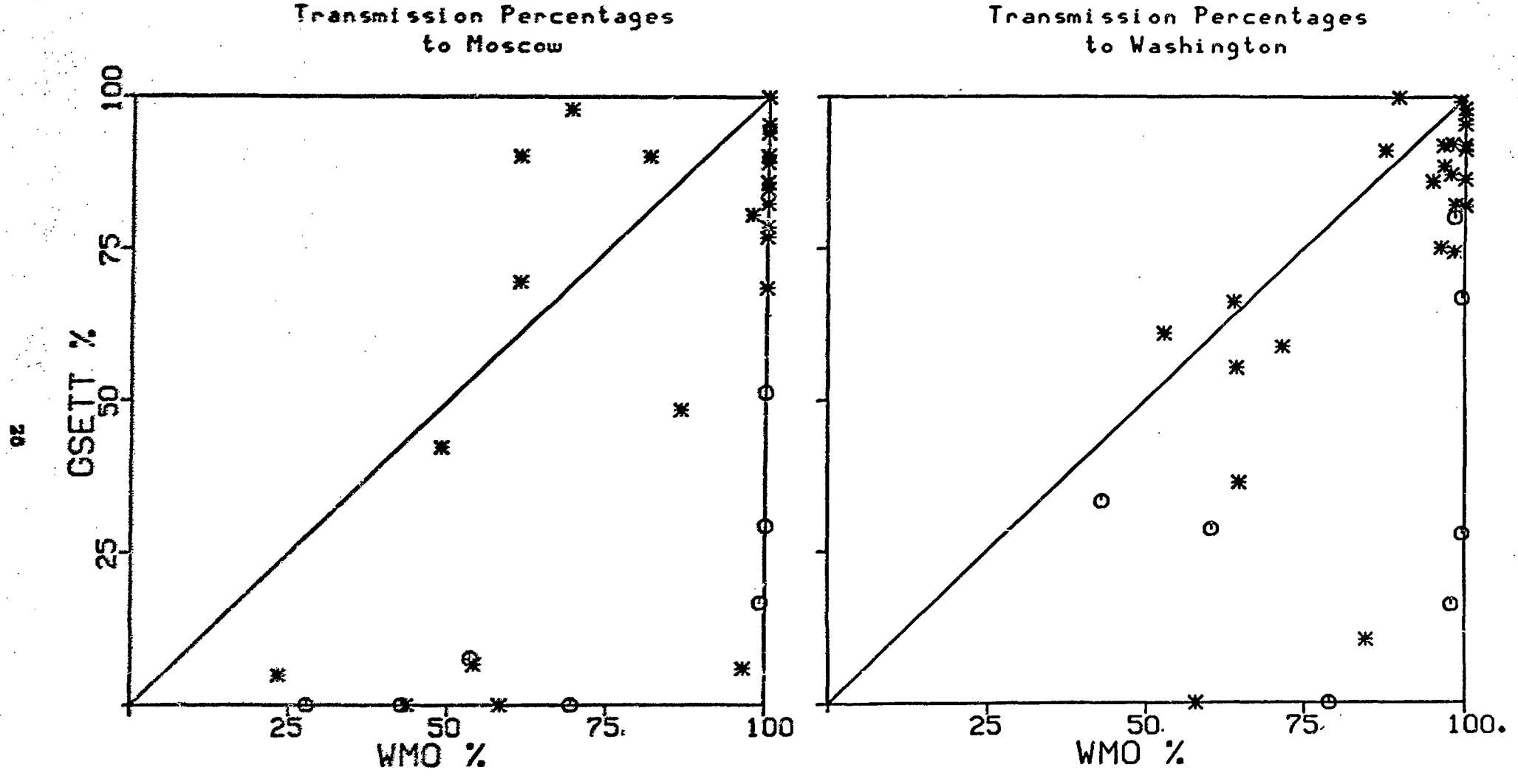


Figure 2-8. Comparison of transmission percentages between the GSETT and the WMO monitoring for Moscow and Washington. The GSETT data refer to seismological messages received at the EIDCs (without re-transmission) and the WMO data refer to meteorological messages received at the WMO Centers. Stars indicate countries that were listed as WMO/GTS users in the procedures for the GSETT (Conference Room Paper 134/Rev.1) prior to the test. Countries entering the GSETT at a later stage are indicated by open circles. The lines correspond to equal transmission percentages for the

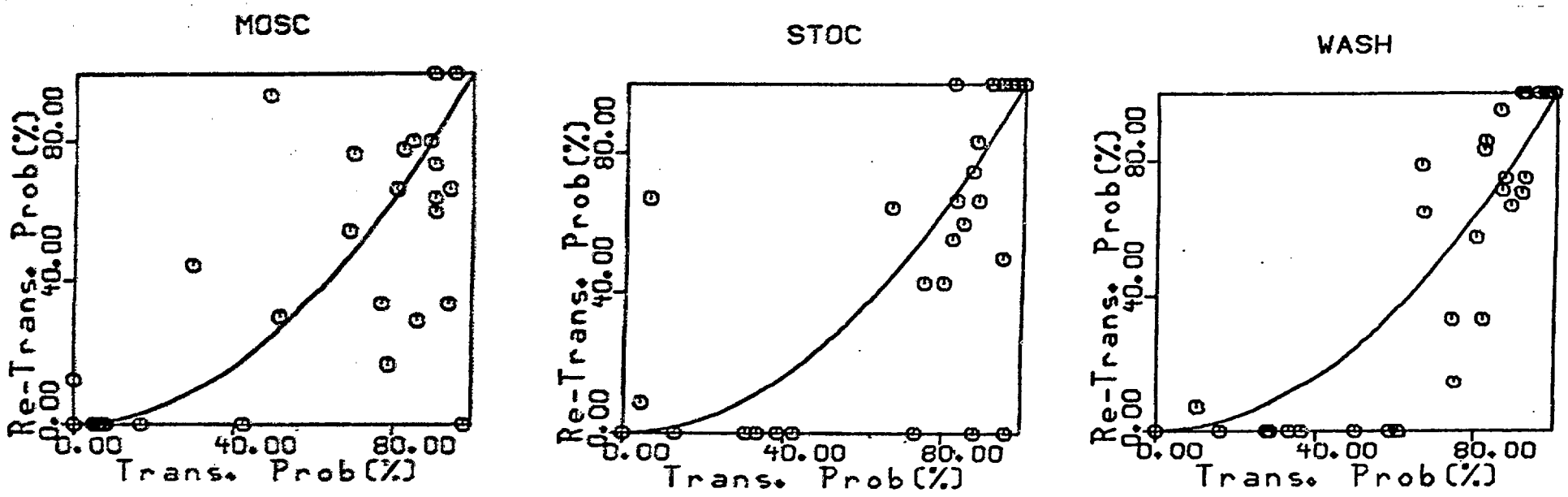


Figure 2-9. The diagram shows the re-transmission percentages against transmission percentages to the EIDC during the GSETT. The curved line indicates quadratic relation.

2.6. DISTRIBUTION OF MOMENTS AND MAGNITUDES AND ESTIMATES OF DETECTION CAPABILITY

A detection threshold, defined for a seismic station or a network by the magnitude at which the cumulative recurrence curve for earthquakes departs from a straight line, is usually adequate as a rough estimate of detection capability. In order to attach high numerical significance to threshold values determined in this way, several factors must be considered.

First, a magnitude determination depends not only on the source but also on the stations providing data for the determination. This makes it difficult to compare detection thresholds computed for different stations and networks. In order to avoid this difficulty, attempts are made here to use the seismic moment as a parameter which is directly related to the source strength and is independent of recording stations or networks.

Second, earthquakes in a given region are regarded as the consequence of a geological process that has a long cycle time, on the order of 100 years or more. Magnitude distributions based on data collected over a short period of time will sample the distribution in a special phase of this cycle. Distributions also may differ from one period to another, and the assumption about an exponential distribution may not be valid at the low end of the magnitude scale. The validity of the exponential distribution is seldom verified by statistical testing.

Data covering only very short time periods are used here as illustrations, including data reported by the NEIS for the year of 1984 and data collected during the Technical Test of the *Ad Hoc* Group of Scientific Experts carried out for two months in the fall of 1984 (GSETT).

The cumulative number of earthquakes as a function of log M_0 is given in *Figure 2-10* for 1984 as reported by NEIS. The curve in *Figure 2-10* starts to taper off at about $M_0=10^{24}$. The "b-value" relation between the cumulative number of earthquakes, N, as a function of moment ($\log_{10}(N)=17.47-0.61\cdot\log_{10}(M_0)$) obtained for a much larger earthquake sample (Chinnery and North, 1975) has been drawn for comparison in the figure. The data for 1984 follow roughly this relation for smaller events, *i.e.*, for M_0 between 10^{25} to 10^{28}. Slope is closer to 0.7 from 10^{24} to 10^{26}. The straight line in *Figure 2-10* implies that the empirical distribution of the moment is exponential. As a generalization, the figure uses the Weibull distribution function. The Weibull distribution was initially introduced in engineering studies of strengths of solids and was used to characterize breaking loads.

Kolmogorov-Smirnov distribution testing was applied to the 1984 data. The hypothesis that the empirical distribution function was that of an exponential or Weibull distribution could be rejected at the 3.2% and 18.0% levels, respectively. Similar results were obtained for the GSETT subset of these data, although the estimated values of the parameters of the Weibull distribution differed slightly.

A simple model is used to describe recorded amplitudes and periods as a function of earthquake source, (S), path, (P), and instrument response (I). This assumes that the amplitude spectrum of a recorded signal, $R(f)$, as a function of frequency, f, can be written as:

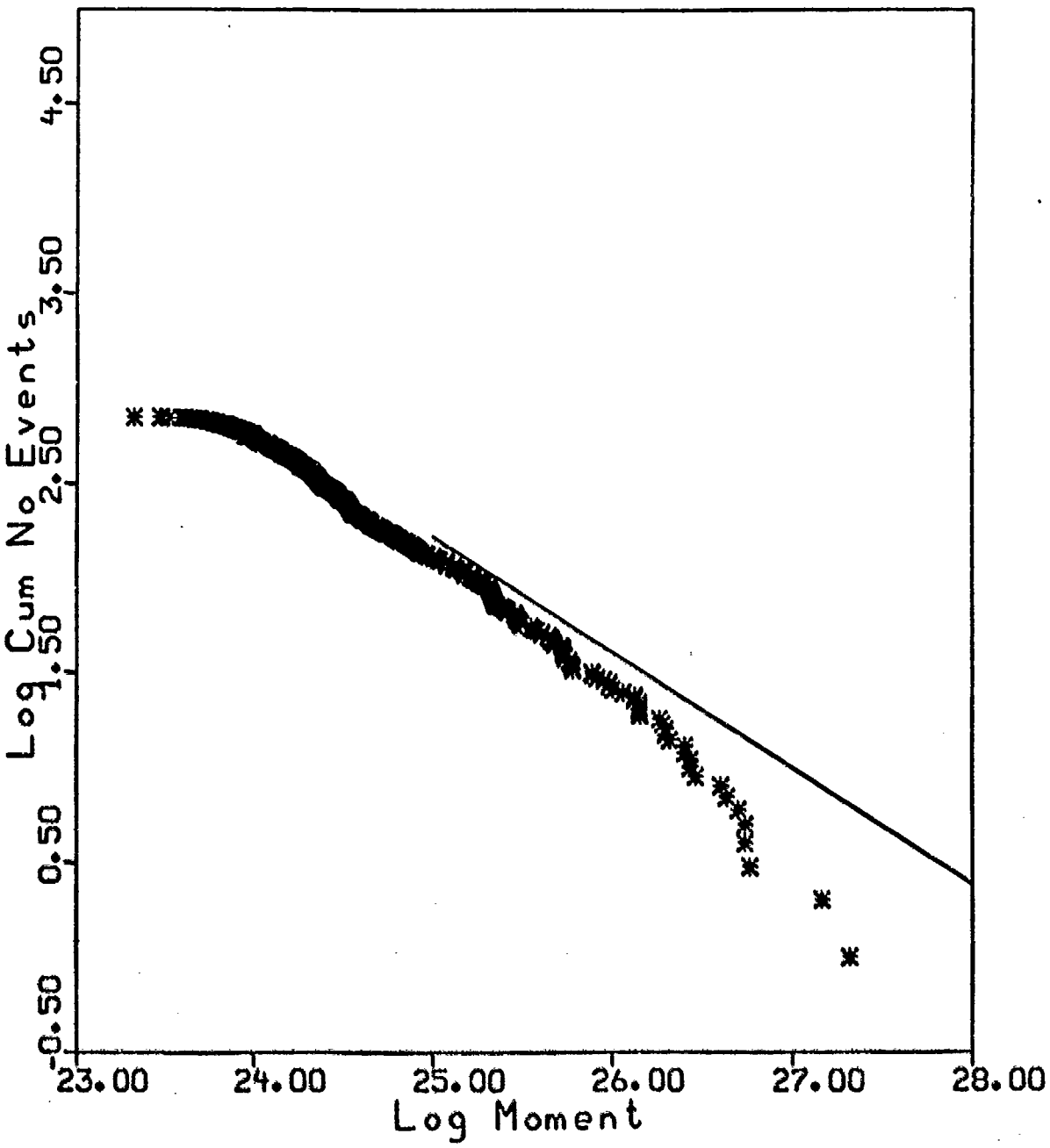


Figure 2-10. The data points show the logarithm of the cumulative number of earthquakes as a function of the logarithm of the seismic moment, M_0, for earthquakes in 1984 as reported by NEIS. The line, $\log_{10}(N)=17.47-0.61\cdot\log_{10}(M_0)$, was estimated from earthqauke data covering more than 40 years of recording (Chinnery and North, 1975).

$$R(f)=S(f)\cdot P(f)\cdot I(f)$$

The source function, S, is assumed to have a cubic roll-off with frequency, based on corner frequency and moment estimates from some 50 earthquakes in California, Turkey, and Tonga-Kermadec (Hanks and Boore, 1984); (Molnar and Wyss, 1972).

The path is assumed to be described by a t^* value for the anelastic attenuation only. The composite spectrum, $R(f)$, usually has a pronounced maximum, the frequency of which, f_{max}, is assumed to correspond to the measured period, $T=1/f_{max}$. Apart from a frequency independent constant, the measured ground amplitude, A, is equal to:

$$A=R(f_{max})/I(f_{max})=S(f_{max})\cdot P(f_{max})$$

The moment M_0 is equal to $A(0)$. This means that a given distribution for M_0 can be used to derive distribution functions for period measurements, $T=1/f_{max}$, and amplitudes, $A(f_{max})$.

Empirical distributions of $\log_{10}(A/T)$, and period data for teleseismic measurements at the RSTN station RSSD, are plotted in *Figure 2-11*. Curves obtained from the model above and the distribution for the moments normalized to the RSSD data are also drawn in the figure and show reasonable agreement with the observed data. These curves represent probability density and have been fitted by eye to the data. Large differences occur, as expected, due to the limited detection capability at smaller amplitude/period ratios and smaller periods.

The model above has also been used to calculate a scaling relation between amplitudes and periods. This relation has been normalized to the data at station RSSD and has also been drawn in *Figure 2-11*, where comparison is made with actual amplitude and period measurements. Although considerable scatter is seen in the observed data, there is reasonable agreement with the calculated curves.

The model above is also used to calculate probability density functions for body wave magnitudes, similar to the amplitude/period ratios in *Figure 2-11*. The Center uses an average value for the amplitude-distance correction, and assumes an instrument response of a WWSSN station rather than that of an RSTN station. This is because comparisons are made with magnitudes reported by the NEIS shown in *Figure 2-12*. The NEIS data cover the same year-long (1984) time period as the moment measurements. The shape of the calculated curve agrees well with the observed data in the magnitude interval m_b=5.0–6.0. The slope of the probability density curve in this interval and the b value for the observed data are close to 1.5 and are clearly different from 1.0. This is a value often obtained from this kind of magnitude data.

The number of earthquakes in the small-magnitude range is an important parameter for test ban monitoring and sometimes this number is partly based on downward extrapolation of the b value curve below the detection threshold. A high b value gives a larger number of earthquakes than a small b value. The calculated probability density curve in *Figure 2-12* is, however, not linear over the entire magnitude range, but tapers off at lower magnitudes, i.e., m_b<5.0, where it approaches a slope of 1.0. In such a case, the number of earthquakes will still be relatively small, even if the b value is high for m_b>5.

Hans Israelsson

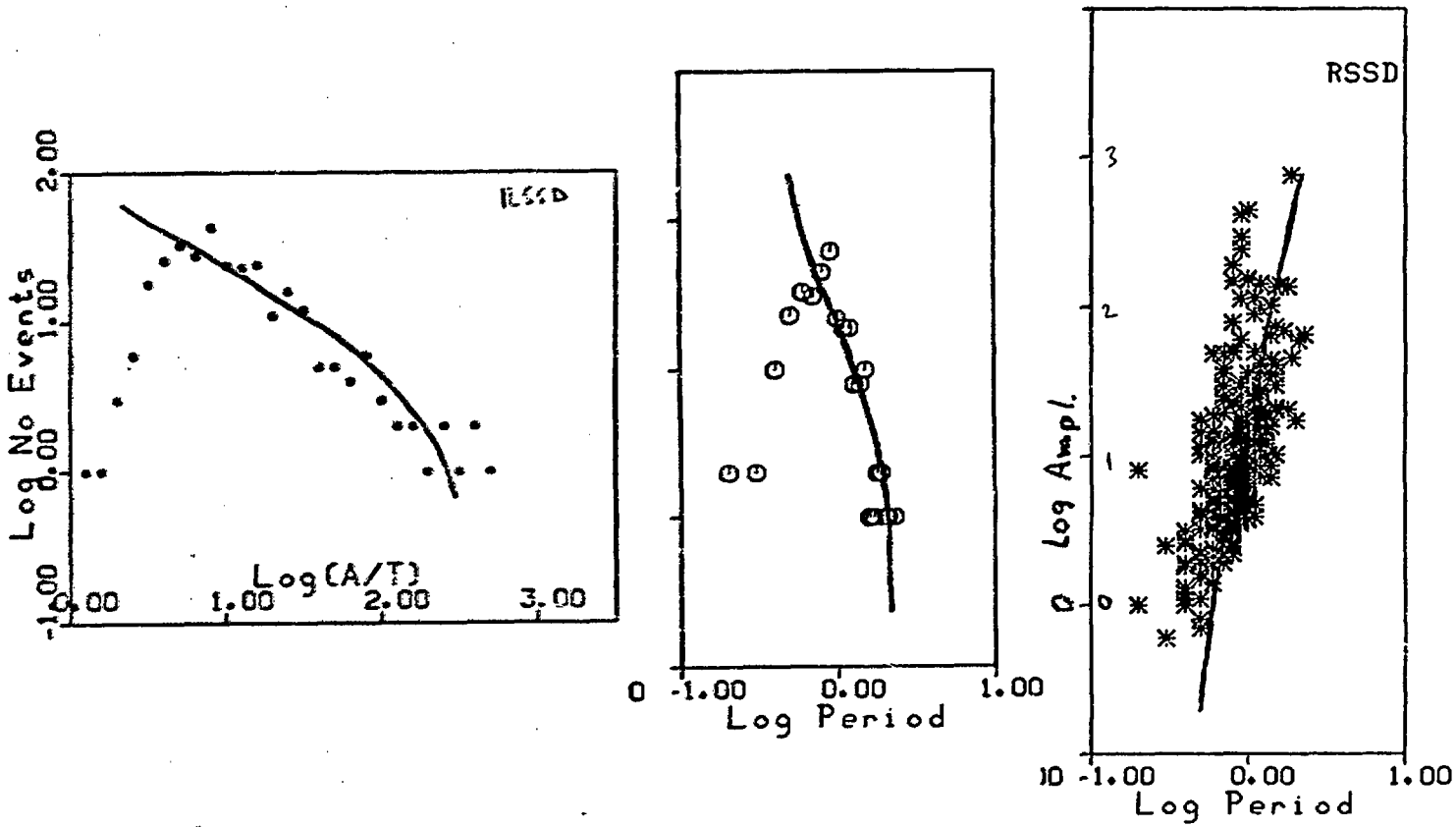


Figure 2-11. Empirical frequency distributions of amplitude/period ratios and periods measured from teleseisms at the RSSD during the GSETT. The curves were obtained as probability densities of amplitude/period ratios and periods from the model discussed in the text. The amplitudes are also plotted against period where the line represents a scaling relation. The curves have been fitted by eye to the observed data. The figure also shows amplitudes plotted versus period with the line representing the amplitude-period scaling obtained from the model discussed in the text.

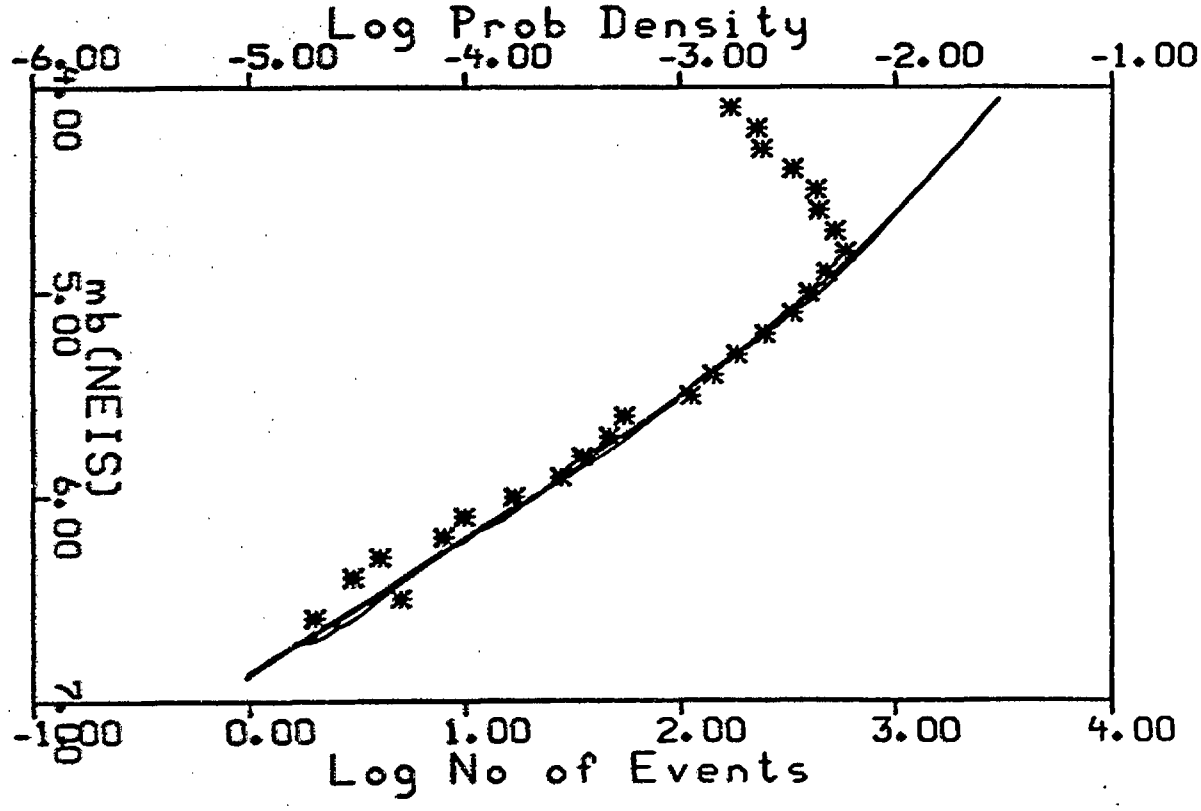


Figure 2-12. Number of earthquakes as a function of body wave magnitude, m_b, reported by the NEIS for 1984. The curve was obtained from the model used in the text and gives the probability density function computed for body wave magnitudes for stations with WWSSN type response.

2.7. REFERENCES

Chinnery, M.A. and R.G. North, "The Frequency of Very Large Earthquakes." *Science,* Vol. 190, pp. 1197-1198, 1975.

Hanks, T.C. and D.M. Boore, "Moment-Magnitude Relations in Theory and Practice." *J. Geophys. Res.*, Vol. 89, pp. 6229-6235, 1984.

Marshall, P.D., D.L. Springer, and H.C. Rodean, "Magnitude Corrections for Attenuation in the Upper Mantle." *J Geophys. J. R. Astr. Soc.*, Vol. 57, pp. 609-638, 1079.

Molnar, P. and M. Wyss, "Moments, Source Dimensions and Stress Drops of Shallow-Focus Earthquakes in Tonga-Kermadec." *J Phys. Earth and Planet. Interiors*, Vol. 6, pp. 263-278, 1972.

Ringdal, F., *Study of Magnitudes, Seismicity, and Earthquake Detectibility Using a Global Network.* NTNF/Norsar, Kjeller, Norway, 1984.

Thomas, John B., *Introduction to Statistical Communication Theory*, John Wiley & Sons, New York, 1969, pp. 670.

Veith, K.F., and G.E. Clawson, "Magnitude from Short Period P-wave Data." *J Bull Seism Soc Am*, Vol. 62, pp. 435-452, 1972.

3. RESEARCH TO IMPROVE ANALYSIS OF SEISMIC DATA

3.1. POLARIZATION PROCESSING ALGORITHM WITH APPLICATIONS TO RSTN DATA

3.1.1. Introduction

Different seismic wave types have distinct particle-motion signatures. Orthogonal three-component seismograms contain the total ground-motion vector as a function of time and thus the polarization characteristics of all arrivals present. However, typical seismograms can be complicated. A given time interval usually contains several types of seismic waves from the same source, arrivals from different origins, scattered energy, and in some cases one type of seismic wave from one source arriving from several directions at once. Polarization processing is an attempt to identify the polarization signatures of different arrivals and separate them based on these signatures. From this, information can be gained on the origin of the seismic waves, the earth's structure, and propagation characteristics of different wave types.

This report describes the design and application of a time domain three-component processor which estimates the ground motion polarization as a function of time and frequency. The assumption is made that each frequency component is elliptically polarized over several cycles of its duration. The data are decomposed into narrow frequency bands and overlapping time windows, and the polarization parameters are estimated separately within each segment.

In the past, several different methods for polarization processing of seismic recordings have been proposed. Flinn (1965) computed the polarization ellipse in sliding time windows using the covariance matrix of the three components of motion. Filtering was achieved by applying the degree of desired polarization as a point-by-point gain control on the input traces. Smart (1977, 1981) proposed a three-component processor in which the parameters of a specified polarization model were determined by a least-squares fit in the frequency domain. This method yielded polarization parameters of a seismic wave as a function of time and frequency, but was not implemented as a filter to pass specific particle motions. Samson and Olson (1980, 1981) designed a technique for estimating the polarization state from the spectral matrix in the frequency domain. The analysis was performed in sliding time windows and could be applied to single-component multichannel array data as well. Polarization filtering was carried out by applying the degree of desired polarization as a frequency- and time-dependent gain function on the input data. Suteau-Henson *et al.* (1985) successfully applied this same technique to three-component ocean-bottom seismograms, RSTN recordings of regional earthquakes and NORESS array data of local events.

The polarization processing algorithm described in this report is novel in that it operates solely in the time domain to estimate the polarization state as a function of time and frequency. This approach obviates the need for spectral estimation within short time windows encountered by frequency domain algorithms. Furthermore, the polarization

filter passes only the vector particle motion with the desired polarization, as opposed to applying a simple gain function to the input traces. The filtered ground motions are constructed by summing together the polarized particle-motion contributions from all time windows and frequency bands. As a result, the filtered seismograms appear very much like recorded ground motions, which is generally not the case for polarization algorithms that apply a simple gain function to the input traces.

This report begins with an outline of the method for doing the frequency and time decomposition of the seismic traces. Next is a description of the algorithm for estimating the polarization parameters and performing the filtering to pass specifically polarized energy. Finally, as examples, the polarization processing is applied to synthetic data and an RSTN recording of an regional earthquake.

3.1.2. Frequency and Time Decomposition

The motivation for decomposing the seismograms as a function of both frequency and time is to separate the contributions of seismic pulses with different arrival times and frequency responses. Clearly, simultaneous multiple arrivals with identical frequency responses cannot be completely separated using a method such as this. The resolution of the processor is determined by the time window lengths and filter bandwidths used in the decomposition.

Figure 3-1 shows a block diagram of the processing algorithm. The main body of the program is contained within two nested loops. In the first loop, the input traces are filtered into narrow frequency bands using a zero-phase filter. The individual bands are overlapped so that a time domain summation of all passband responses yields a flat amplitude spectrum.

This concept is illustrated in *Figure 3-2*, where the amplitude spectra as well as time responses are shown for seven frequency bands. A time domain summation of all the impulse responses is also shown as the bottom trace on the right side. The amplitude spectrum of this trace is the flat portion between 1 and 10.5 Hz. In this example, a sixth-order butterworth filter was applied consecutively in both directions to give a zero-phase response. The bandwidth Δf is a constant number of octaves and is related to the center frequency f_c by

$$\Delta f = \frac{f_c}{ncycles}$$

where *ncycles* is an input parameter. The parameter *ncycles* controls the bandwidth and the time resolution of each bandpass filter. In *Figure 3-2* the parameter *ncycles* has been set to 3. The duration of the filter impulse response clearly decreases with frequency. The frequency resolution is controlled by the width of the passbands and the amount of overlap between them. The overlap is simply related to the fall-off rate of the filter outside the passband. A summation of all the passbands is always unity.

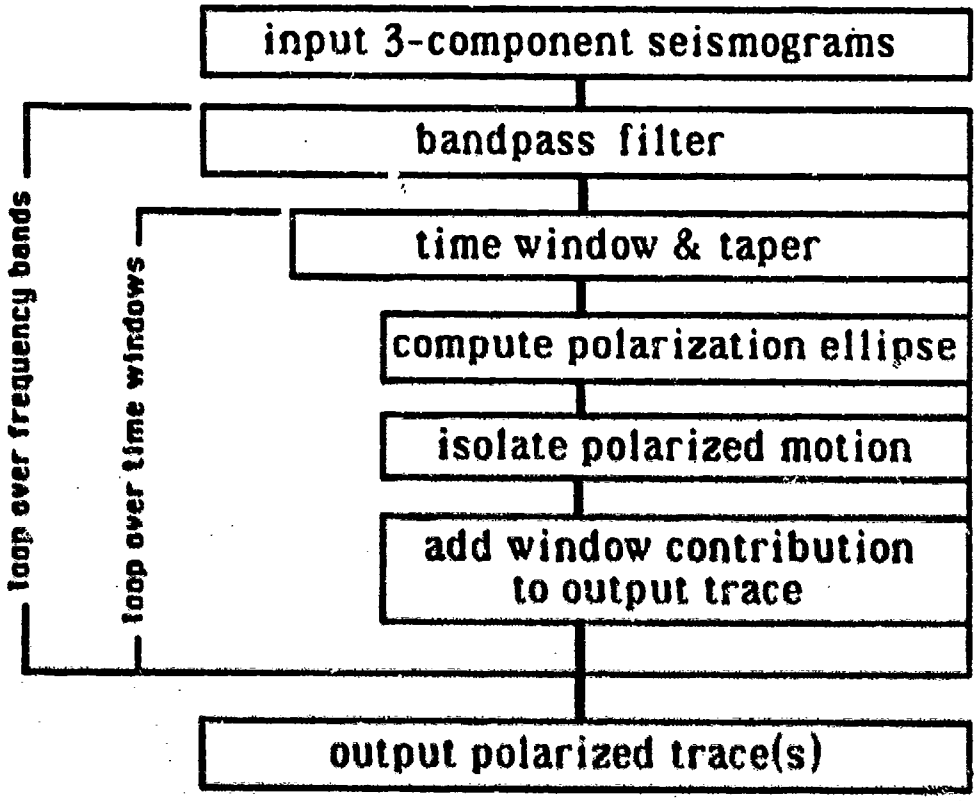


Figure 3-1. Block diagram of the polarization processing algorithm. The two main loops are over frequency and time.

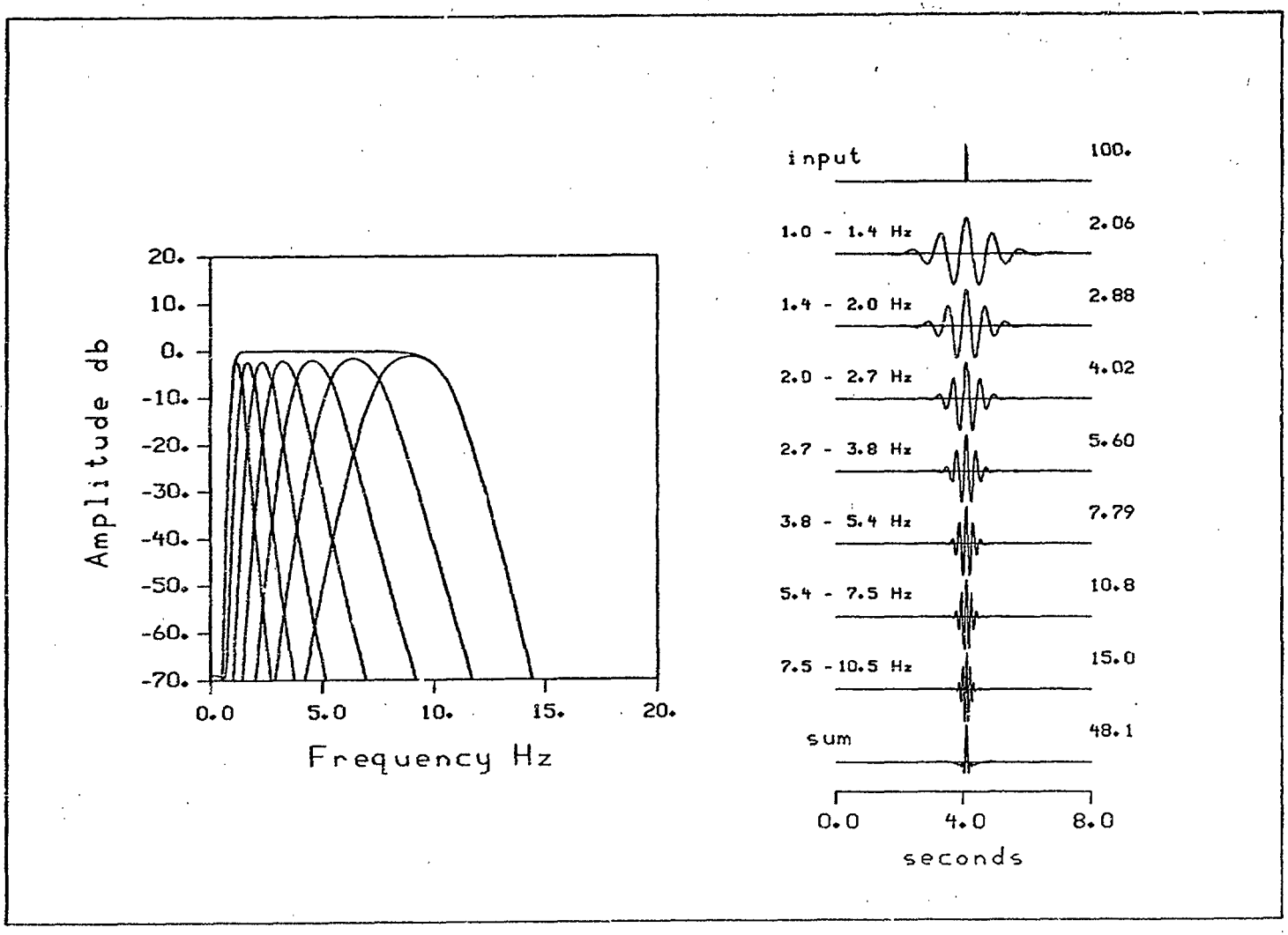


Figure 3-2. **Frequency decomposition using seven bandpass filters. The upper trace on the right side shows a spike input. The impulse response of each bandpass filter is shown below; the time resolution is a constant number of cycles. The bottom trace is a summation of all the impulse responses; its amplitude spectrum is flat over the band 1 to 10.5 Hz.**

The inner loop in *Figure 3-1* is over time windows. *Figure 3-3* schematically shows this concept. Identical tapered windows are applied to all three orthogonal input traces after bandpass filtering. The time resolution is determined by the window length. Adjacent windows are overlapped so that a summation of all segments gives an envelope of unity. The overlap between adjacent windows is 50% and thus the taper length is also 50%. A cosine taper is used. This large overlap gives a substantial redundancy which tends to stabilize the estimation process. The time window length used, $N\delta t$, is related to the time resolution of each passband. This is, in turn, determined by the band center frequency and the parameter *ncycles*:

$$N\,\delta t = \frac{1}{\Delta f} = \frac{ncycles}{f_c}$$

where N is the number of samples per segment and δt is the sampling interval.

Thus, the window length is inversely proportional to the passband width. This is the normal trade-off between time and frequency resolution encountered in spectral estimation. A long time window has good frequency resolution, but will smooth over arrivals of several seismic pulses, which could have different polarization characteristics. A short time window has better time resolution, but the estimation variance and sensitivity to noise are increased. The actual values used for the time and frequency resolution are data dependent. The processing algorithm sets the time window length equal to some multiple number of cycles of each passband center frequency. In applications done to date, the results are quite robust with respect to the window length. A time segment of 3 to 5 cycles of the passband center frequency has worked well in all cases.

3.1.3. Polarization Estimation

The polarization ellipse is estimated for each time segment in each frequency band. The covariance matrix R_{ij} is evaluated via

$$R_{ij} = \frac{1}{N}\sum_{t=1}^{N} S_{it}S_{jt} \qquad i,j=1..3$$

where i and j are trace indices, t is time index, N is the number of samples per window, and S_{it} is sample t of trace i. The covariance matrix R_{ij} is 3x3, real, symmetric and positive semidefinite (the eigenvalues are positive and may be zero). R_{ij} is the matrix of coefficients for a quadratic form which is an ellipsoid. This ellipsoid is the best fit to the data points in a least-squares sense. The algebraic eigenproblem for R is to find the eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$ and eigenvectors $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ which are the nontrivial solutions to

$$[R - \lambda^2 I]\,\mathbf{u} = 0$$

where I is the 3x3 identity matrix. The eigenvectors are orthogonal and unit length.

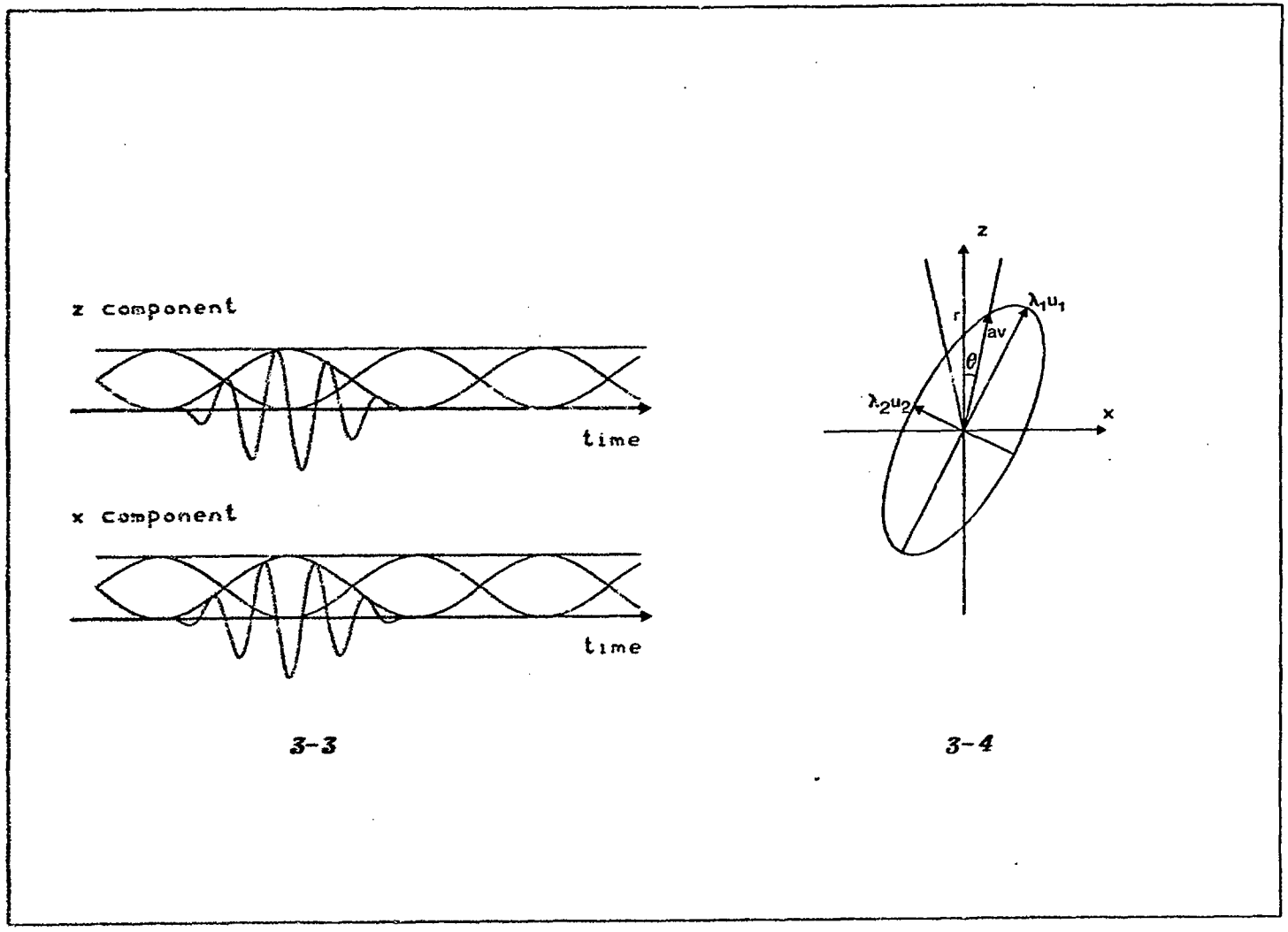


Figure 3-3. Time decomposition using identical overlapping cosine-tapered windows applied to each component.

Figure 3-4. A schematic of the polarization ellipse in two dimensions. The principle axes are $\lambda_1 u_1$ and $\lambda_2 u_2$. The angle θ represents the aperture function, measured from the pass direction x=x in this case.

The three principle axes of the ellipsoid are formed by $\lambda_i \mathbf{u}_i$, where $i = 1...3$. The principle-axis lengths are λ_i and the directions are the eigenvectors $\mathbf{u}_i$. The squared eigenvalue λ_i^2 is the variance or power of the signal in the direction of $\mathbf{u}_i$ and λ_i is the standard deviation or amplitude. The eigenvalues are ordered such that $\lambda_i \geq \lambda_j$ for $i<j$. Purely rectilinear ground motion has only one non-zero eigenvalue; $\lambda_i=0$ $i\neq 1$. Purely elliptical polarization has two non-zero eigenvalues, $\lambda_1 \geq \lambda_2$, $\lambda_3=0$, and circular motion is the special case $\lambda_1=\lambda_2$, $\lambda_3=0$. Elliptical polarization also has a sense of rotation, prograde or retrograde depending on the convention for positive. In real applications, all three eigenvalues are generally non-zero and non-equal, so the polarization is ellipsoidal. The covariance matrix is usually well-conditioned because random noise and windowing effects are uncorrelated from trace to trace. The eigenvalue decomposition of a 3x3 real symmetric matrix is very fast using standard mathematics libraries such as IMSL.

Once the eigenvalues and eigenvectors of the polarization ellipsoid are estimated, the polarization state of the ground motion in the data segment is completely determined. For example, the degree of rectilinearity is given by the relative lengths of the principle axis

$$rectilinearity = 1 - \frac{\lambda_2+\lambda_3}{2\lambda_1}$$

and the direction of dominant rectilinear motion is the major axis $\mathbf{u}_1$. A useful way of displaying the rectilinear polarization content of three-component seismograms is to plot, as a function of both frequency and time, the ellipticity and the orientation of the major axis of the ellipsoid for each window.

3.1.4. Polarization Filtering

Polarization filtering is performed by passing only the vector component of motion with a desired polarization state. The following is a description of the filtering algorithm used for passing rectilinear motion. There is no one unique and correct way of separating individual particle motion signatures when several unspecified pulses are present simultaneously. This particular approach has been designed primarily to isolate rectilinear polarization in P-wave codas.

First, the polarization ellipse is computed in each time segment and frequency band as described above. A unit direction vector $\mathbf{r}$ is used to specify the desired orientation of rectilinear motion to pass. The pass direction is generally not parallel to any of the principle axes of the ellipsoid. Next, the noise contribution is removed by subtracting the third and smallest eigenvalue λ_3 from all the eigenvalues,

$$\lambda_1^2 \equiv \lambda_1^2 - \lambda_3^2$$

$$\lambda_2^2 \equiv \lambda_2^2 - \lambda_3^2$$

$$\lambda_3^2 \equiv \lambda_3^2 - \lambda_3^2 = 0$$

This operation is justified by the observation that the covariance matrix of uncorrelated random noise is diagonal with $\lambda_1=\lambda_2=\lambda_3$. Seismic waves have polarization states that are purely elliptical or rectilinear, and the motion is treated as purely elliptical from this point forward. In actual seismograms, simultaneous arrivals with different polarization states often contribute to the same polarization ellipse, which causes the third eigenvalue to become non-zero. Discarding the third eigenvalue therefore assumes that the desired motion to be passed is in the plane of the two largest principle axes. *Figure 8-4* shows a schematic of a polarization ellipse.

The polarization filter contains an aperture function for passing motions which may be close to, but not exactly in, the desired orientation. The aperture has the form $\cos^n\theta$ where the angle θ is measured from the pass direction $\mathbf{r}$. Any rectilinear motion polarized at an angle θ from $\mathbf{r}$ is passed with a reduced amplitude $\cos^n\theta$. The value n, which is the aperture width, is an input parameter; typical values range from 1 to 50. If $n=1$, then the aperture is at its widest and the amplitude passed is just the cosine or projection. For large n, the aperture is very small. Rectilinear motion at 30° from $\mathbf{r}$ is scaled by a factor 0.15 for $n=30$.

The motion passed by the filter has orientation and amplitude which are found by combining the polarization ellipse with the aperture function. Call $a\mathbf{v}$ the vector from the origin in *Figure 8-4* to any point on the ellipse; a is the amplitude of the ellipse in direction $\mathbf{v}$. Thus $a\mathbf{v}$ satisfies the equation of the ellipse with principle axes $\lambda_1\mathbf{u}_1$ and $\lambda_2\mathbf{u}_2$. The multiplication of $a\mathbf{v}$ by the aperture function $\cos^n\theta$ is $a(\mathbf{v}\cdot\mathbf{r})^n$, since $\mathbf{v}\cdot\mathbf{r} = \cos\theta$. The rectilinear motion passed by the filter, then, is the maximum of the ellipse multiplied by the aperture function, i.e., the maximum of $a(\mathbf{v}\cdot\mathbf{r})^n$ over the entire ellipse. Call the direction of this maximum $\hat{\mathbf{v}}$ and the amplitude of the ellipse in this direction $\hat{a}$. Then $\hat{a}\hat{\mathbf{v}}$ is always on the ellipse, but generally not in any of the directions $\mathbf{r}$, $\mathbf{u}_1$, or $\mathbf{u}_2$. If the ground motion is purely rectilinear, this maximum motion will be in the direction of the major axis, even with the effect of the aperture. The motion passed has an amplitude $\lambda_1(\mathbf{u}_1\cdot\mathbf{r})^n$ and orientation $\mathbf{u}_1$. If the ground motion is rectilinearly polarized in the pass direction $\mathbf{r}$, then the motion passed is simply $\lambda_1\mathbf{r}$ since $\mathbf{u}_1\cdot\mathbf{r} = 1$.

The ground particle motion $\hat{S}$ passed by the filter within each window is formed by the orthogonal input components rotated to direction $\hat{\mathbf{v}}$:

$$\hat{S}_i = \sum_{i=1}^{3} \hat{v}_i S_{ij}.$$

Since the amplitude of the motion passed by the aperture was computed as $\hat{a}(\hat{\mathbf{v}}\cdot\mathbf{r})^n$, a scaling must be introduced to normalize the amplitude of the particle motion passed,

$$\hat{S}_i = \hat{S}_i \cdot \frac{\hat{a}(\hat{\mathbf{v}}\cdot\mathbf{r})^n}{\left[\frac{1}{N}\sum_{i=1}^{N}\hat{S}_i^2\right]^{1/2}}.$$

Now $\hat{S}_i$ is rectilinearly polarized in direction r and has amplitude determined by the aperture and the polarization ellipse.

The final step in the filtering algorithm is the introduction of a scaling term related to the ellipticity,

$$\hat{S}_i = \hat{S}_i\left(1-\frac{\lambda_2}{\lambda_1}\right)^m$$

This scaling is applied as a point-by-point gain to the computed particle motion for each window. If the polarization ellipse is rectilinear, $\lambda_1 \gg \lambda_2$ and the scaling is unity. For circular polarization $\lambda_1=\lambda_2$ and the scaling is zero. The parameter m is an input to the program. This scaling term controls the filter response to particle motions which are not rectilinear.

When the particle motion to be passed has been computed, the window contribution is added onto to the output trace at the correct time lag. The processing algorithm then repeats the above steps for each time and frequency loop summing the individual contributions to produce the final output trace. The algorithm for passing elliptically polarized surface waves is similar to the rectilinear algorithm described above. The main difference is that the computed motion is elliptical so two output traces are generated.

3.1.5. Application to Synthetic Data

Figure 3-5 shows the application of three-component polarization processing to an idealized test example. Three orthogonal input traces of 500 samples each contain four separated arrivals. The first three arrivals are rectilinearly polarized with incidence angles 0, 15, and 30° from vertical. The fourth arrival is circularly polarized in the vertical-north plane; the north component is 90° out of phase from the vertical. Random noise has been added and all three traces zero-phase bandpassed between 1 and 7.5 Hz. The upper three input traces in *Figure 3-5* have signal-to-noise ratio of 50. The fourth trace shows the polarization-filtered output designed to pass only particle motion rectilinearly polarized in the vertical direction. The aperture function was $n=30$, and the ellipticity scaling parameter was $m=1$. The time window length used was four cycles of each passband center frequency. The first arrival has been passed by the filter virtually unscathed. The second and third arrivals have been reduced by the aperture function $\cos^n\theta$ since, although rectilinear, their orientation is not vertical. The fourth arrival is eliminated by the ellipticity scaling term

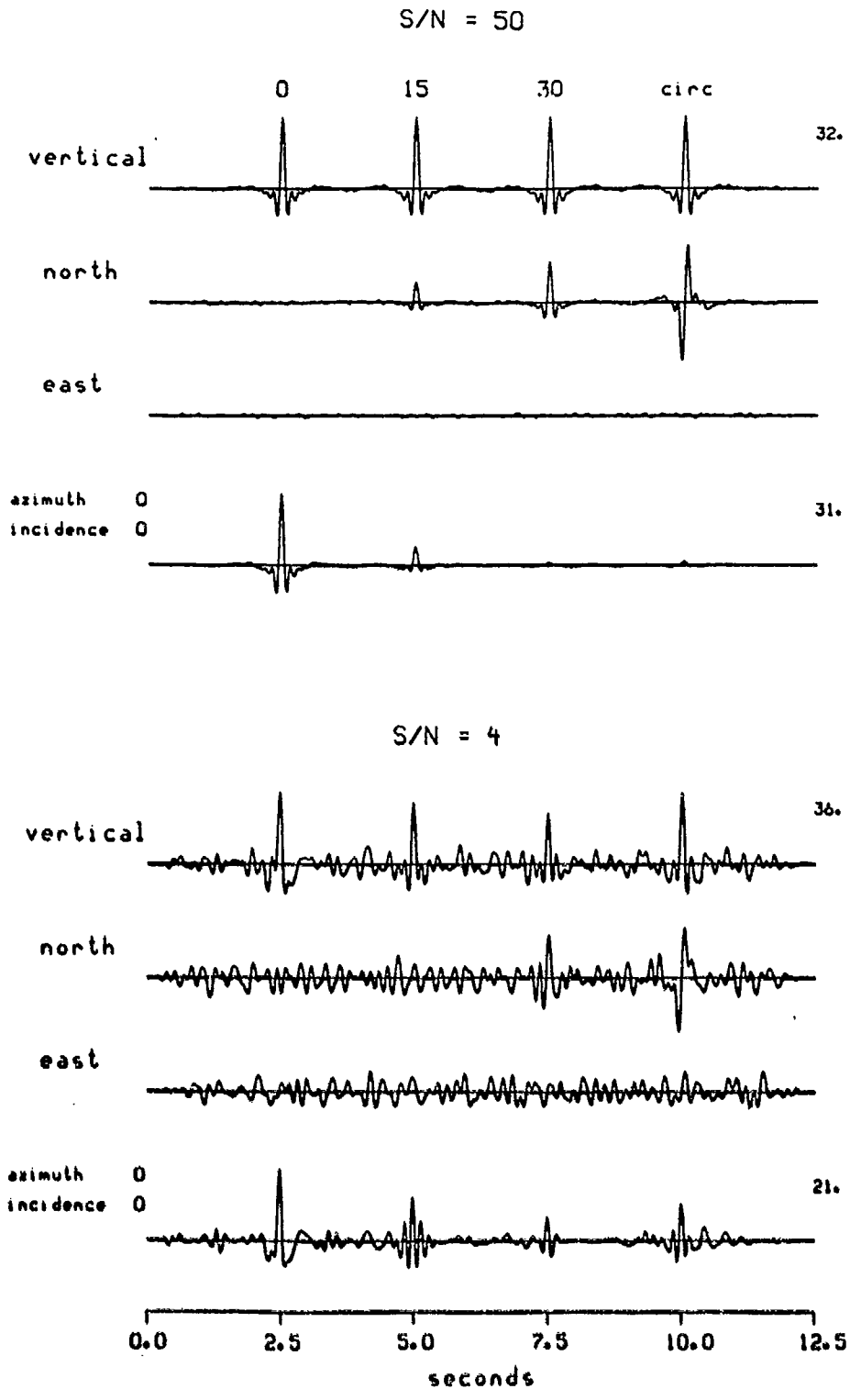


Figure 3-5. **Application of polarization filtering to a simple test example. The input traces contain four arrivals; three are rectilinear with varying orientations from vertical and one is circularly polarized. The filtered output is the rectilinear particle motion in the vertical direction. The filter aperture function allows some motion to pass which is close to but not exactly in the desired polarization state. The bottom set of traces are identical to the top set but the random noise level is higher.**

$$(1-\frac{\lambda_2}{\lambda_1})^m$$

because it is circularly polarized and $\lambda_1 = \lambda_2$.

The bottom part of *Figure 3-5* shows the same example except that the noise level has been increased to S/N = 4. The peak amplitude of the processed trace is smaller than for the low-noise case. In general, the effect of random noise is to increase the magnitude of the second and third eigenvalues relative to the first eigenvalue. Therefore, noise distorts the polarization ellipse and decreases the ellipticity of pure rectilinear signals. The algorithm attempts to compensate for noise by subtracting the smallest eigenvalue from the two largest and scaling by the ellipticity term. The net effect of additive noise is to "smear" the performance by rejecting any signal that would be passed in the noiseless case and passing any signal that would be otherwise be rejected. If several randomly polarized seismic pulses are contained within the same window, the polarization ellipse is affected in a manner similar to the addition of random noise.

3.1.6. Application to an RSTN Regional Earthquake Recording

The second application of polarization processing is to an RSTN recording of the Goodnow, N.Y. earthquake of 7 October 1983 (m_b=5.1). The top three traces in *Figure 3-7* are the three components of the initial P-wave coda recorded on the short-period instrument at station RSCP in Tennessee. The epicentral distance was 12°. The initial arrival is at approximately 4.0 seconds on this record.

Figure 3-6 shows the ellipticity, incidence angle and azimuth of the rectilinear motion as a function of time and frequency. The bottom trace is just the observed vertical component for reference. The ellipticity, which is the degree of rectilinearity, is quite variable over time and frequency. The values prior to about 3.8 seconds correspond to the ambient noise and are of little interest. The ellipticity is quite high at all frequencies for the initial P-wave arrival at 4.0 seconds. However, the incidence angle at 4.0 seconds is a strong function of frequency. The higher frequencies are generally more vertically incident than the lower frequencies over the entire record duration. The incidence angle in the band .5 - 2.0 Hz at 4.0 seconds is about 34° and the azimuth here is very close to the theoretical one of 223°. The azimuth for energy which is almost vertically incident is poorly determined because the horizontal vector component is very small and easily obscured by noise.

The fourth trace in *Figure 3-7* shows a polarization filtered output. In this case, rectilinear P-wave motion emerging at an azimuth of 223° and an incidence angle of 34° from vertical has been passed by the processor as the dominant direction of rectilinear motion in the time around 4.0 seconds in the band .5 - 2.0 Hz. This time corresponds to the initial pure P-wave signal. The results showed that the higher frequencies are polarized more vertically than the lower frequencies. The fourth trace has relatively less high frequency energy content than the vertical input trace. The fifth trace in is simply a high-passed version of the fourth trace.

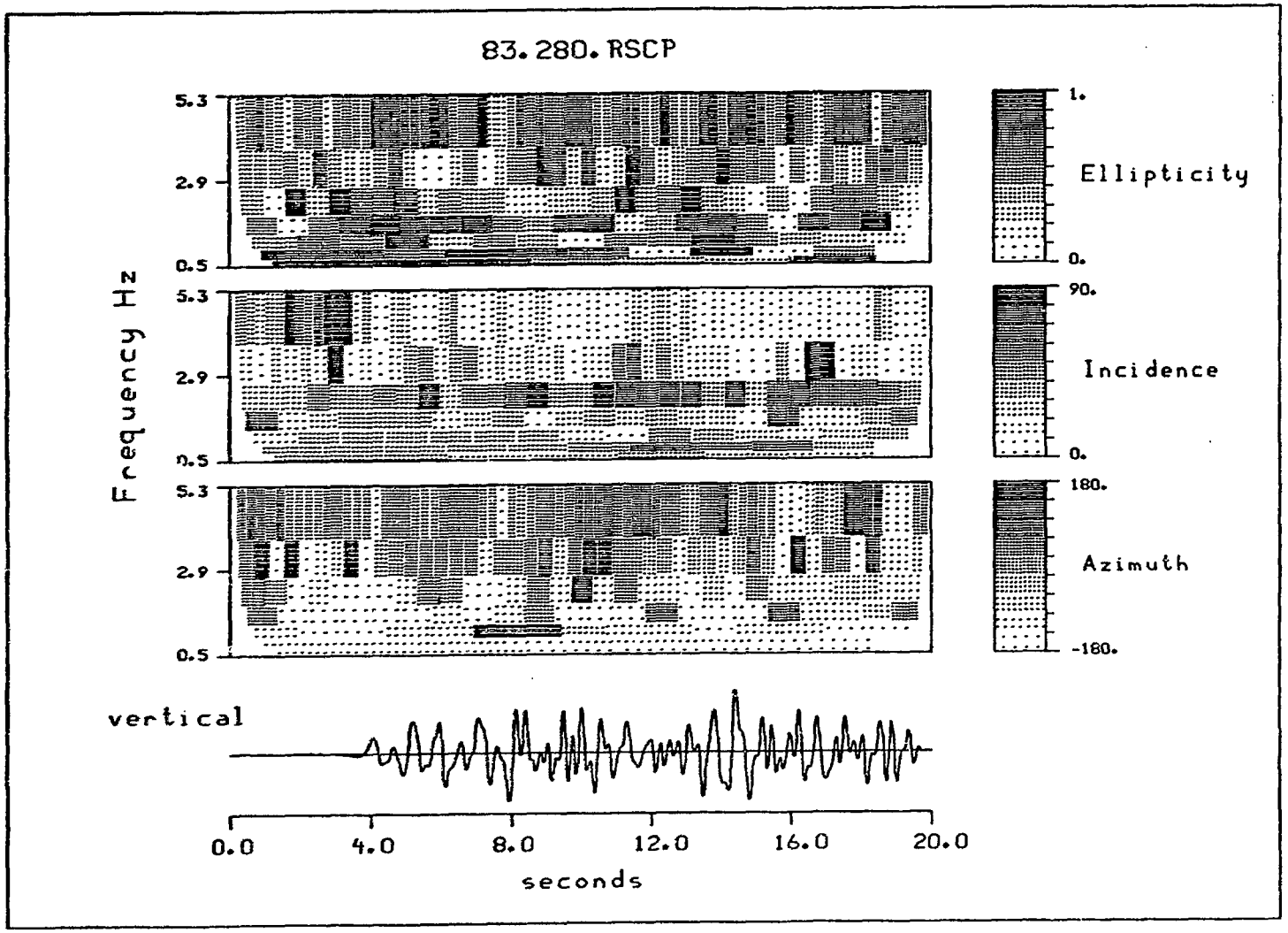


Figure 3-6. **Polarization parameters as a function of frequency and time for a segment of the RSCP recording of the 7 October 1983 Goodnow N.Y. earthquake. The bottom trace shows the P-wave coda observed on the vertical component.**

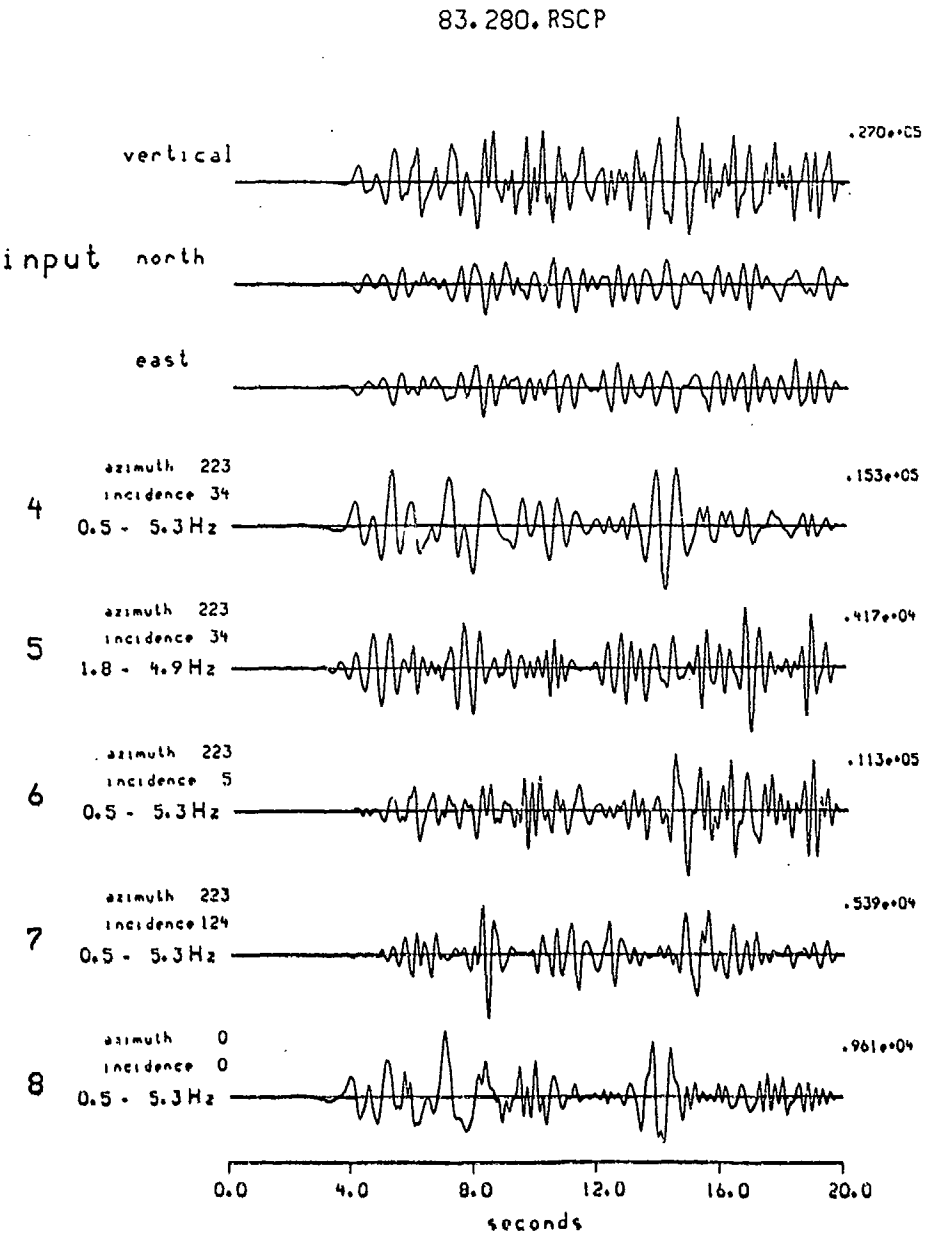


Figure 3-7. **Polarization filtering applied to the 7 October 1983 Goodnow earthquake at RSCP. The top three traces show the 3 input components. The other traces are outputs from polarization filtering. Trace 4 is the broadband rectilinear motion with an incidence angle of 34°. Trace 5 is a bandpassed version of trace 4. Trace 6 is the broadband rectilinear motion at an incidence of 5°. Trace 7 is the rectilinear motion in the S_V orientation. Trace 8 is the vertical input component scaled by the ground motion ellipticity as a function of frequency and time.**

The sixth trace in *Figure 3-7* shows the wide-band rectilinear motion passed for an incidence angle of 5°, which is approximately the dominant polarization direction of the higher frequency energy. A comparison of the amplitudes of the fifth and sixth traces shows more high-frequency energy emerging at 5° than at 34. The seventh trace shows the rectilinearly polarized energy in the S_V direction. The incidence angle used in this case was 124°, 90° from the low-frequency incident P-waves. Several isolated packets of energy are apparent in the S_V motion which lag the rectilinear P-wave energy in time. All the results presented above were obtained using an aperture parameter of n=30, an ellipticity parameter of m=1 and a window length of 4 cycles. The bottom trace in the figure shows the vertical component of motion with a point-by-point gain function applied. This gain is the frequency and time dependent ellipticity term with m = 10. The gain is near unity if the particle motion is rectilinear.

3.1.7. Summary

An algorithm has been outlined for estimating the ground particle-motion polarization of three-component seismograms and filtering them on this basis. In this approach, the polarization parameters are evaluated in the time domain as functions of time and frequency. The assumption is made that each frequency component is elliptically polarized over several cycles of its duration. Such a time domain approach obviates the need for spectral estimation within short time windows associated with frequency domain techniques. Polarization filtering is performed by passing only the vector particle motion with the desired polarization state. An aperture function is used, allowing motion to pass which is polarized close to, but not exactly in, the desired orientation. The polarization processor was implemented to estimate and filter rectilinear particle motions on synthetic data and on RSTN recordings of a regional earthquake.

The synthetic tests indicated that performance of the algorithm is sensitive to random noise when high-resolution short time windows are used. Investigations are currently being performed on methods of improving the processing performance on noisy recordings.

The second application was to the P-wave coda of a m_b=5.1 earthquake in New York State recorded at RSCP. Results showed a high degree of particle-motion rectilinearity at the time of the direct P-wave arrival at all frequencies. However, the angle of incidence at this time was a strong function of frequency. The higher frequencies were more vertically polarized than the low frequencies, generally throughout the entire record. Subsequent analyses of recordings at several RSTN stations have shown a similar frequency-dependent polarization direction. Polarization filtering of this event showed 3 or 4 distinct packets of rectilinearly polarized energy in the P as well as S_V directions. The S_V energy was delayed relative to the P pulses.

Future plans for research on this topic include applying the polarization processing to P-wave coda of teleseismic recordings in an attempt to separate the contributions of near-receiver secondary phases from near-source depth phases, implementing a filtering algorithm for estimating and passing surface-wave particle motions, and investigating the cause for the observed changes in P-wave polarization angles with frequency.

Andy Jurkevics

3.2. EXPANDED USE OF COMPUTERS IN REGIONAL DATA ANALYSIS

Computer usage in seismology has been primarily confined to the so-called *number-crunching* tasks. These include tasks such as forward and inverse modeling, and the processing of real time seismic data. The success of these methods relies on the many seismological observations that can be described mathematically in terms of source and path effects. However, some seismological problems are interpretational and depend on the ability to use concepts, relationships, and pattern recognition, as well as drawing upon the Center for Seismic Studies experience and the experience of others. The analysis of data from a regional seismic network involves both of these approaches. Numerical methods are used for such tasks as event detection and location, and interpretational, or hueristic, methods are used to distinguish real events from noise triggers, or local events from teleseisms, or quarry blasts from local tectonic events. The heuristic methods are generally undertaken by human analysts since the rules the analyst uses are difficult to quantify and thus difficult to program. However, as the volume of data handled steadily increases, it will become necessary to automate and program these heuristic rules so that the analyst's time and skills are used efficiently.

The goal of this project is to tap the skills and experience of human seismic analysts and use these for improved automated seismic analysis. Over the course of the next few Quarterly Reports, the Center will show its progress on a number of sub-projects involving the use of heuristic rules. These include:

- Improvements to detection processing algorithms including the use of three-component data to constrain azimuth and angle of incidence, as well as the use of Walsh transforms and frequency counters to discriminate regional and teleseismic events.

- Improvements to post-detection processing including pattern recognition to discriminate different types of events, and automatic association of events based on recognized moveouts and waveforms.

One of the tools being investigated for this project is the sonogram. A sonogram is essentially a plot of signal spectra versus time. Unlike an ordinary spectrum which is a transform over a single time interval, the sonogram provides a running spectrum along a seismic trace. An example of a sonogram for an event which occurred near Ottawa and recorded on station RSNY is shown in *Figure 3-8*. In many ways, the sonogram simulates the process an experienced seismic analyst goes through in analyzing a seismic event. The analyst looks at changes in amplitude and frequency along a seismic trace to identify an event and its associated phases. Much more information is contained in a sonogram than is necessary for an analyst to interpret a signal. The question becomes (1) what simple parameters does a human analyst visually extract from a seismic signal in order to identify and classify a seismic event and (2) can these parameters be extracted from the sonogram in order to characterize the event for future comparison with other events. Ideally,

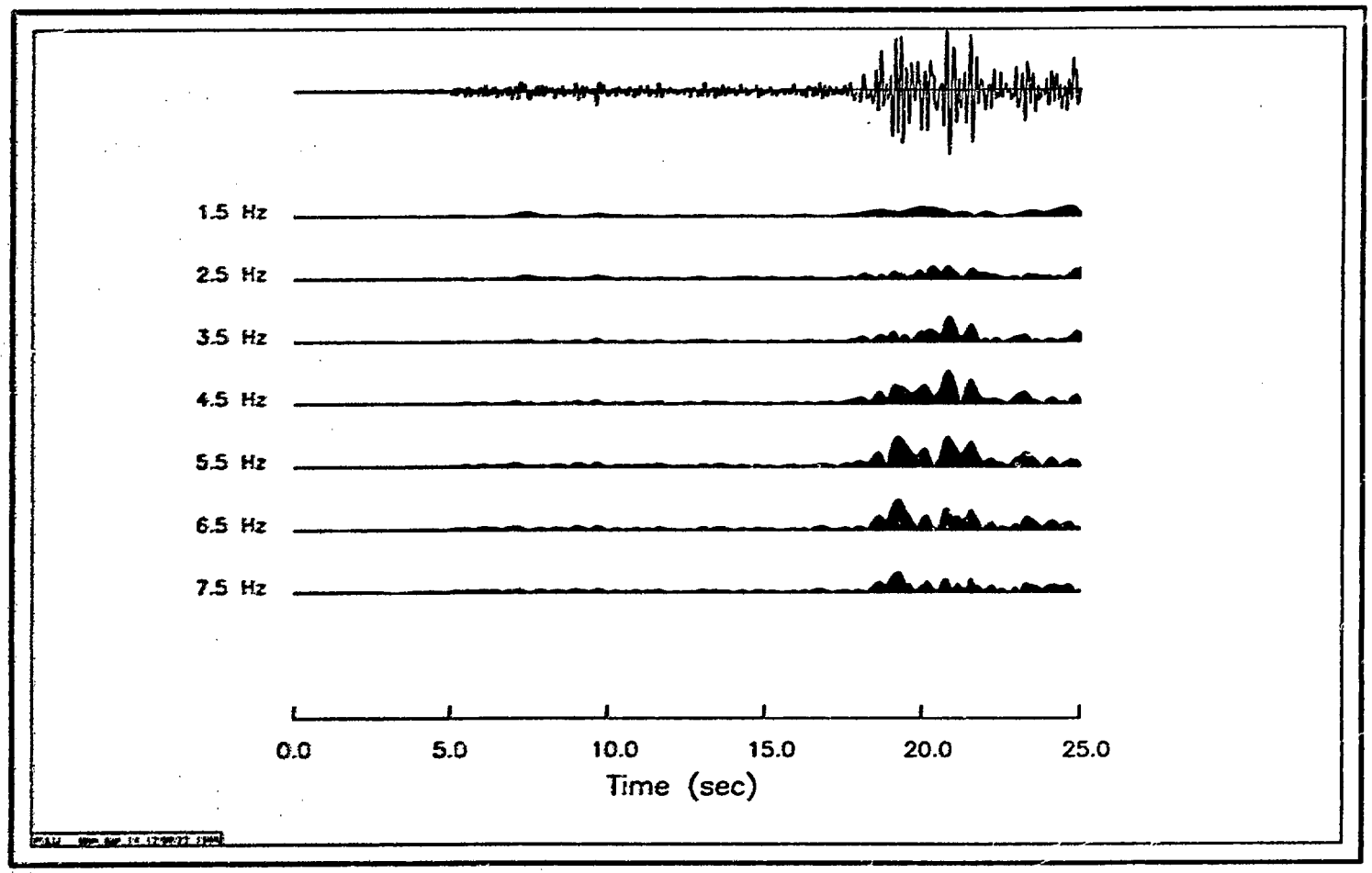


Figure 8-8. Sonogram of an earthquake which occurred near Ottawa and was recorded at station RSNY. The unfiltered seismogram is shown at the top. Spectral content for frequencies from 1.5 Hz to 7.5 Hz is shown below.

the computer would have a *dictionary* of event characteristics which it would use in a correlation process to identify a new event. Thus, the computer may eventually be able to automatically identify a seismic signal and have this identification available as part of the routine post-detection output.

Although the possibilities for heuristic processing of seismic signals are virtually limitless, the Center will be focusing on this automatic identification problem in the work and reporting on its progress in future Quarterly Reports.

Jay J. Pulli

3.3. REFERENCES

Flinn, E.A., "Signal Analysis Using Rectilinearity and Direction of Particle Motion." *Proc. IEEE,* Vol. 53, No. 12, pp. 1874, 1965.

Samson, J.C., and J.V. Olson, "Some Comments on the Descriptions of the Polarization States of Waves." *Geophys. Jour R.A.S.,* Vol. 61, No. 1, pp. 115, 1980.

Samson, J.C., and J.V. Olson,"Data-adaptive Polarization Filters for Multichannel Geophysical Data." *Geophysics,* Vol. 46, No. 10, pp. 1423, 1981.

Smart, E., "A Three-Component, Single-Station, Maximum-Likelihood Surface Wave Processor," Teledyne Geotech report SDAC-TR-77-14, 1977.

Smart, E., "Regional Phase Processors," Teledyne Geotech report SDAC-TR-81-1, 1981.

Suteau-Henson, A., G.H. Sutton, and J.A. Carter, "Polarization State Analysis of Three-Component and Array Seismograms," Roundout Associates Incorporated, manuscript in preparation, 1985.

4. STUDIES WITH REGIONAL SEISMIC DATA

4.1. LOCAL ATTENUATION AND SITE EFFECTS AT RSTN STATIONS

Knowledge of the spatial and frequency dependence of seismic wave attenuation in the earth serves two important purposes in seismology. First, it allows one to correct observed seismograms for propagation effects and retrieve information about the source strength. Second, it allows one to include realistic propagation effects in numerical simulations for comparison with observed waveforms. Attenuation can be measured with a number of different techniques, depending on the frequency band of interest. One common technique is to measure seismic wave amplitudes from a given source over a great circle path that includes two stations. However, the measured attenuation value will only be valid for that particular path. Any variations in attenuation which may exist along that path will be averaged into the final Q value. Additionally, such measurements at high frequencies often show large variations, which is often due to wave scattering. As seismology moves closer to handling realistic three-dimensional heterogeneous earth models, it becomes necessary to describe seismic parameters on a *point-by-point* basis, rather than by *path-averaged* approximations. One way to estimate seismic wave attenuation over short distances and at high frequencies is by measuring the amplitudes of scattered waves in a small volume surrounding the source and receiver. These scattered waves, known as *coda waves*, arrive at a recording site after the passage of the primary P- and S-waves and comprise the *tail* of a local or regional seismogram. Coda waves have been studied in a number of tectonically active and stable regions around the world and their behavior has been successfully used to estimate properties of the source, path, and recording site (*e.g.*, Aki and Chouet, 1975; Aki, 1980; Herrmann, 1980; Pulli, 1984; Phillips, 1985).

This study is aimed at determining the local attenuation and site effects at the RSTN stations RSNY, RSCP, RSON, and RSSD. These stations are located within or near zones of active seismicity and/or quarrying. Thus, numerous local sources are available for analysis. There are a number of models available for interpreting the coda waves generated by these local sources. All of these models can be expressed in the form:

$$A(f|\,t) = S(f) \cdot R(f) \cdot P(f|\,t)$$

Here, $A(f|\,t)$ is the RMS amplitude of the coda at frequency f and lapse time t (measured from the source origin time), $S(f)$ is the source effect, $R(f)$ is the receiver (or site) effect, and $P(f|\,t)$ is the path effect. A simple form of $P(f|\,t)$ is

$$P(f|\,t) = t^{-a}\, e^{-\pi f t / Q_c(f)}$$

where a equals 1.0 for body waves and 0.5 for surface waves. $Q_c(f)$ refers to the total Q of coda waves propagating in an ellipsoidal volume surrounding the source and receiver.

Since the path term $P(f|\,t)$ is the only one to depend on lapse time t, the logarithm of the RMS amplitude of the coda versus lapse time at a fixed frequency f is thus inversely proportional to $Q_c(f)$

$$\log_{10}[A(f|\,t)\cdot t^{\alpha}] = C - \frac{(\log_{10}e)\pi f}{Q_c(f)}t$$

C is a constant which equals $\log_{10}[S(f)\cdot R(f)]$. Thus, in order to determine $Q_c(f)$, the seismogram is first bandpass filtered. The RMS amplitude of the coda is calculated and multiplied by t^{α} and the slope of the logarithm of this quantity versus t is calculated.

In this study, four frequency bands have been considered. These bands are outlined in Table 4-1. The bands were chosen to be the same or similar to those chosen in the above referenced works to allow comparison of the results. An example of the procedure for the analysis of coda waves is illustrated in *Figure 4-1*. The event being analysed is an aftershock of the 7 October 1983 Goodnow, NY earthquake (main shock m_b of 5.2). This aftershock had m_b 4.2. The seismogram has been filtered through band number 4 with a center frequency of 12 Hz. After filtering, the envelope of the entire seismogram is calculated. Next, this envelope function is low-pass filtered below 0.2 Hz to provide smooth RMS amplitudes. For this example, the value of Q_c obtained is 1100. For the approximately dozen events examined so far at station RSNY, the measured values of Q_c obey the relationship

$$Q_c(f) = 750\, f^{0.2}$$

These values of Q_c are somewhat lower than the values of *Lg* Q obtained by Hasegawa (1985) and Gonca and Dean (1986) for eastern North America. Hasegawa's measurements indicate that the Q of *Lg* waves obeys the relationship $Q_{Lg}(f) = 900\, f^{0.2}$ while Gonca and Dean found that $Q_{Lg}(f) = 1000\, f^{0.35}$. A possible explanation for the lower

Table 4-1

FILTER PARAMETERS FOR CODA ANALYSIS

Band	Center Frequency (Hz)	Bandwidth (Hz)
1	1.50	1.00
2	3.00	2.00
3	6.00	4.00
4	12.00	8.00

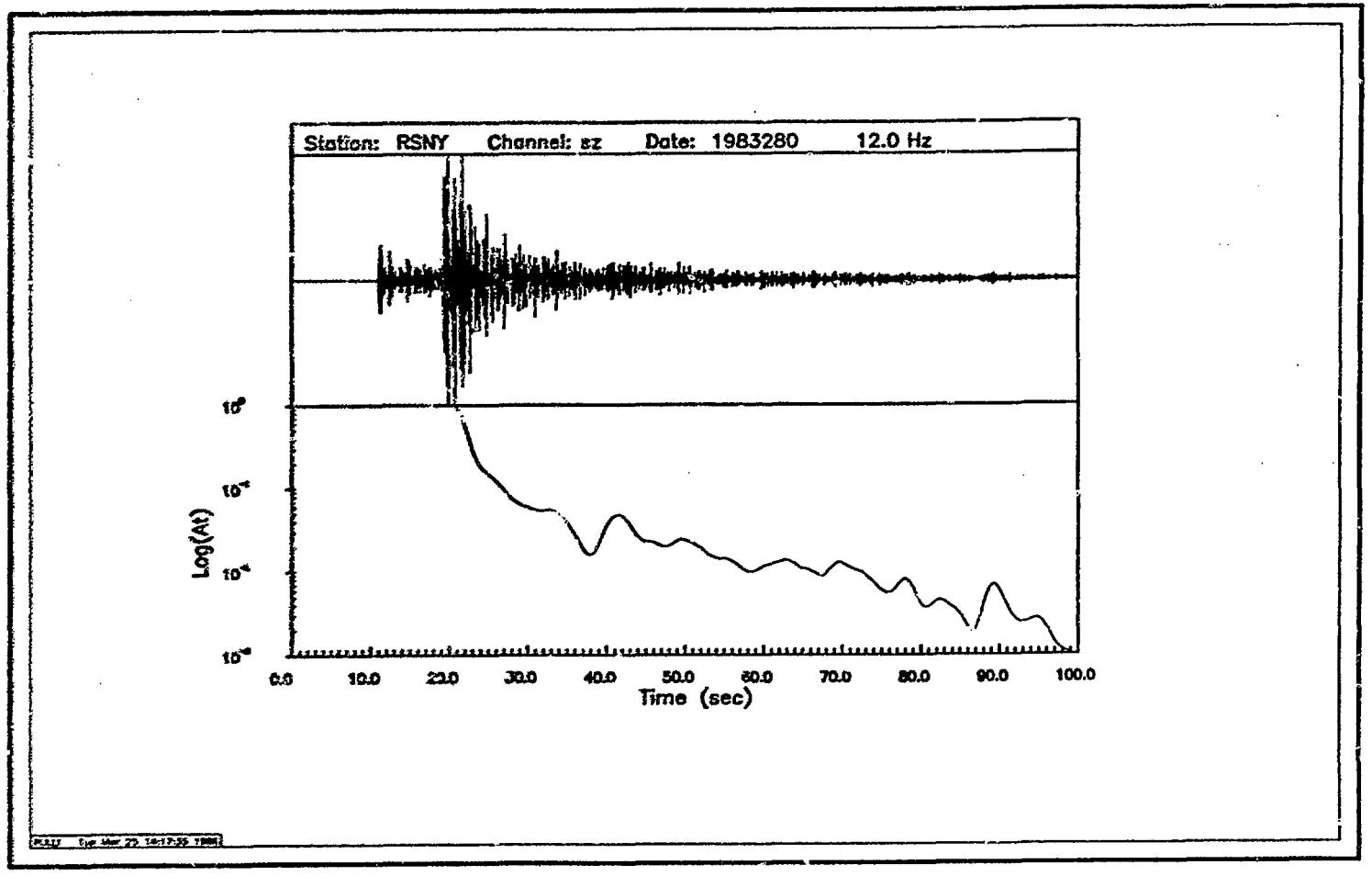


Figure 4-1. Example of coda analysis for a local earthquake. The event is an aftershock of the 1983 Goodnow, NY earthquake recorded at station RSNY. The seismogram has been bandpass filtered with a center frequency of 12 Hz. Below is shown the logarithm of the envelope function multipled by lapse time. The slope of this line is inversely proportional to Q_c at this frequency. A value of 1100 was obtained for this example.

values obtained so far in this study is that the Q values are measured over much smaller propagation distances. Studies of seismic wave attenuation in the Central and Northeastern United States have shown that larger values of Q are obtained when measurements are made over longer distances. Presumably, this is due to small variations in Q which are averaged out over long distances and lead to a bias toward the larger values. For the example show in *Figure 4-1* the measurement was made for lapse times between 30 and 90 seconds. This time interval corresponds to an ellipsoidal volume of about 200 km in lateral extent around the source and receiver.

As this project continues, the analysis method will be applied to the other RSTN sites and will compute both path effects and site effects. It should be mentioned that the theory of coda wave propagation and attenuation is one that is not completely understood at this time. Questions remain as to whether single scattering or multiple scattering constitutes the coda. In order to address these uncertainities, alternative models of the coda will be applied to the data and the results obtained from each model compared.

Jay J. Pulli

4.2. DISCRIMINATION OF QUARRY BLASTS FROM LOCAL EARTHQUAKES

Discrimination between quarry blasts and natural earthquakes at regional distances has long been a problem faced by seismic analysts. As regional seismic networks were installed across the United States, analysts gradually 'learned' to recognize the characteristics of seismic signals generated by quarry blasts and chemical explosions in their study areas. These recognizable characteristics, such as the absence of dilatational P-wave first motions and the presence of large-amplitude Rg waves, allowed experienced analysts to reduce the amount of time spent on such events and devote their efforts to the analysis of naturally occuring earthquakes.

Under a Comprehensive Test Ban Treaty (CTBT), vast amounts of seismic waveform data will have to be analysed, and derived seismic parameters must be quickly reported to central data facilities. Much of this effort can be handled by computers incorporating event detectors and post-detection processors. However, there will inevitably be reporting stations and small arrays located near areas of active quarrying. The automated systems should be able to discriminate between events from the quarrying areas and events of interest to CTBT monitoring. In addition, there is the issue of discriminating between chemical blasts and nuclear explosions at similar locations.

During the 1984 GSE Technical Test, the United States contributed seismic parameter data from six North American stations. According to US/GSE/37, among the analysis problems encountered, *the greatest procedural difficulty involved the analysis and reporting of local events. The sheer number of locals caused a serious analysis problem.* Later review showed that the majority of these local events were quarry blasts and chemical explosions. Analysts adopted procedures based on signal characteristics to identify these local events, but these procedures were invoked during the analyst review process and thus still involved considerable amounts of human intervention and time.

Automation of these procedures would be beneficial to the analyst's tasks.

A number of investigators have examined the problem of quarry blast versus local earthquake discrimination, with little success. For example, Gupta *et al.* (1984) examined the spectral characteristics of regional phases generated by quarry blasts, nuclear explosions, and small earthquakes and found large variability from one event to another. This large variablility prevented the establishment of a single discriminant or set of discriminants which could be universally applied to regional seismic data. In approaching this problem, the Center for Seismic Studies will not be looking for one discriminant which works in all areas. Instead, the Center is concentrating on the study of regionally dependent source and path effects which can be applied to specific stations for specific seismic zones.

This study reexamined the local waveform data for quarry blasts and regional earthquakes recorded during the GSE Technical Test. The starting point is the waveform data from the RSTN station RSSD near Rapid City, SD, chosen because of the large percentage of reported local events and the observed distinct characteristics of quarry blasts in this area. For example, *Figure 4-2a* shows the short-period vertical recording of a quarry blast which occured on 11 December 1984 at 22h 30m UTC. This blast was located approximately 200 km from RSSD. A striking feature of this recording is the low-frequency surface wave component arriving at a lapse time of about 90 seconds. These surface waves have an apparent group velocity of about 2 km/sec. *Figure 4-2b* shows the same recording, low-pass filtered to eliminate frequencies above 1.5 Hz. In this frequency band, these surface waves constitute much greater energy than the body wave phases. The P-wave phases are largely invisible in this passband. Likewise, *Figure 4-2c* shows the same recording high pass filtered to eliminate frequencies below 1.5 Hz. Note that only the body wave phases and their respective codas show up in this passband. A simple way to represent these distinct characteristics is to calculate the difference between these two passbands. Thus, subtraction of the high frequency envelope from the low frequency envelope produces a single characteristic which identifies this quarry. *Figure 4-2d* shows this difference of the envelope functions. Note that the body wave phases appear as negative amplitudes whereas the surface wave phases appear as positive amplitudes. Implementation of such a characteristic could be easily incorporated in the post-detection processor as a simple correlation.

Currently, a large number of regional events recorded at the RSTN stations during the GSETT are being examined. Although the characteristic observed above has also worked at station HSCP, other characteristics which may be unique to each station or seismic region are being investigated. This is equivalent to the so-called 'learning process' mentioned in the first paragraph. This learning process can be automated so that a station processor can accumulate knowledge automatically with time and determine its own characteristics with time. This is the ultimate goal for this project.

Jay J. Pulli
Antoinette Campanella
Michael Tiberio

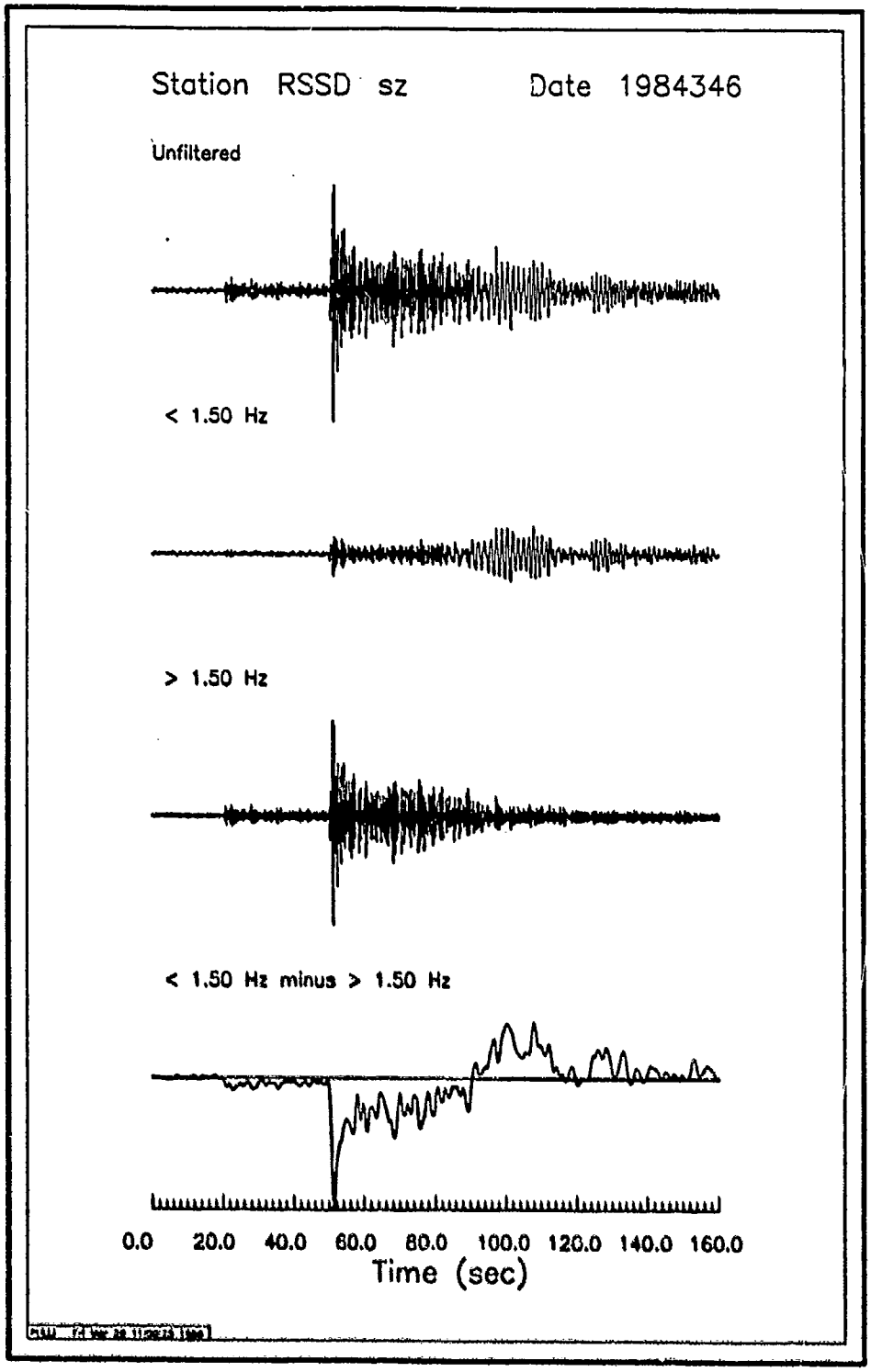


Figure 4-2. Recording of a quarry blast on the RSTN station RSSD. a) Unfiltered trace. b) Signal low pass filtered below 1.50 Hz. c) Signal high pass filtered above 1.50 Hz. d) Envelope of low pass filtered trace minus envelope of high pass filtered trace.

4.3. SEISMOGRAM MODELING CAPABILITY

To model the effects of source depth and propagation to regional distances, the Center for Seismic Studies has implemented an algorithm to obtain full wave solutions in horizontally layered media by the superposition of normal modes (Harvey, 1981). A partial evaluation of the major computational program **FINDPOLE** has been made for application to various problems currently being addressed by the Center research staff. In addition, some synthetic seismograms have been generated to illustrate the effect of varying source depth.

Table 4-2 lists the input parameters and computer requirements for a number of simulations in a 4-layer New England crustal model (Foley *et al.*, 1985), a 3-layer model derived for the Canadian Shield (Taylor, 1985), and a similar layer over halfspace. Table 4-3 lists the layer parameters for each crustal model.

As expected, the run times in VAX 780 cpu hours increase approximately as the square of the maximum frequency (fmax). Increasing the depth to the cap and the maximum time window (tmax) has an almost linear effect on the run times in these ranges. Little is to be gained by restricting the solution to include only P-SV modes, since its run time is comparable to that for the complete solution. The presence of the cap layer, which is necessary to the locked mode approximation for the inclusion of body wave phases, can be eliminated for conventional surface wave calculations. Consequently, run times are reduced considerably.

Table 4-2

FINDPOLE parameters

Model	Cap depth (km)	modes	tmax (sec)	fmax (hz)	cpu (hrs)	output size (mbytes)
NE	54	psvsh	100	2	1	0.25
NE	54		100	5	5.6	1
NE	54		200	2	1.9	0.5
NE	135		100	10	41	12
NE	135		100	5	10	3
NE	135		100	2	2.2	0.5
NE	54		100	10	21	5
CS	125	psvsh	100	2	1.2	3.5
CS	125	psv	100	2	0.8	0.2
CS	125	sh	100	2	0.3	0.2
LOHC	136	psvsh	100	2	1.1	0.5
LOHC	136	psv	100	2	0.9	0.3
LOHC	136	sh	100	2	0.3	0.2
LOH		psvsh	100	2	0.2	0.1
LOH		sh	100	2	0.05	0.05

Table 4-3			
Velocity models			
New England (NE)			
Vp (km/sec)	Vs (km/sec)	Density (gm/cc)	Thickness (km)
5.31	3.07	2.75	0.88
6.06	3.5	2.75	12.21
6.59	3.8	2.8	21.51
8.1	4.68	3.3	20.0
16.0	9.0	3.5	cap
Canadian Shield (CS)			
6.09	3.52	2.75	10.0
6.66	3.85	2.8	30.0
8.04	4.65	3.3	85.0
16.0	9.0	3.5	cap
Layer over halfspace LOHC and LOH (no cap)			
6.4	3.7	2.75	36.0
8.1	4.7	3.3	100.0
16.0	9.0	3.5	cap

Figure 4-9 shows vertical seismograms for the first case in Table 4-2. The vertical channels shown here for 4 depths have been filtered by the RSTN short period instrument response. The relative amplitudes are plotted here to demonstrate the dramatic effect of source depth on the amplitudes of L_g and R_g. Recent studies (Eissler and Kanamori, 1985; Kafka, 1985) have shown that the amplitudes of these phases relative to the body wave amplitudes are extremely sensitive to source depth, particularly for events in the upper 3 to 5 km of the crust.

The capability to compute full-wave seismograms represents an important analysis tool for interpreting the complex wave features observed at regional distances. Whereas recorded motions contain the effects of geologic irregularities within the crust and upper mantle, for most cases the recorded motions are dominated by seismic phases that result from horizontal wave guides in the earth. By assigning representative regional properties to layered-earth models, synthetic seismograms can be made that closely resemble recorded motions (*e.g.*, Herrmann and Nuttli, 1975; Bache et al., 1980; Bache *et al.*, 1981). This capability is useful for interpreting recorded events and serves to further establish reliabilities and deficiencies inherent in the idealized model. The capability is also useful for revealing trends to be expected in regional observations. The Center is currently applying the full-wave methods for interpreting key events in an effort to explain observed signals and determine source parameters. Also, in accordance with the Center's research plan, studies are underway to seek improved discriminants for determining source depth from regional observations.

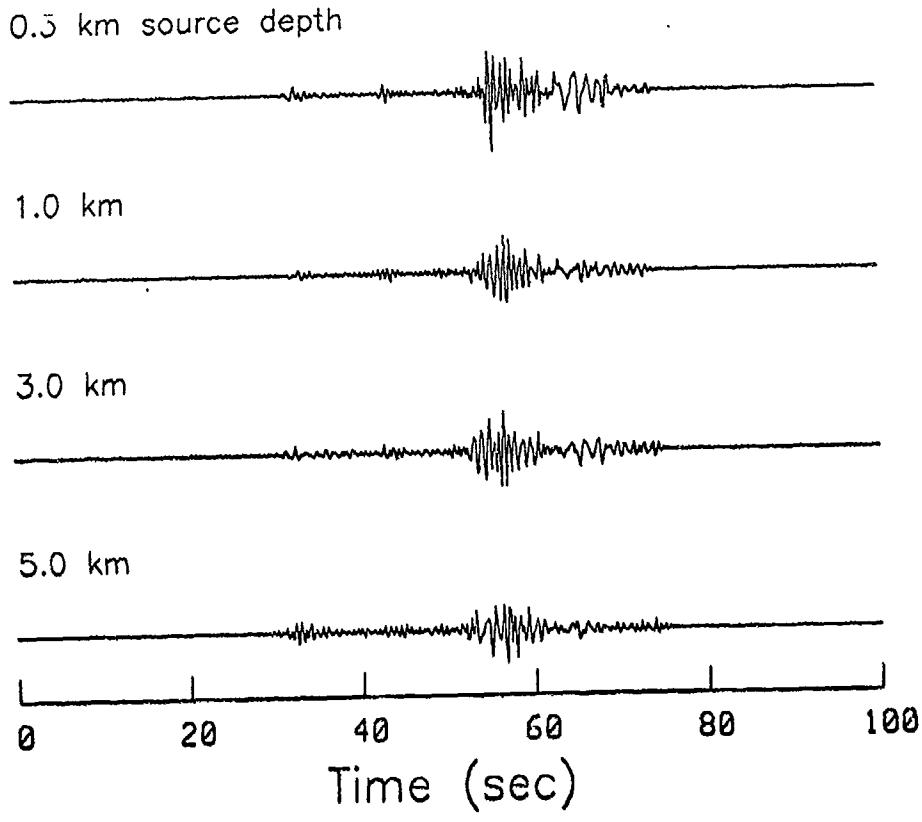


Figure 4-9. **Calculated vertical response of an RSTN short period instrument to a vertical strike-slip source at a distance of 200 km and depths ranging from 0.5 to 5 km in the New England earth structure of Table 4-3.**

Simulations are currently being made for New England and Canadian Shield crustal models to assist in projects relating to:

1. Evaluation of propagation and source effects on recordings of the 15 October 1985 Ohio and 19 October 1985 New York earthquakes at station RSNT.

2. m_b and L_g-magnitude biases between NTS and East Kazak nuclear explosions.

3. Discrimination of quarry blasts from earthquakes recorded at RSTN stations.

4. Modal analysis of surface waves at regional distance for depth information.

5. Modeling of regional phases recorded at RSTN and NORESS.

Paul Dysart
Gerald Frazier

4.4. MAGNITUDE SCALING OF HIGH-FREQUENCY EARTHQUAKE SPECTRA

4.4.1. Introduction

The smallest events recorded at regional distances generally have dominant frequencies above 1 Hz. The maximum signal-to-noise ratio for body phases along high Q paths is expected to occur near the corner frequency of the source because the displacement amplitude spectrum for ground noise in this frequency range (f > 1 Hz) decays roughly as inverse frequency squared. Consequently, the magnitude dependence of the source near the corner frequency is directly relevant to event detection. Also, high-frequency signals are potentially useful for determining source parameters (*e.g.*, depth and variable frequency magnitude), and discriminating between earthquakes and explosions. For example, spectral differences between earthquakes and explosions at frequencies above the corner frequencies have been proposed as a possible discriminant by Evernden *et al.* (1985).

This report reviews the high-frequency response of earthquakes in an effort to identify trends and to evaluate those characterizations used by Evernden *et al.* (1986). First, source models derived from idealized representations of the fracture process are reviewed and some apparent deficiencies are noted. To resolve these deficiencies, an alternate scaling relation that increases the sensitivity of high-frequency motion to earthquake magnitude is proposed. High-frequency recordings have become much more abundant in recent years, particularly at close distances, and the interpretation of these data is actively being pursued by several researchers. Review of these studies indicates that the high-frequency properties of earthquakes are not well understood. Some researchers speculate that, in many cases, the corner frequencies observed for small earthquakes may frequently contain more information about the recording site than about the size of the source. Also, the rate of decay of the high-frequency asymptote is still uncertain, although decay with inverse

frequency squared is usually assumed.

4.4.2. Source Models

Conventional earthquake models produce far-field spectral amplitudes that approach two asymptotes, one at high frequencies and one at low frequencies, as illustrated in *Figure 4-4*. The mid-frequency at which these asymptotes intersect is termed the corner frequency, f_c. The low-frequency asymptote is flat in displacement amplitude, A(f), and the high-frequency asymptote decays as the inverse square or cube of frequency depending on the model. This general behavior can be expressed by

$$A(f) = \frac{A(0)}{(1+f/f_c)^p} \tag{1}$$

where p = 2 or 3 denotes the model-dependent decay rate of the high-frequency asymptote. This simplified relation represents only the asymptotic form of the source spectrum and does not include model-dependent details of the source such as azimuthal variations resulting from radiation pattern and directivity of the dynamic fracture.

The far-field displacement averaged over azimuth is proportional to the moment rate of the source, and hence the low-frequency asymptote A(0) scales directly with seismic moment, M_0. The moment is given by

$$M_0 = (\textit{shear modulus})\,(\textit{source area})\,(\textit{average offset}) \tag{2}$$
$$\approx \Delta\sigma\, L^3$$

where $\Delta\sigma$ is the static stress drop, L is the characteristic dimension of the source, and $\approx$ denotes proportionality. The length and width are assumed to be of comparable size. The average offset then scales with the product of stress drop and source dimension to obtain the above scaling relation for seismic moment (Kanamori and Anderson, 1975).

The corner frequency at the intersection of the two asymptotes corresponds to a time constant for the idealized fracture. Earthquakes may likely have several characteristic times including:

1. source duration, t_L

2. slip duration at individual points (rise time) or time separation of fracture irregularities, t_r

3. local transition time for the individual fracture irregularities, t_δ

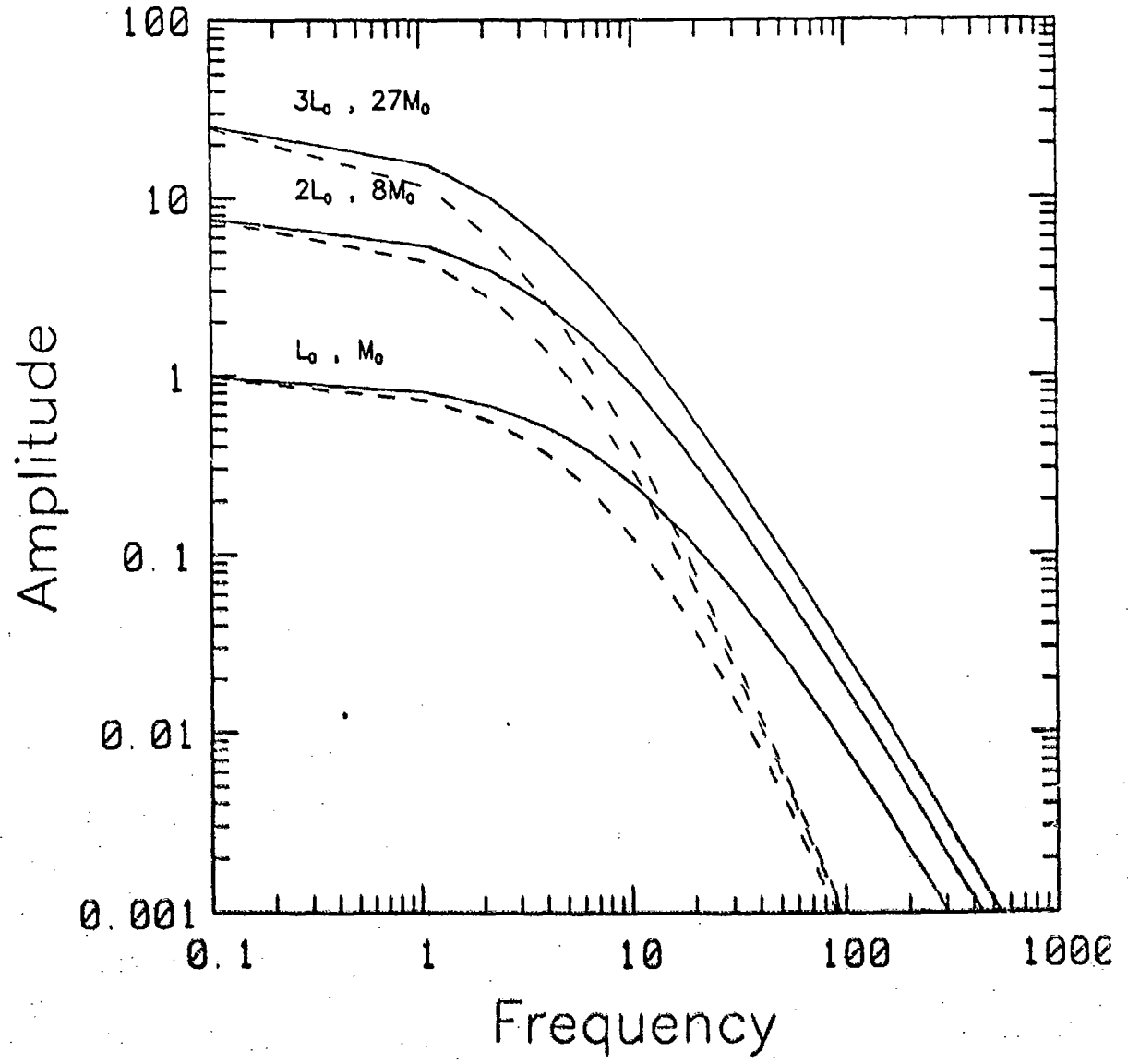


Figure 4-4. **Spectral asymptotes for conventional frequency-cubed and frequency-squared source models.**

For most idealized source models the time constant is the duration time for source fracture, which scales with source dimension, *i.e.*,

$$f_c \approx L^{-1} \tag{3}$$

The scaling relations presented in (2) and (3) are substituted into the general asymptotic form of (1) to obtain the scaling relation for the high-frequency asymptote

$$A(f \gg f_c) \approx \Delta\sigma \frac{L^3}{(fL)^p}. \tag{4}$$

First consider models that produce spectral amplitudes decaying at high frequencies as inverse frequency cubed. Substituting p = 3 into (4) gives

$$A(f \gg f_c) \approx \Delta\sigma\, f^{-3}. \tag{5}$$

That is, the frequency-cubed models with a constant stress drop produce the same high-frequency radiation independent of earthquake size. Archambeau's model (1972) used by Evernden *et al.* (1986) has this characteristic for the P-wave radiation. As pointed out by Aki and Richards (1980, p 822), this model does not appear to explain the entire radiated signal. To compensate for the longer durations associated with larger earthquakes, the amplitude of the high-frequency time signal would actually have to decrease with increasing magnitude.

This point is expanded as a basis for subsequent comments on the scaling of high-frequency motions. The time envelope of the far-field radiation from the source, narrow band filtered at frequency $f \gg f_c$, is approximated by a quadratic time function, *i.e.*,

$$U(t, f \gg f_c) = 2U_{max}\, t_L^{-2}\, t\, (t_L - t) \tag{6}$$

where $t_L = f_c^{-1} \approx L$ is the source duration. This function, illustrated in *Figure 4-5*, depicts the gradual growth and decay in fault area that contributes to the signal as a function of time. The spectral amplitude is then the area under the narrow band signal envelope:

$$A(f \gg f_c) = \int_0^{t_L} U(t, f \gg f_c)\, dt$$

$$= \frac{1}{3}\, t_L\, U_{max} \tag{7}$$

$$\approx L\, U_{max}$$

For $A(f \gg f_c)$ independent of L, as in the frequency-cubed models with constant stress drop (5), the implausible result is that the high-frequency envelope U_{max} must decrease with increasing source size.

It appears that earthquake models with a high-frequency asymptote proportional to inverse frequency cubed do not include the effects of abrupt changes in fracture velocity. Inclusion of a stopping phase in idealized source models results in a high-frequency asymptote proportional to inverse frequency squared. For these models with p = 2 and $f_c \approx L^{-1}$ the high-frequency asymptote (4) becomes

$$A(f \gg f_c) \approx \Delta\sigma \, L \, f^{-2} \tag{8}$$

Here the high-frequency asymptote scales directly with the source dimension. The maximum amplitude of the high-frequency time signal U_{max} from (7) is now invariant of the earthquake size, but the initial amplitudes ($t \ll t_L$) still decrease with increasing size to accommodate the increases in source duration with increasing magnitude. This behavior, illustrated in *Figure 4-5*, is still not consistent with observations and appears to constitute a serious deficiency.

As mentioned before, it is the stopping phase that causes the inverse frequency-squared asymptote. The stopping phase occurs along the exterior perimeter of the simulated fracture, which scale linearly with the source dimension; hence, the spectral amplitude of high frequencies scales linearly with source dimension. It seems much more likely that irregularities in the dynamic fracture are distributed over the entire source area. As discussed below, provision for small scale irregularities in the fracture process increases the sensitivity of the high-frequency motions to source dimension.

Several studies have demonstrated that the recordings of strong ground motions from potentially damaging earthquakes (M > 5.5) can be modeled by combining the recordings from small earthquakes in sufficient number to match the moment of the large event (*e.g.*, Hartzell, 1978; Hadley and Helmberger, 1980; Hadley *et al.*, 1982). Time delays for the small earthquakes are assigned to mimic the gross fracture velocity of the large earthquake. With this procedure, the very low-frequency motions add together in phase, while the high-frequency motions combine with essentially random phase. If N small earthquakes of length $\tilde{L}$ and moment $\tilde{M}_0$ are used to construct one large earthquake of length $L > \tilde{L}$ then moment $M_0 = N \tilde{M}_0$. Note from (2) that the moment and the low-frequency amplitudes scale with length cubed to give

$$N = \frac{M_0}{\tilde{M}_0} = \left(\frac{L}{\tilde{L}} \right)^3 \tag{9}$$

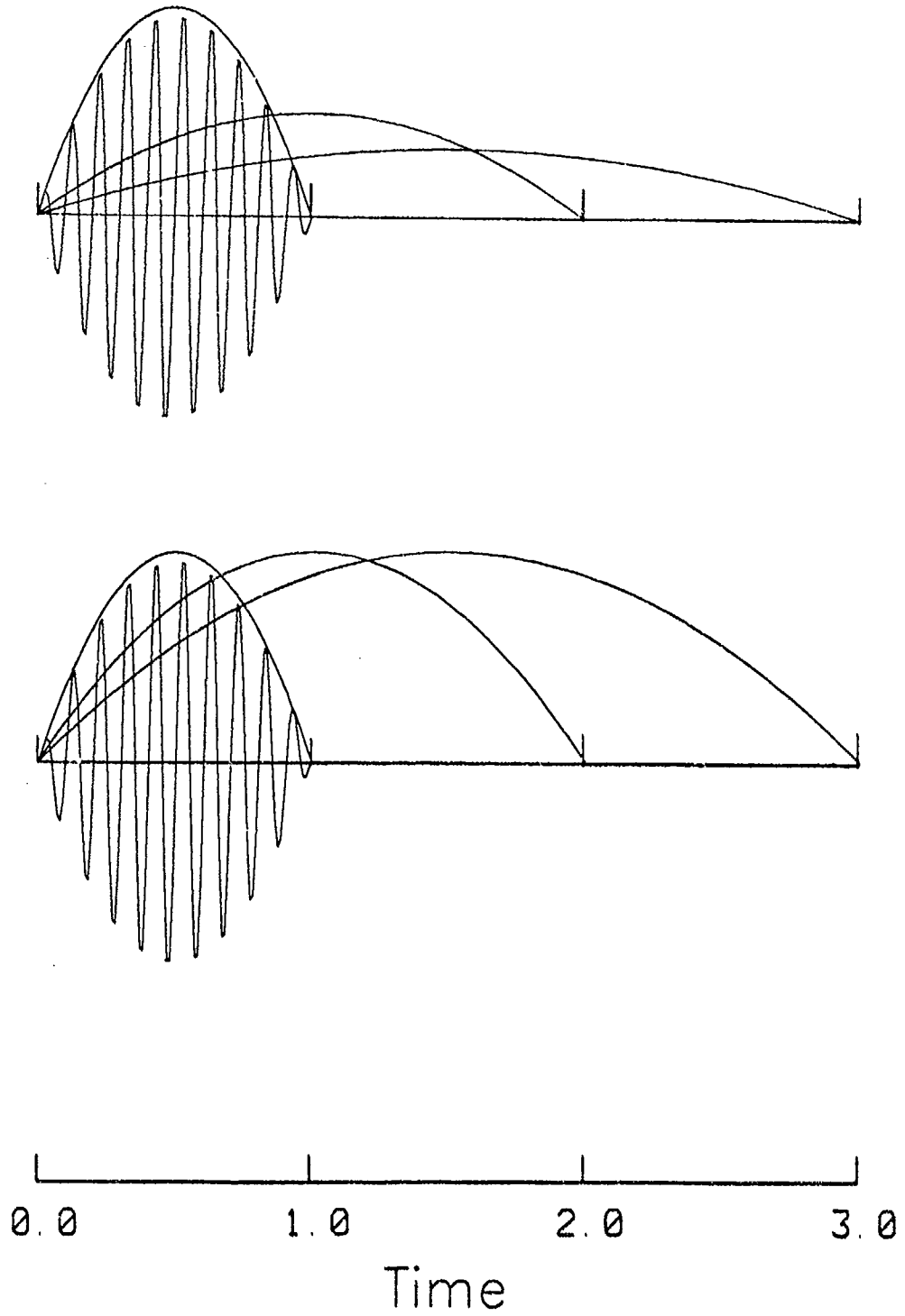


Figure 4-5. Time envelope of narrow band far field source radiation for $f \gg f_c$ for the two spectral source models shown in *Figure 4-4*.

The very low-frequency amplitudes for the small earthquakes $\tilde{A}(f \ll f_c)$ combine in phase to produce the low-frequency amplitudes for the large earthquake, *i.e.*,

$$A(f \ll f_c) = N\,\tilde{A}(F \ll \tilde{F}_c) \qquad (10)$$
$$= \left(\frac{L}{\tilde{L}}\right)^3 \tilde{A}(f \ll f_c)$$

In contrast, the high-frequency amplitudes combine with random phase to give

$$A(f \gg f_c) = \sqrt{N}\,\tilde{A}(f \gg \tilde{f}_c) \qquad (11)$$
$$= \left(\frac{L}{\tilde{L}}\right)^{3/2} \tilde{A}(f \gg f_c)$$

Hence, there exists a reasonable basis for considering that the high-frequency amplitudes for earthquakes scale with length to the three-halves power. This scaling relation differs from that derived for both the frequency-cubed models (5) and the frequency-squared models (8). In addition, the degree of coherence or randomness of high-frequency phase may provide a magnitude-dependent parameter which effects the scaling of corner frequency with source dimension and stress drop. Introducing the result of (8) into (4) yields

$$f_c \approx \left(\frac{L^{q-3}}{\Delta\sigma}\right)^{1/q} \qquad (12)$$

where q is the scaling power of L with high-frequency amplitude. For a constant stress drop and q=3/2 this relation requires a scaling of $f_c \approx L^{-3/4}$.

4.4.3. Observations

A study of the Mammoth Lakes earthquake sequence (Archuleta *et al.*, 1982) showed a scaling of stress drop with seismic moment. This and other studies (*e.g.*, Chouet *et al.*, 1978; Rautian *et al.*, 1978; Bakun *et al.*, 1976) suggests that the concept of similarity among earthquakes (Aki, 1967) does not apply to events in the low magnitude range (M_L < 5). The idea of asperities and barriers (Aki, 1984; Das and Aki, 1977; Takeo, 1983) has been useful in explaining some of the observations by providing mechanisms which constrain the characteristic dimensions of the source region while allowing variable slip functions and high-frequency behavior. However, these mechanisms, although active in some regions, cannot be universally applied to all small events.

The independence of the various parameters used to describe the earthquake source and the practical ability to measure them from high-frequency recordings has recently come under question. The effects of site response and near-receiver attenuation have been found to have a severe effect on the shape of high-frequency spectra (Cranswick *et al.*, 1985; Dysart, 1985; Frankel, 1982). Consequently, observed scaling relations are sometimes spurious and not related to the physical parameters implied by a single source model or scaling law. These effects, together with a typically wide physical scatter in the estimated parameters, especially stress drop estimates, make observed scaling relations from high-frequency data difficult to interpret.

Some efforts have been made to improve source parameter estimates by limiting the number of independent source parameters, correcting for site response, and using more stable estimation procedures. A recent study by Andrews (1985) has shown seismic spectra which are more consistent with conventional scaling laws. Using only two independent source parameters computed from spectral integrals he made the following observations in his study of the 1980 Mammoth Lakes sequence.

1. The spectral shape is independent of earthquake size, and adequately described using seismic moment and corner frequency.

2. An f^{-2} falloff rate above the corner frequency provides a good fit to all the spectra.

3. There is no clear correlation between stress drop and size.

4. There is a strong correlation between Brune (1970) and Hanks (1981) stress drops.

5. Stress drop has a log-normal distribution of stress drops with a standard deviation of 2.

These most recent observations support the concept of similarity among earthquakes, and argue strongly in favor of a frequency-squared source model with constant stress drop, and the scaling of high-frequency amplitudes with source dimension as shown in near-field studies of peak ground accelerations (*e.g.*, Campbell, 1981).

4.4.4. Summary

Simplified earthquake models in the literature fail to explain important observations of earthquakes, such as the increase of the high-frequency time envelope with increasing magnitude, and the scaling relationship between high-frequency amplitude and source dimension. The frequency-cubed model used by Evernden *et al.* (1986) appears to require high-frequency motion to increase with decreasing magnitude for a constant stress drop. Their model therefore opposes the weight of earthquake observations which points to a ω^{-2} falloff rate, and a scaling of the high-frequency envelope with magnitude. This report presented a basis for scaling high-frequency amplitudes with the source dimension to the three-halves power which partially alleviates this dilemma. Spectral amplitudes at high frequencies are expected to decay with inverse frequency-squared as the result of irregularities in the fracture process.

Earthquake spectra are not easily observed at high frequencies because of the severe effects introduced by wave transmission. Near-field observations of the Mammoth Lakes earthquakes demonstrate high-frequency decay as inverse frequency squared (Andrews, 1985). The stress drop estimates for these earthquakes appear to be log- normally distributed with no obvious correlation to corner frequency.

The near-field measurements of peak accelerations suggest a scaling of high-frequency amplitudes proportional to some power of L. Similarly, the success in modeling large earthquakes from the superposition of smaller events points to a scaling of high-frequency amplitudes with a greater power of source dimension than implied by the conventional frequency-squared or frequency-cubed models.

Gerald Frazier
Paul Dysart

4.5. REFERENCES

Aki, K., "Asperities, Barriers, Characteristic Earthquakes and Strong Motion Prediction." *J. Geophys. Res.*, 89, pp. 5867-5872, 1984.

Aki, K., "Scattering and Attenuation of Shear Waves in the Lithosphere." *J. Geophys. Res.*, 85, pp. 6496-6504, 1980.

Aki, K. and Chouet, B., "Origin of Coda Waves: Source, Attenuation, and Scattering Effects." *J. Geophys. Res.*, 80, pp. 3322-3342, 1975.

Aki, K. and P. G. Richards, *Quantitative Seismology: Theory and Methods*, W. H. Freeman, San Francisco, 1980.

Andrews, D. J., *Objective Determination of Source Parameters and Similarity of Earthquakes of Different Size*, in press, 1985.

Archambeau, C. B., "The Theory of Stress Wave Radiation from Explosions in Prestressed Media." *J. R. Astro. Soc.*, 29, pp. 329-366, 1972.

Archuleta, R. J., E. Cranswick, C. Mueller and P. Spudich, "Source Parameters of the 1980 Mammoth Lakes, California, Earthquake Sequence." *J. Geophys. Res.*, 87, pp. 4595-4607, 1982.

Bache, T. C., H. J. Swanger, and B. Shkoller, *Synthesis of L_g in Eastern United States Crustal Models*, System, Science and Software Semi-annual Technical Report SSS-R-81-4668, September 1980.

Bache, T. C., H. J. Swanger, B. Shkoller and S. M. Day, *Simulation of Short Period L_g, Expansion of Three-dimensional Source Simulation Capabilities and Simulation of*

Near-field Ground Motion from the 1971 San Fernando, California Earthquake, System, Science and Software Final Technical Report SSS-R-81-5081, July 1981.

Bakun, W. H., C. G. Bufe and R. M. Stewart, "Body Wave Spectra of Central California Earthquakes." *Bull. Seismol. Soc. Am.*, 66, pp. 363-334, 1976.

Brune, J. N., "Tectonic Stress and the Spectra of Seismic Shear Waves from Earthquakes." *J. Geophys. Res.*, 75, pp. 4997-5009, 1970.

Campbell, K., "Near-source Attenuation of Peak Horizontal Acceleration." *Bull. Seismol. Soc. Am.*, 71, pp. 2039-2065, 1981.

Chouet, B., K. Aki and M. Tsujiura, "Regional Variation of the Scaling Law of Earthquake Source Spectra." *Bull. Seismol. Soc. Am.*, 68, pp. 49-80, 1978.

Cranswick, E., R. Wetmiller and J. Boatwright, "High-frequency Observations and Source Parameters of Microearthquakes Recorded at Hard-rock sites." *Bull. Seismol. Soc. Am.*, 75, pp. 1535-1567, 1986.

Das, S. and K. Aki, "Fault plane with Barriers: a Versatile Earthquake Model." *J. Geophys. Res.*, 82, pp. 5648-5670, 1977.

Dysart, P. S., *Moment-Radius-Stress Drop Relations and Temporal Changes in the Regional Stress from the Analysis of Small Earthquakes in the Matsushiro Region, Southwest Honshu, Japan,* Ph. D. Dissertation, Virginia Polytechnic Institute and State University, Blacksburg, Va., 1985.

Evernden, J. F., C. B. Archambeau and E. Cranswick, "An Evaluation of Seismic Decoupling and Underground Nuclear Test Monitoring Using High Frequency Seismic Data." manuscript submitted to *Rev. Geophys*, 1986.

Foley, J. E., C. Doll, F. Filipkowski and G. Lorsbach, *Northeastern U. S. Seismic Network Bulletin*, No. 34, Weston Observatory, Dept. of Geology and Geophysics, Boston College, 1985.

Frankel, A., "Source Parameters and Scaling Relationships of Small Earthquakes in the Northeast Caribbean." *Bull. Seismol. Soc. Am.*, 71, pp. 1173-1190, 1981.

Frankel, A., "The Effects of Attenuation and Site Response on the Spectra of Microearthquakes in the Northeast Caribbean." *Bull. Seismol. Soc. Am.*, 72, pp. 1379-1402, 1982.

Gonca, J.H. and W.C. Dean, *Lg Earthquake Spectrum Function and Q for Eastern North America from Regression of RSTN Data,* in preparation, 1986.

Hadley, D. M., D. V. Helmberger, and J. A. Orcutt, "Peak Acceleration Scaling Studies."

Bull. Seismol. Soc. Am., 72, pp. 959-979, 1982.

Hadley, D. M. and D. V. Helmberger, "Simulation of Strong Ground Motions." *Bull. Seismol. Soc. Am.*, 70, pp. 617-630, 1980.

Hanks, T. C. and R. K. McGuire, "The Character of High Frequency Ground Motion." *Bull. Seismol. Soc. Am.*, 71, pp. 2071-2096, 1981.

Hartzell, S. H., *Interpretation of Earthquake Strong Ground Motion and Implications for Earthquake Mechanisms*, Ph.D. Dissertation, Univ. of Cal., San Diego, La Jolla, Calif, 1978.

Harvey, D. J., "Seismogram Synthesis Using Normal Mode Superposition: the Locked Model Approximation." *Geophys. J. R. Astr. Soc.*, 66, pp. 37-59, 1981.

Hasegawa, H., "Attenuation of Lg Waves in the Canadian Shield." *Bull. Seis. Soc. Amer.*, 75, pp. 1569-1582, 1985.

Herrmann, R.B., "Q Estimates Using the Coda of Local Earthquakes." *Bull. Seis. Soc. Amer.*, 70, pp. 447-465, 1980.

Herrmann, R. B. and O. W. Nuttli, "Ground-motion Modelling at Regional Distances for Earthquakes in a Continental Interior, I. Theory and Observations." *Int. J. Earthq. Engng. Struct. Dyn.*, 4, pp. 49-58, 1975.

Kanamori, H. and D. Anderson, "Theoretical Basis of Some Empirical Relations in Seismology." *Bull. Seismol. Soc. Am.*, 65, pp. 1073-1095, 1975.

Phillips, W.S., *The Separation of Source, Path, and Site Effects on High Frequency Seismic Waves: An Analysis Using Coda Wave Techniques*, PhD Thesis, MIT, pp. 195, 1985.

Pulli, J.J., "Attenuation of Coda Waves in New England." *Bull. Seis. Soc. Amer.*, 74, pp. 1149-1166, 1984.

Randall, G. E. and S. R. Taylor, *Inversion of Rayleigh Wave Phase Velocities*, Lawrence Livermore National Laboratory, UCID-20504, pp. 13, 1985.

Takeo, M., "Source Mechanisms of the Usu Volcano, Japan, Earthquakes and Their Tectonic Implications." *Phys. Earth Planet. Int.*, 32, pp. 241-264, 1983.

5. YIELD ESTIMATION

5.1. YIELD ESTIMATION USING $m_b(L_g)$

In support of the DARPA Seismic Review Panel, classified yield and P-wave magnitude data were assembled for use in analyzing Nuttli's $m_b(L_g)$ determinations for yield estimation. In addition, WWSSN film records from selected Eurasian stations were collected for most of the recent Shagan Test Site events. These records were provided to Nuttli for his use in expanding the database of $m_b(L_g)$ determinations for important Soviet events, to provide a broader base for comparing $m_b(L_g)$ yield estimates with those from isotropic moment (M_0) measurements as well as conventional P-wave magnitude (m_b) measurements.

Data from Rondout Associates for NTS Lg magnitudes were also considered, as were a number of Lg magnitude determinations for Shagan events made by Alexander from GRFO records.

An $m_b(L_g)$-yield relation for NTS was derived from the original Nuttli data set. This relationship was used to predict the yield of approximately 30 of the larger Shagan events. Using the yields predicted from Nuttli's $m_b(L_g)$, an $m_b(P)$-yield relationship was derived for Shagan for direct comparison with the official AFTAC formula to arrive at an estimate of the formula bias. This information was used by both the DARPA and AFTAC Panels in arriving at their estimates of the formula bias between NTS and the Shagan River Test Site.

Analysis of the combined $m_b(L_g)$, $m_b(P)$ and M_0 data for United States and Soviet explosions is continuing, in an attempt to identify systematic differences between the test sites which might influence the accuracy of the yield estimates of Soviet tests based on NTS $m_b(L_g)$-yield relations. Results of this research will be reported in a separate classified report to be issued for the first quarter of 1986.

George Bulin

5.2. ESTIMATES OF SEISMIC COUPLING AND NUCLEAR EXPLOSION YIELDS USING SEISMIC MAGNITUDE AND MOMENT DATA

Observations of the spectral character of direct compressional waves from underground nuclear explosions, both in the near and far field distance ranges, suggest that the purely explosive-produced P-wave has a form that can be represented by a simple step function pressure equivalent. Here, however, the "elastic radius" for the equivalent source is significantly larger than the radius at which high strain flow and fracture effects cease, and corresponds to the (mean) radius where strain-dependent losses, at relatively low strain levels in the range 10^{-4} to 10^{-6}, cease.

Using this simple equivalent source representation and the assumption that the (reversable) work done on the medium outside the elastic radius is proportional to yield, it

follows that m_b values depend directly on yield and upon two source-dependent coupling parameters K_c and K_E, which are termed the non-linear and elastic coupling parameters, respectively.

On the other hand, M_s is dependent only on yield and the non-linear coupling parameter K_c, while the moment M_0 is simply related to M_s and is dependent only on K_c, and the Poisson ratio at the source elastic radius, as well as yield. The coupling parameter K_E may be expressed in terms of the elastic velocities at the elastic radius of the source, while K_c is strongly dependent on a variety of local material properties, including porosity, water content, pressure, and the fracture density within the material surrounding the explosion.

It is found that the available magnitude-yield data for United States nuclear explosions, in a wide variety of media, require four sets of the coupling constants K_c and K_E for a simultaneous description of both the m_b and M_s observations. The resulting empirical magnitude-yield curves correspond to a high-coupling set for m_b and M_s, two intermediate sets, and a low-coupling set. The high-coupling set is shown on *Figure 5-1* as an example.

These magnitude-yield curve sets provide a means of estimating yield using both m_b and M_s together, provided tectonic release effects and differences in test site absorption, relative to NTS, are taken into account. On the other hand, if the range of coupling at some other test site is assumed to be within the range sampled by the United States data, then it is possible to estimate the (combined) difference in near-site absorption and tectonic effects between NTS and any other site.

This inference method is applied to Soviet explosion data, with the objective of first estimating absorption and tectonic release differences. Both (m_b, M_s) and (m_b, M_0) data indicate a difference in absorption between NTS and the Kazakh test site, which is such that observed m_b values from Kazakh will be larger, by 0.31 - 0.35 magnitude units, than those from NTS for explosions of the same yield as shown in (*Figures*5--2and5--3). Next, using a most probable value of 0.32 for the m_b bias between the sites, the yields of the largest Kazakh explosions were estimated using the appropriate sets of (corrected) coupling curves, where it is required that the tectonic-release-corrected M_s (or M_0), as well as the m_b observations, give the *same* estimate of yield, in order to achieve consistent m_b and M_s (or M_0) yield estimates. This procedure results in estimates of Soviet test yields, with the largest values all very near 150 kt. In particular, by this analysis two events since 1978 have values somewhat larger than the 150 kt value (i.e., near 170 and 180 kt), with the remaining large events having most probable yield values at, or slightly below, 150 kt as shown in *Figure 5-4*.

Charles Archambeau

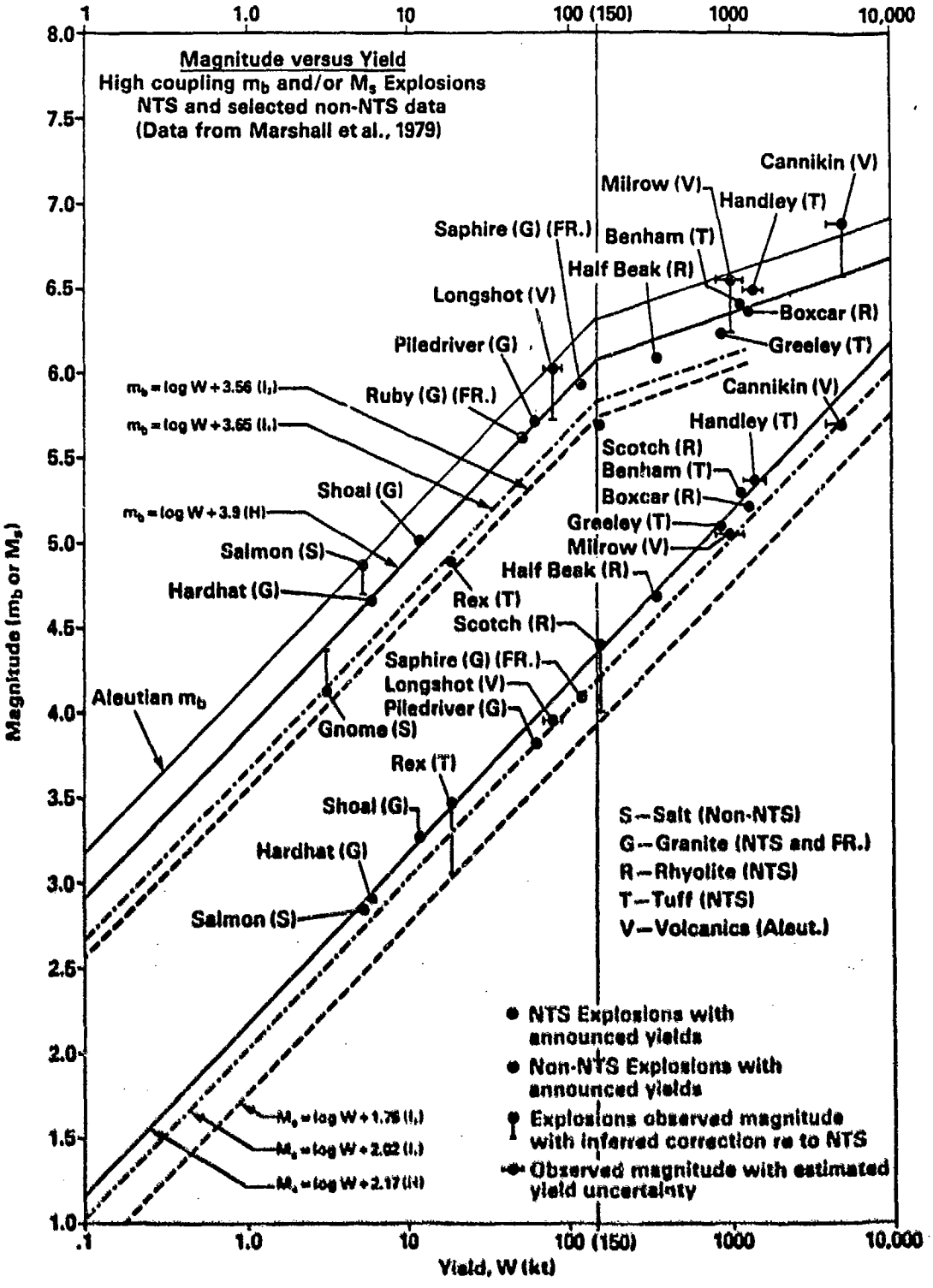


Figure 5-1. Magnitude vs. Yield. High coupling M_b and/or M_s Explosions. NTS and selected non-NTS data.

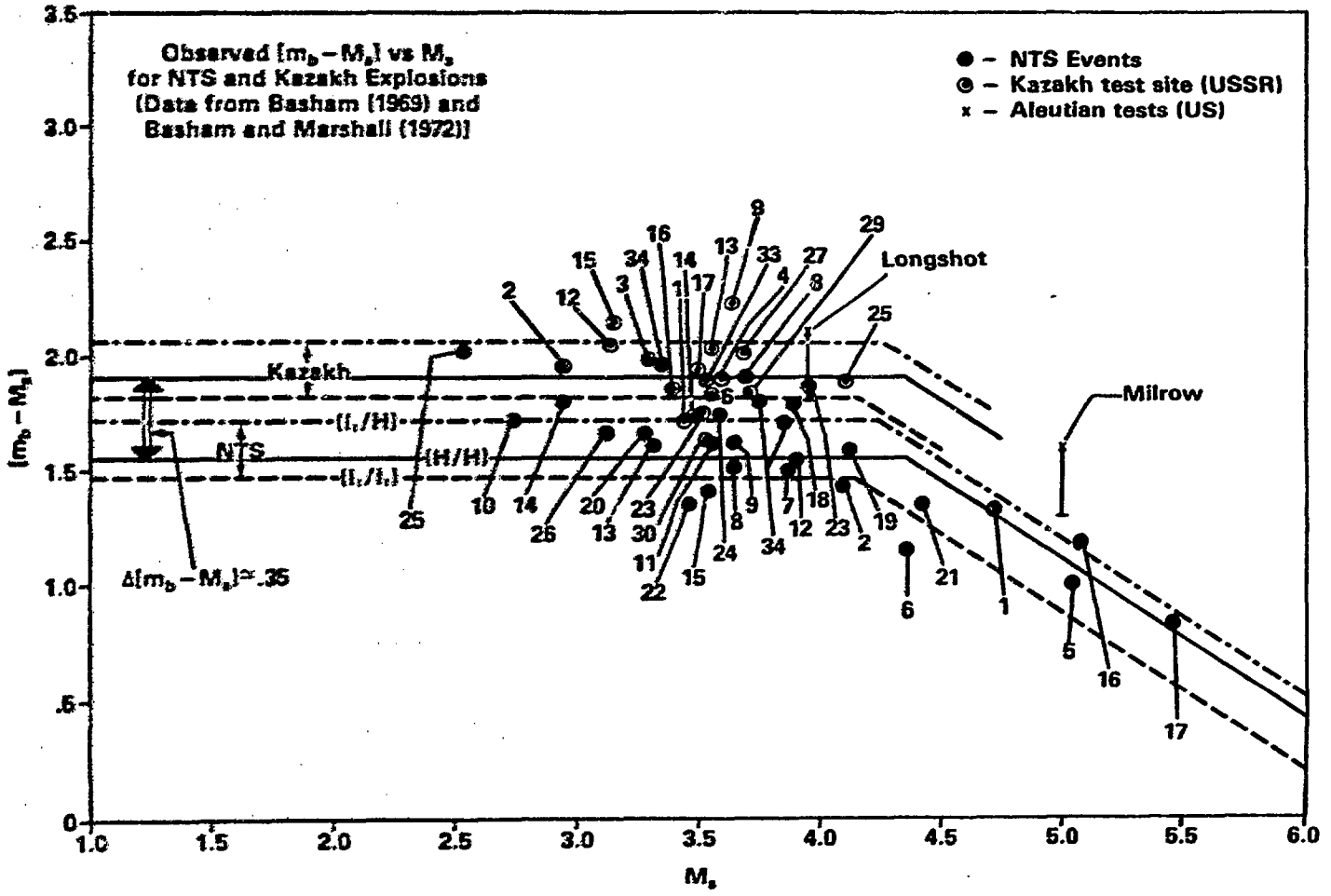


Figure 5-2. Observed $[M_b - M_s]$ vs. M_s for NTS and Kazakh Explosions.

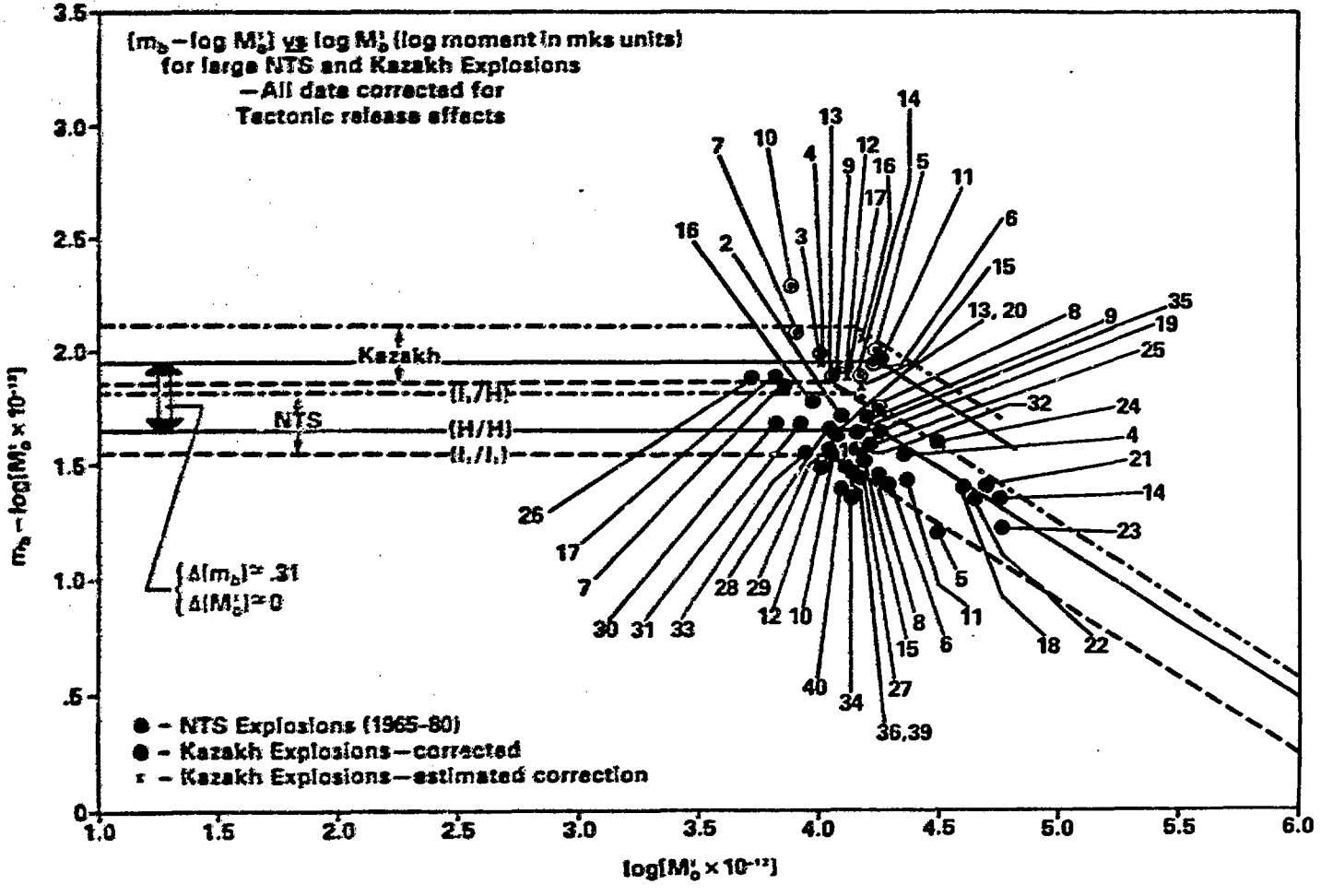


Figure 5-9. $[M_b - \log M_0^1]$ vs. log M_0^1 (log moment in mks units) for large NTS and Kazakh Explosions. All data corrected for Tectonic release effects.

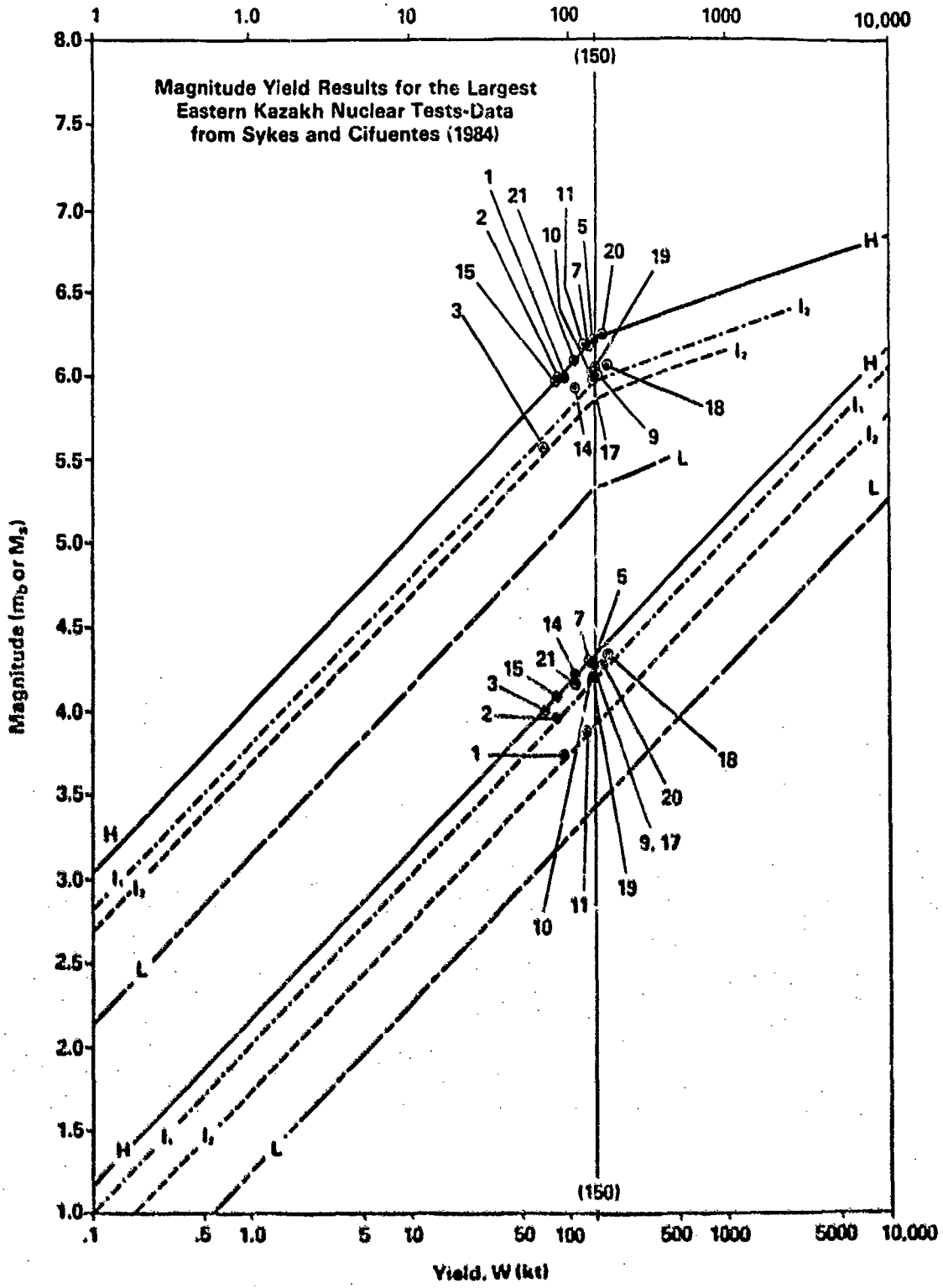


Figure 5-4. Magnitude Yield Results for the Largest Eastern Kazakh Nuclear Tests.

5.3. NUMERICAL TESTS OF SURFACE WAVE PATH CORRECTIONS

In determining path corrections, an approximation is made which separates the observed Rayleigh wave spectrum into source region and path dependent functions (Stevens, 1985). The formulation and suggested approximations were given by McGarr and Alsop (1967). If one assumes no mode conversion or reflection and normal incidence to the boundary, the approximation is equivalent to assuming that the total horizontal energy flux remains constant during the transmission of each mode across the boundary. This approximation will be referred to as the Conservation of Lateral Energy Flux (CLEF).

CLEF was originally proposed to estimate the transmission coefficient of teleseismic Rayleigh waves crossing continental margins at near normal incidence. Bache *et al.* (1978) applied this approximation to near source boundaries on the assumption that the center of radius of curvature of the boundary between source and receiver roughly corresponded to the source location. This geometry reduces the effect on geometric spreading when crossing the boundary. The Gaussian Beam approximation for Surface Waves (GBSW) also conserves the lateral energy flux. In addition, it includes the effect of lateral heterogeneity on geometrical or optical spreading. Both approximations assume that the elastic transition zone is gradual compared to the wavelengths involved so that mode conversion and reflection is negligible.

Since neither approximation was intended for the use made by Bache *et al.*, it is important that their validity be tested for limited source regions with possibly sharp boundaries such as low velocity basins. With this in mind, Rayleigh wave seismograms were calculated for explosive sources at depth in a finite vertical cylinder with contrasting elastic properties embedded in a vertically stratified propagation media. The technique couples lateraly inhomogeneous finite-element calculations of the source region with Green's functions for teleseismic Rayleigh waves using the elastodynamic Representation Theorem (RT).

The source is modeled in the finite-element code by adding a stress-glut of constant moment to the dilatational stress at the centroids of two adjacent elements along the z-axis of the mesh. Stresses and displacements associated with the total wave field propagating out of the source region are monitored at a series of surrounding nodes located in the exterior medium of the finite-element geometry. The RT method is then used to calculate the teleseismic Rayleigh waves by convolving these stress and displacement time histories with the appropriate Green's functions for exterior propagating medium. The Green's functions are fundamental mode Rayleigh waves due to point forces at depth and are calculated by a modal program for the vertically inhomogeneous propagation medium. The elastic properties of the exterior region of the finite-element geometry are the same as those on the surface on which the representation theorem is applied in the teleseismic propagation medium. Because the RT coupling is performed after the wave field has propagated out of the source material, multiply reflected and converted phases are taken into account. These Rayleigh waves for various source material contrasts are compared with the CLEF and GBSW approximations.

The source region medium is a cylinder with a radius of 1.8 km and extends from the free surface down to a depth of 1.8 km. The source cylinder is embedded in a half space of

the same elastic properties as that of the upper part of the propagating medium. The RT surface, over which the propagation Green's functions are integrated, surrounded the plug at a distance of 0.3 km into the half space of propagation material. The RT surface is composed of 11 nodes or rings per side with a spacing of 0.2 km. The propagation model is CIT109, which is based on western United States observations and has a 14 km thick upper crust. The three elastic materials used in the source medium cylinder are characteristic of the shot point properties at three NTS areas; Climax Stock, Pahute Mesa, and Yucca Flat. The seismic velocities, densities, and other material properties for the source media are given in the following table.

In order to demonstrate that modeling an explosion by applying a stress glut at two elements along the the vertical axis was sufficient, an RT calculation was made in which the source medium was the same as the surrounding medium. Then its analytic Rayleigh wave analog was calculated, a point explosion at a depth of 0.4 km in the CIT109 model. The time domain records of both calculations as seen through a long period WWSSN instrument at 3000 km were essentially indistinguishable. The maximum error was less than 2%.

Previous reports have focused on time domain comparisons between RT and CLEF. Since the phase of the CLEF calculation is not used to estimate the explosion and tectonic release moments, this report is restricted to comparisons of amplitude spectra and spectral ratios. *Figure 5-5*, shows the spectra and spectral ratios of the analytic and RT synthetics for the pure CIT109 model without the WWSSN-LP instrument. The frequency range is 0 to 2.5 hz. The spectra on the right side of the figure is decimated 4 to 1 for clarity of display. The agreement between the two calculations is excellent out to the roll-off frequency of the cosine filter which was applied to the finite element forcing functions before their decimation to 0.2 second time steps. The original finite element time step was considerably smaller, by nearly a factor of 13.

Figure 5-6, shows in greater detail the longer period portions of the spectra of these two runs with and without the instrument. At the top, the instrument for these

Table 5-1

Material Properties of Source Region Models

SOURCE MODEL (S)	α km/s	β km/s	ρ gm/cm^3	$\frac{\mu_{CIT}}{\mu_S}$	$\frac{(\lambda+2\mu)_{CIT}}{(\lambda+2\mu)_S}$	α^2/β^2
YUCCA	2.35	1.30	1.86	10.7	10.2	3.27
PAHUTE	4.00	1.90	2.30	4.06	2.86	4.43
CLIMAX	5.33	2.78	2.67	1.63	1.39	3.68
CIT109	6.20	3.51	2.74	1.00	1.00	3.12

frequency ranges is on the right. The latter frequency range gives the best detail of the frequencies of interest in estimating seismic moments using surface wave path corrections. At periods greater than 50 seconds, the RT calculation has larger amplitudes than the analytic or exact synthetic. This can be considered the long period upper limit to the accuracy of the RT calculations.

Figure 5-7, displays the spectra and spectral ratios of RT calculations for the three source media and their CLEF approximations. The agreement between RT and CLEF spectra is good, especially considering the rational behind the approximations. The three cases all show the same character in that CLEF spectra is smaller than RT at long periods and greater at higher frequencies. The crossover is somewhere between 20 and 10 seconds period. Because of the small vertical extent of the inhomogeneous source region relative to the wavelengths of the surface waves, the difference between the spectra of CLEF and GBSW is negligible for periods greater than two to four seconds. Although not shown here, at frequencies higher than 0.25 hz, the GBSW approximation is better.

Peter Glover
David G. Harkrider

5.4. REFERENCES

Bache, T. C., W. L. Rodi, and D. G. Harkrider,"Crustal Structures Inferred from Rayleigh Wave Signatures of NTS Explosions." *Bull. Seism. Soc. Am.*, 68, pp. 1399-1413, 1978.

McGarr, A. and L. E. Alsop, "Transmission and Reflection of Rayleigh Waves at Vertical Boundaries." *J. Geophys. Res.*, 72, pp. 2169-2180, 1967.

Stevens, J. L., "Estimation of Scalar Moments from Explosion-Generated Surface Waves," S-CUBED Scientific Report No. 2, submitted to DARPA, AFGL-TR-85-0097 (SSS-R-85-7232), April 1985.

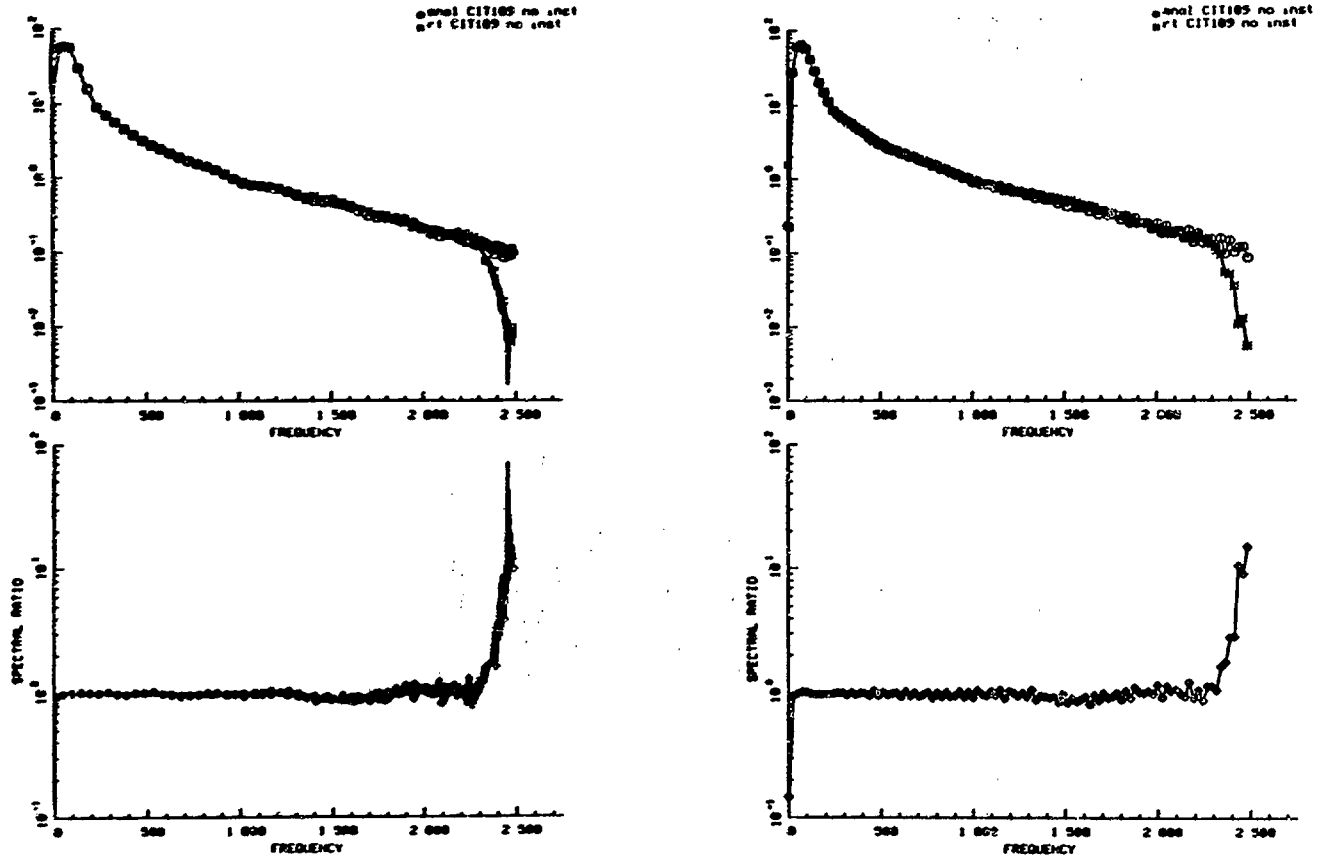


Figure 5-5. **Spectra and spectral analysis of the analytic and RT synthetics for the pure CIT109 model without the WWSSN-LP instrument.**

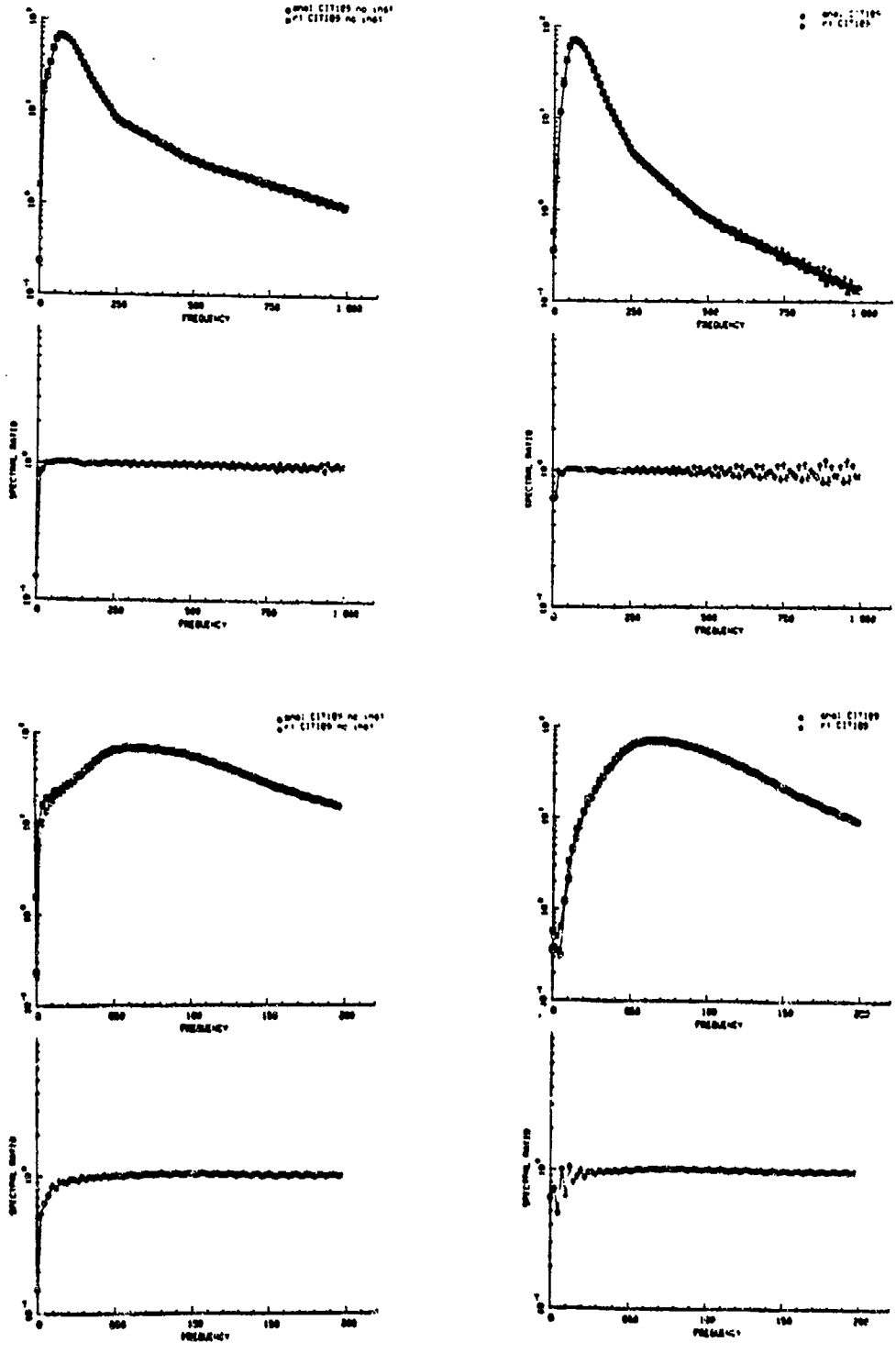


Figure 5-6. Longer period portions of spectra with and without the instrument.

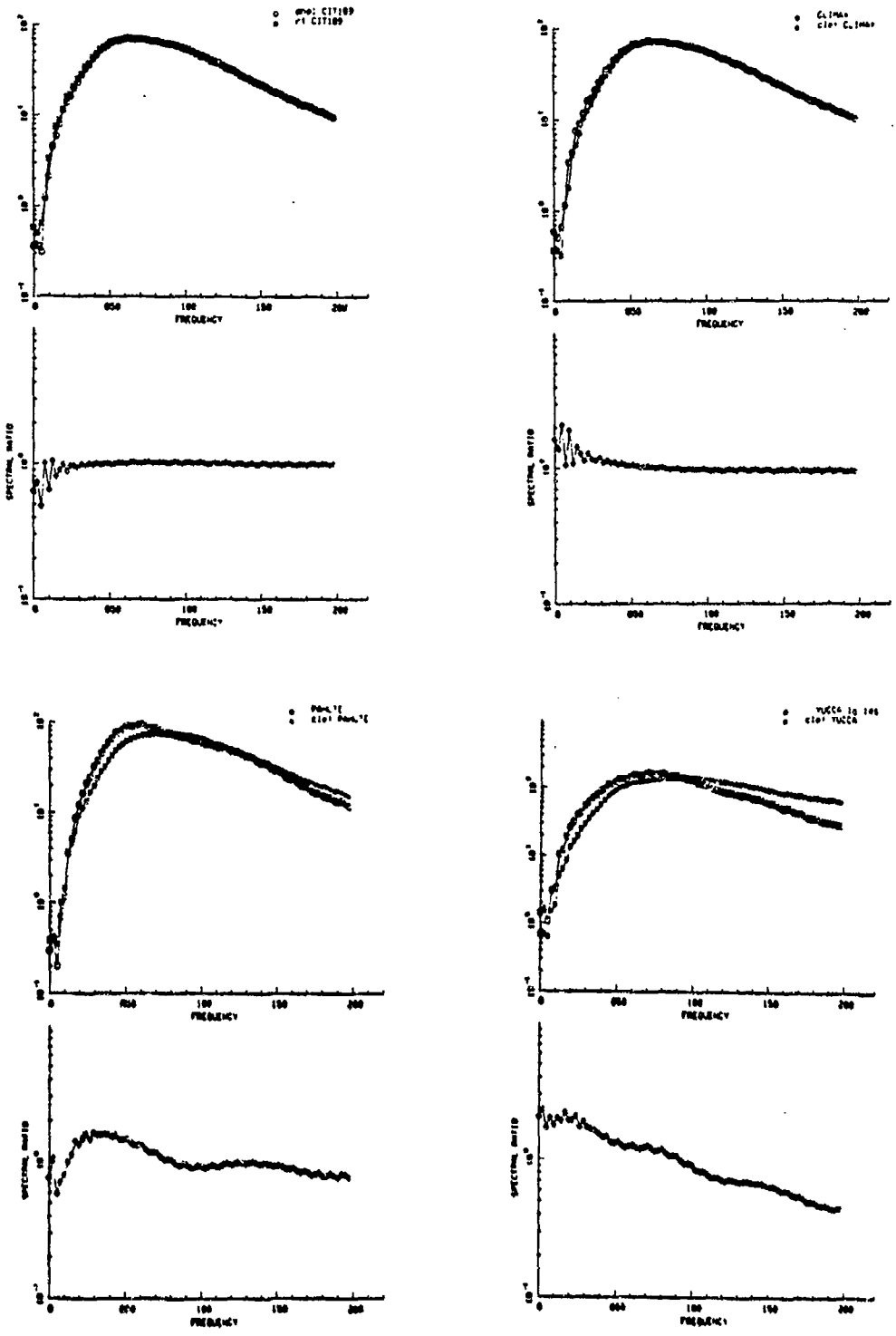


Figure 5-7. Spectra and spectral ratios of RT calculations for the three source media and their CLEF approximations.

6. SYSTEM AND DATABASE DEVELOPMENT

6.1. ACQUISITION OF NORESS DATA AT THE CENTER FOR SEISMIC STUDIES"

6.1.1. Introduction

The Norwegian Regional Seismic Array System (NORESS) is a sophisticated, unmanned array of seismometers that continuously transmits digital seismic waveforms to locations where they are recorded for subsequent analysis. NORESS is designed to detect high-frequency seismic waves that are characteristic of underground explosions and small earthquakes recorded within approximately 2000 kilometers of the source regions. The system produces high quality digital recordings for verification research. The NORESS database is important in research aimed at improving methods to detect and discriminate earthquakes and underground nuclear explosions at regional distance ranges, and determine source characteristics of the detected events.

The array has been in operation since October 1984. It was officially dedicated on 3 June 1985. NORESS consists of 25 seismometer sites arranged in a tapered array of concentric circles which are logarithmically spaced and within an area three kilometers in diameter. Data from the array are transmitted continuously at 32Kbits per second to four receiving sites around the world, one of which is the Center for Seismic Studies. This report contains a description of the array and the hardware and software procedures employed at the Center for the collection of the NORESS data.

6.1.2. Description of the Physical System

NORESS is located 100 kilometers north-northeast of Oslo, about 20km east-southeast of Hamar, Norway. NORESS is co-located with a much larger seismic array, called NORSAR (Norwegian Seismic Array), which was developed earlier through a cooperative project between Norway and the United States Defense Advanced Research Projects Agency.

There are 25 seismometer sites comprising NORESS, of which four are three-component short-period instruments (TSP) and 21 are vertical-component instruments (SP). In addition, located at the center site is a three-component, broadband (BB) instrument generating mid-period and long-period data. All of the TSP and SP elements are sampled at 40 samples/second, the mid-period element is sampled at ten samples/second, and the long-period at one sample/second.

Each of the 24 outlying seismometers (those other than the hub) is installed in a vault, a partially buried concrete and fiberglass housing which provides physical protection. Vaults are rigidly anchored to bedrock. This ensures high quality reception of seismic signals of interest and minimizes interference caused by pressure fluctuations in the air or local disturbances of the soil. In addition to the seismometer(s), the vault contains electronics to amplify seismic signals, circuits to convert signals from analog to digital form, authentication circuits, seismometer calibration circuits, and optical

transmitters and receivers.

The hub is the operational center of the array as well as its physical center. It has two 60m deep holes for its seismometers. The downhole instrument package contains a complete set of electronics in addition to the seismometers. Two cables connect the hub to each outlying seismometer. One, a conventional wire cable, transmits electric power. The other, an optical fiber cable, transmits data. Use of the optical data transmission eliminates electrical interference from lightning, power lines, or similar sources. The hub also services all the other seismometers by providing electric power, timing signals, calibration commands and by receiving, assembling and transmitting all data. A telephone line carries the data to the Norwegian analysis center in Kjeller; a dish antenna transmits data via satellite to three data centers in the United States.

The geometrical configuration of the NORESS array is shown in *Figure 6-1.* The 25 seismometer sites are arranged in concentric circles around the hub. The circles are referred to by the mnemonic names of A, B, C, and D with radii of 150m, 300m, 700m, and 1500m, respectively. The seismometer sites are numbered sequentially, starting at one in the north and increasing clockwise. These numbers are the lower identifying numbers on *Figure 6-1.* The upper identifying numbers are the channel identifying numbers (CID) represented in hexadecimal notation where the individual binary bits of the CID have the following meaning:

SP: 0 through 1F hex
TSP: 20 through 3F hex
BB: C0 through DF hex

Binary bits 0 through 5 are called the channel ID field and bits 6 and 7 are the frame ID field (bit 0 is the least significant). The frame type is determined according to Table 6-1.

Table 6-2 summarizes the configuration of the whole NORESS array as of 30 September 1985. Because of the research nature of the array, its configuration changes from time to time. Table 6-2 should not be used for subsequent time periods. The first column in the table is the component sequence number followed by the CID and the site mnemonic. Next is the instrument sequence number and the component identifier. The subsequent numbers describe the distance from the array center in the north-south and east-west direction and the elevation above sea level. All distance measurements are give in meters. The last column shows the instrument type.

Table 6-1

CID bit 7	6	Frame Type
0	0	SP Z-axis ID
0	1	SP N-axis ID
1	0	SP E-axis ID
1	1	Broad Band ID

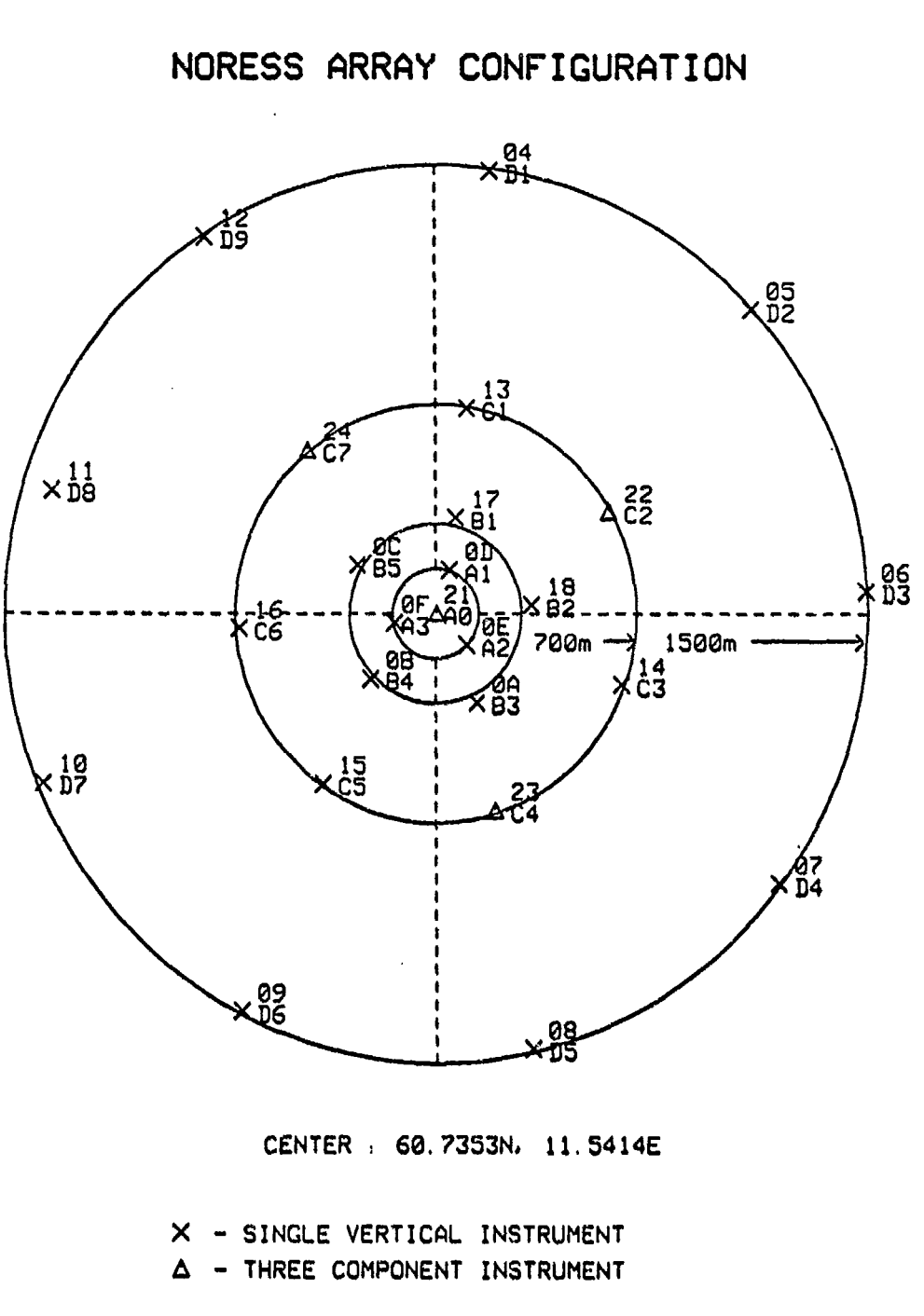


Figure 6-1. NORESS Array Configuration.

Table 6-2 NORESS Seismic Array - Lat 60.7353, Lon 11.5414								
Seq.	CID	Site	Seq.	Comp.	NS(M)	EW(M)	ELEV(M)	INSTRUMENT
1.	21	A0	1.	SPZ	3	4	302	GS-13
2.	61	A0		SPN	3	4	302	GS-13
3.	A1	A0		SPE	3	4	302	GS-13
4.	0D	A1	2.	SPZ	146	49	291	GS-13
5.	0E	A2	3.	SPZ	-103	108	311	GS-13
6.	0F	A3	4.	SPZ	-30	-143	296	GS-13
7.	17	B1	5.	SPZ	321	70	299	GS-13
8.	18	B2	6.	SPZ	30	334	315	GS-13
9.	0A	B3	7.	SPZ	-298	143	314	GS-13
10.	0B	B4	8.	SPZ	-217	-228	299	GS-13
11.	0C	B5	9.	SPZ	163	-272	289	GS-13
12.	13	C1	10.	SPZ	687	109	299	GS-13
13.	22	C2	11.	SPZ	341	603	339	GS-13
14.	62	C2		SPN	341	603	339	GS-13
15.	A2	C2		SPE	341	603	339	GS-13
16.	14	C3	12.	SPZ	-238	647	352	GS-13
17.	23	C4	13.	SPZ	-657	208	311	GS-13
18.	63	C4		SPN	-657	208	311	GS-13
19.	A3	C4		SPE	-657	208	311	GS-13
20.	15	C5	14.	SPZ	-569	-396	299	GS-13
21.	16	C6	15.	SPZ	-48	-687	303	GS-13
22.	24	C7	16.	SPZ	548	-447	275	GS-13
23.	64	C7		SPN	548	-447	275	GS-13
24.	A4	C7		SPE	548	-447	275	GS-13
25.	04	D1	17.	SPZ	1480	192	305	GS-13
26.	05	D2	18.	SPZ	1015	1098	372	GS-13
27.	06	D3	19.	SPZ	76	1493	453	GS-13
28.	07	D4	20.	SPZ	-901	1189	379	GS-13
29.	08	D5	21.	SPZ	-1451	335	348	GS-13
30.	09	D6	22.	SPZ	-1326	-681	352	GS-13
31.	10	D7	23.	SPZ	-566	-1368	337	GS-13
32.	11	D8	24.	SPZ	414	-1336	301	GS-13
33.	12	D9	25.	SPZ	1257	-802	278	GS-13
34.	C1	E0	26.	IPZ	-2	3	247	KS-36000-04
35.	C1	E0		IPN	-2	3	247	KS-36000-04
36.	C1	E0		IPE	-2	3	247	KS-36000-04
37.	C1	E0	27.	LPZ	-2	3	247	KS-36000-04
38.	C1	E0		LPN	-2	3	247	KS-36000-04
39.	C1	E0		LPE	-2	3	247	KS-36000-04

6.1.3. Hardware Connection at the Center

Communication of the NORESS data to the Center relies on a radio frequency satellite transmission link. The first leg of that link uses direct transmission from the NORESS hub station to an Intelsat Atlantic satellite, which retransmits the data to the Intelsat receiving station at Etam, West Virginia. The signal then is relayed via a

domestic satellite, Comsat, to the Center. The receiving antenna is located on the top of the building in which the Center resides. From there, the signal is sent via cable to the 14th floor computer room.

Figure 6-2 illustrates the physical connection from the Comsat signal conditioning equipment to the DEC PDP 11/44 computer, which is dedicated to the NORESS data acquisition. The data is routed through the Loral Data Systems Data Processor, Analyzer, and Display System Mark II (d/pad mk II). The d/pad is a microprocessor-controlled decommutator that performs all of the functions of the central front end of a ground station. The d/pad processes the serially-encoded, noise-contaminated data stream and outputs reconstructed parallel data in real-time. The effective data rate at this point is 32 Kbits/second. The data is captured by a host adapter subsystem consisting of two printed circuit boards: one AXIS, Inc. DR11 FIFO (first-in-first-out) Buffer Module and one MDB Systems, Inc. MDB-DR11-W Direct Memory Access (DMA) Module, as shown on *Figure 6-2*. The d/pad is programmable to the extent that a data frame sync pattern can be entered and it will search the incoming bit stream for that sync pattern and present complete frames to the subsequent electronics. The FIFO Buffer Module provides synchronizing logic for aligning incoming data blocks with d/pad frame and master frame boundaries. It provides buffering for a maximum of 65536 16-bit words of d/pad data organized in FIFO memory. The MDB-DR11-W DMA Module is a 16-bit parallel interface compatible with the DEC PDP/11-44 Unibus. It is used to provide a high-speed DMA data path for d/pad data from the FIFO Buffer to the computer's processor and memory.

6.1.4. Data Format

The basic unit of data transmission is a 50-word frame (16 bits per word); the first word of each frame is a constant, fixed bit pattern: F325 hex. A collection of 40 frames constitutes a master frame. The master frame contains real-time data as well as delayed data. The delayed data time difference is 261 seconds (at this writing). The master frame has the following configuration:

Real time hub status frame
20 real-time data frames
Delayed hub status frame
18 delayed data frames.

The three frames from any TSP element are always transmitted consecutively in either the real or delayed section. Frames from the center TSP and the BB instruments are always transmitted both in the real-time as well as delayed. No frame is sent twice in either section and idle frames replace any element frame which is not in sync so that bit synchronisation can be maintained at the master frame level.

The CID order of the components received in the real time transmission section of the master frame are as follows: 4, 6, 8, 10, 12, 22, 62, A2, 14, 16, 24, 64, A4, 18, B, D, 21, 61, A1, C1. The CID order of the components in the delayed section of the master frame are as follows: 5, 7, 9, 11, 13, 23, 63, A3, 15, 17, A, C, E, F, 21, 61, A1, C1.

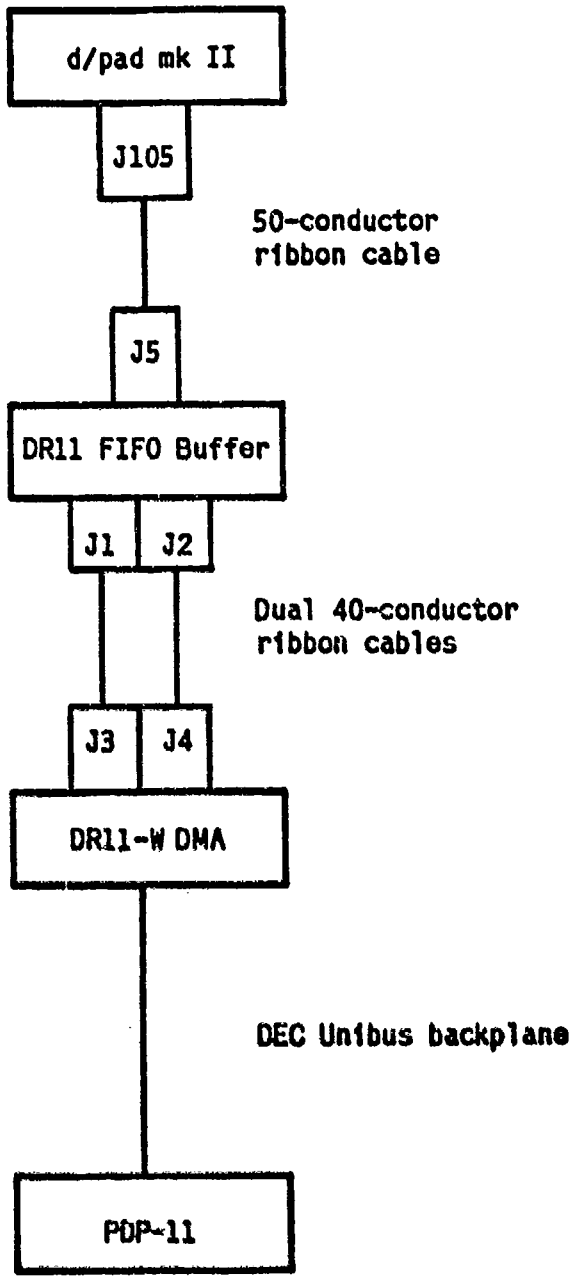


Figure 6-2. Host Adapter Block Diagram.

A more complete description of the data is available from: "Regional Seismic Array System Hub Processor", 20 June 1984, published by Sandia National Laboratories. What follows here are the highlights from that description.

Table 6-3 shows the hub status frame format.

The master frame number is increased by one for every subsequent frame and it is a 14 bit counter, thus it recycles about once every 4.55 hours. The 5 words of GMT contain the time in the following order: milliseconds, seconds, minutes, hours, and days. Each digit of every time element is coded into four binary bits, thus a hexadecimal representation of the time becomes human readable but, of course, the computer must shuffle the bits around in order for the time words to be computer readable.

The contents of status bytes and words are described in great detail in the above mentioned reference; they are concerned mostly with the engineering state-of-health (SOH) information of the array.

Table 6-3 Hub Status Frame Format
Frame Sync Word 1 = F325 hex
Frame Sync Word 2 = 0CDA hex
Master Frame Number
GMT Words 1, 2, 3, 4, 5
Hub Analog Status Words 1 through 10
Port 0, 1 Status Byte
Port 2, 3 Status Byte
Port 4, 5 Status Byte
Port 6, 7 Status Byte
Port 8, 9 Status Byte
Port 10, 11 Status Byte
Port 12, 13 Status Byte
Port 14, 15 Status Byte
Port 16, 17 Status Byte
Port 18, 19 Status Byte
Port 20, 21 Status Byte
Port 22, 23 Status Byte
Port 24, 25 Status Byte
General Status Word
Uplink Time Position Word
Last Hub Command Word
Last Command Byte, Last Command Element
13 Words of Multiplexed Digital Status
1 Word, Content Undefined

The other frames consists of either an SP-type or BB-type data with the exception that if an instrument is not available, then an idle frame appears in its place.

Table 6-4 gives the format of the SP frame.

Time words one and two represent different significant parts of the same 28 bit counter; one is the reverse representation of the other. The AID is the 4 bit array ID which should always be = 1. The CID is in the upper order byte of the next to the last word in the frame. Each seismic data word is a 16 bit gain ranged digitized representation of the sensor analog signal. The most significant two bits of the seismic data word are encoded by the element processor to indicate the gain select value, which is a multiplier of the low order 14 bit value, as shown by Table 6-5.

The format of the BB frame is given in Table 6-6.

The GTSOH and DH SOH status words represent coded indicators of the monitoring of voltage, temperature, and pressure of the seismometers. All the other items in the frame are similar to the ones for the SP frame.

If any component is not in sync, then an idle frame is transmitted instead of the standard frame. The format of the idle frame is given in Table 6-7.

Note that bit 14 of the master frame number, the good/bad flag, is set to zero for the idle frame.

Table 6-4 SP Frame Format
Frame Sync Word 1 = F325 hex
Master Frame Number
Time Word 1
Time Word 1, Channel Status
40 Words of Short Period Data
Channel SOH
Command
Command, AID, Time Word 2
Time Word 2
CID, FA hex
6 Bits Undefined, 10 Bit Authentication Number

Table 6-5		
Bit		Gain Select
15	14	
0	0	128
0	1	32
1	0	8
1	1	1

Table 6-6 BB Frame Format
Frame Sync Word 1 = F325 hex
Master Frame Number
Time Word 1
Time Word 1, BB Channel Status
10 Words of IPZ Data in Order Collected
10 Words of IPN Data in Order Collected
10 Words of IPE Data in Order Collected
LPZ Data Word
LPN Data Word
LPE Data Word
GTSOH Status
DH SOH Status Word 1
DH SOH Status Word 2
DH SOH Status Word 3
4 Unused Words, All = 0
Command
Command, AID, Time Word 2
Time Word 2
CID, FA hex
6 Bits Undefined, 10 Bit Authentication Number

Table 6-7 Idle Frame Format
Frame Sync Word 1 = F325 hex
Master Frame Number (bit 14 = 0)
48 Words, All = 5555 hex

6.1.5. Acquisition Software

The remaining parts of this article describe software development plans as prepared in July of 1985. While the objectives have changed only in small ways since that time, many details have changed in major ways as of the current date (April 1986).

The data collection is a two-stage process which leads to the ultimate archiving at the Center. The first stage is the collection of the data in real-time through the electronics described in the previous sections. The software, described below, gathers the statistics of the data collection, such as master frame number continuity, time word synchronization, etc., and strips off all of the unnecessary items from the data stream and then writes the data onto tape. Some days after the data collection, the detection list for the corresponding time period arrives from the NORESS site via electronic mail. This list is matched with the raw data tape and the data segmentation program extracts the time

period indicated by the detection list. The data archiving software is then invoked for the segmented waveform data.

Additionally, a quick-look capability exists which, when invoked from a Tektronix 4014 graphics terminal, displays the last two minutes of data collected on all available SPZ components. This display takes a few minutes to complete but every time it is invoked, it gives the latest available data.

6.1.6. Data Collection

As data enters the computer, it is captured by the software driver which uses a standard DEC DR11-W protocol. The data enters the computer memory and is made available to the data collection software. Statistics are gathered related to the data quality. The most elementary check is already made by the d/pad, which synchronizes the bit stream according to the master frame sync. Thus, the data would not even get to the computer memory unless the master frame is in sync. There are many other checks which are possible and the following list shows the ones that are done by the data collection software. The basic unit of time for which the statistics are entered into the log is the rollover of the master frame number, which takes about 4.55 hours. Discontinuity of the master frame number will also cause a separate entry.

(1) The continuity of the master frame number is checked. The master frame number should increase by precisely one between subsequent master frames. The real-time and delayed master frame numbers differ by 261 in the same master frame.

(2) Frame syncs exist for every frame within the master frame.

(3) Master frame number bits 14 and 15 are consistent. The master frame number appears in the status frame where these top two bits should be zero. The same number also appears as the second number in every other frame but with the top two bits set to one. Bit 14 is used as the good/bad flag and bit 15 is zero in the hub status frame but a one in every other frame.

(4) No port status byte in the hub status frame should have an unknown element type. This condition is indicated by the three least significant bits of the port status byte being all ones.

(5) The uplink time position word in the hub status frame has a particular format which can be checked. This word contains a bit and byte time related to the uplink transfer of the real hub status frame. The bit and byte time are located in the seven most and the nine least significant bits of the word, respectively. The bit time should never exceed seven and the byte time starts with 12C hex at the top of the command frame and counts down to zero.

(6) Every component occurs at most once in the master frame. If it does not, then it must be replaced by the idle frame. Also, all three TSP components must occur consecutively in either the real-time or the delayed sections.

(7) The hub digital status word in the hub status frame should have the value of 1FF hex. The individual bits represent bit sync lock, carrier lock, etc. indicators and each bit should have a one as its value.

(8) The value of the AID in all of the SP and BB frames should be one.

(9) The value of the Channel Status represents the clip flag, calibration flag, and motor status. This value should be zero.

(10) The four unused words of the BB frame, words 42 through 45, should be all zero.

(11) The CID order should be constant except for the idle frames and the low order byte of a valid CID = FA hex. Although the port connection order at NORESS is indeed different than originally planned, there is no reason that it should change often. Any change of CID order should be reported in the statistics because it will trigger actions in the subsequent software.

(12) GMT continuity must be checked. The GMT should only increase by one second for every master frame. A substantial deviation is a sign of error. This one second increase is not always precise in practice. It has been observed that the GMT gains a millisecond about every 70 seconds of real-time data. Certain time corrections must also be folded into the data presentation but this should be the function of the data segmentation software.

(13) The time words one and two can be checked for consistency with each other. They both appear in the same positions in the TSP and BB frames and they are each split into two different words in the frame. They represent different significant parts of the same 28 bit frame counter. Time word one contains the 2**0 through 2**23 bits in reverse order. Time word 2 contains frame count bits 2**0 through 2**19 and 2**24 through 2**27. The overlapping bits can be checked against each other and a disagreement logged as an error.

(14) Time words one and two should be checked for continuity between adjacent master frames. They combine to make up the 28 bit frame counter which is incremented by one for every master frame. The continuity from frame to frame must be checked and reported in the statistics if there is a deviation. The frame counter is independent across CID's, thus they have to be checked for every CID.

(15) Every TSP and BB frame contain an authentication number which can be used as a measure of goodness. The authentication plans of the NORESS data are currently in progress. When complete, they will become part of the statistical measure.

After the data are examined and the statistics logged, the data of seismic importance are saved in chronological order. At the start-up, the real-time CID frames will be replaced by idle frames in order to maintain the time continuity. The center TSP and BB instruments are transmitted both in which frame to keep. If both are the same, it does not matter which is kept. If they are different, then the one which passes the authentication will be saved. If they both fail the authentication, then the frame which has the least number of problems according to the above described statistical measures will be saved.

The data is buffered on disk, the amount of data buffered being a function of resource availability. If there is enough room on disk, then perhaps a whole day's worth of data can be buffered before putting it on tape. Preliminary estimates indicate that after the unnecessary part of the data is discarded, the data saved will amount to less than 250 MBytes per day. With standard compression techniques and judicious use of magnetic tape, approximately two tapes will hold a day of data. Automatic scripts can also be developed so that if the data overflows the designated disk area, then some of it can be transferred to another computer's disk area by use of the local area network. This would be especially useful on weekends in order to keep the operator intervention of the

data collection process to a minimum.

The result of this operation is the continuous, time-ordered, compressed data which will be the input to the time segmentation process described in the next section. The only processing done thus far is just the gathering of the data transmission statistics. The data on tape will not be in the Center official format. Even the data values are left in the raw, gain ranged, 16-bit format in order to minimize the length of magnetic tape.

6.1.7. Data Segmentation and Archiving

The data are also collected and analyzed by seismologists in Norway. This involves passing the data through an event detector, and by beam forming and filtering techniques, a detection and event list is produced. This list is electronically transferred to the Center using the ARPANET computer network. The whole process takes a few days (3-4 at this writing). By way of example, *Figure 6-9* contains part of the list for day 305, 1 November 1985.

The NORESS bulletin contains 80 bytes of text for each detected phase and the layout is as follows:

```
byte 01-01   detection type:                                  field 01
                2 - P phase (unassociated P has LAT/LON specified
                      associated P does not include LAT/LON
                      on this line)
                3 - E unidentified secondary phase (not used in
                      localization)
                4 - LG phase (LAT/LON not specified)
                5 - solution with origin time, region and LAT/LON.
                    The solution is based on preceding P and LG
                    phases (line 2 and 4)

byte 02-80 for type 2,3 and 4 ----------------------------------------

byte 02-02   ***** not used *****
     03-08   arrival date  mmddyy                             field 02
     09-09   ***** not used *****
     10-19   arrival time  hh.mm.ss.d                         field 03
     20-20   ***** not used *****
     21-22   NORESS identification flag (see below)           field 04
     23-24   ***** not used *****
     25-26   phase  P,LG or E                                 field 05
     27-27   ***** not used *****
     28-34   amplitude, nm zero to peak  9999.99              field 06
     35-35   ***** not used *****
     36-39   period (seconds)  9.99                          field 07
     40-40   ***** not used *****
```

41–44	velocity (km/sec.) 99.9	field 08
45–45	***** not used *****	
46–50	azimuth (deg.) 999.9	field 09
51–51	***** not used *****	
52–56	SNR (signal to noise ratio) STA/LTA 999.9	field 10
57–57	***** not used *****	
58–62	LAT (Latitude) 99.9N	field 11
63–63	***** not used *****	
64–69	LON (Longitude) 999.9E	field 12
70–70	***** not used *****	
71–73	Magnitude (not implemented yet)	field 13
74–74	***** not used *****	
75–80	detection number from RONAPP (RONAPP – automatic processing package)	field 14

byte 02–80 for type 5 ———————————————————

byte 02–02	***** not used *****	
03–08	origin date mmddyy	field 02
09–09	***** not used *****	
10–19	origin time hh.mm.ss.d	field 03
20–22	***** not used *****	
23–23	NORESS identification flag (see below)	field 04
24–29	***** not used *****	
30–55	region name (not implemented yet)	field 05
56–57	***** not used *****	
58–62	LAT (Latitude) 99.9N	field 06
63–63	***** not used *****	
64–69	LON (Longitude) 999.9E	field 07
70–70	***** not used *****	
71–73	Magnitude (not implemented yet)	field 08
74-80	***** not used *****	

Starting on day 293, 20 October 1985, the bulletins based on the real-time processing of the NORESS data are reviewed by an analyst at NORESS. The review process is essentially a check of the detection information contained in the bulletins against the plots produced by the event processing, with complete sets of seismograms, beams, and FK-spectra. Each detection of types 2 and 4, or regional event solution (type 5) is characterized by the analyst. The notation entered into the bulletin as the identification flag has the following meaning.

Characterization of detection output (column 21):

(1) Acceptable results, as far as can be judged from the output,

(2) Probably regional event, with an undetected phase seen in the plots,

	DATE	TIME				AMP	PER	VEL	AZI	SNR	LAT	LON	ML	DPX
2	11 185	2.41.27.3	A2		P	206.76	0.27	12.1	346.0	7.9	82.6N	42.3W		39817
2	11 185	2.45.10.5	A1		P	879.99	0.51	15.8	71.6	27.2	37.0N	98.6E		39818
2	11 185	3. 7. 8.1	A1		P	203.16	0.58	11.2	206.6	6.3	39.3N	2.2W		39820
2	11 185	3.26.35.8	A3		P	122.82	0.37	22.4	26.6	5.4	31.0N	159.6E		39821
2	11 185	3.47.42.5	A3		LG	338.16	0.37	3.3	323.1	9.8		39822		
2	11 185	4.43.27.0	A3		P	104.66	0.40	16.7	0.0	4.9	57.0N	169.2W		39825
2	11 185	5. 6.56.1	A1		P	606.08	0.37	25.0	0.0	18.2	21.5N	169.2W		39826
2	11 185	5. 7. 2.9	A1		P	772.57	0.45	35.4	45.0	5.1	12.9N	144.9E		39827
3	11 185	5. 7. 7.7			E	1537.47	0.57	35.4	45.0	2.8		39828		
2	11 185	5.10. 9.0	A1		P	400.69	0.61	22.4	26.6	7.9	31.0N	159.6E		39829
2	11 185	6.49.14.4	A2		P	215.03	0.26	15.8	108.4	4.7	19.9N	69.0E		39831
2	11 185	8.22.57.0	A1		P	3106.16	0.59	22.4	26.6	53.2	31.0N	159.6E		39833
3	11 185	8.23. 8.1			E	1117.69	0.34	22.4	26.6	6.3		39834		
2	11 185	9. 3.33.4	U3		P	168.66	0.26	10.0	126.9	4.6	46.1N	34.0E		39836
2	11 185	10.18.26.1	H4		LG	261.64	0.25	3.3	258.7	6.5		39839		
2	11 185	10.56.15.9	A3		LG	186.89	0.24	3.3	53.1	5.9		39843		
2	11 185	11. 0.25.1	A1		P	253.09	0.38	6.1	104.0	4.3		39844		
3	11 185	11. 1. 8.4			E	269.19	0.36	4.2	94.8	3.4		39845		
4	11 185	11. 1.15.5	A2		LG	471.13	0.36	4.0	76.0	2.4		39846		
5	11 185	10.59.32.8		A	61.4N	18.0E	1.8							
2	11 185	11.13.36.3	A1		P	844.50	0.31	6.6	203.2	17.2		39847		
4	11 185	11.14. 3.6	A1		LG	586.02	0.31	3.4	195.9	8.5		39848		
5	11 185	11.13. 0.8		A	58.9N	10.5E	1.6							
2	11 185	11.55.25.1	A1		P	1167.68	0.83	35.4	45.0	8.6	12.9N	144.9E		39850
2	11 185	11.56.35.7	U3		P	228.35	0.23	8.6	121.0	5.1	50.8N	32.0E		39852
2	11 185	12.17.11.5	A1		P	217.50	0.24	6.2	209.7	30.0		39855		
3	11 185	12.17.25.7			E	509.61	0.28	3.8	202.6	7.8		39856		
4	11 185	12.17.30.1	A3		LG	712.36	0.27	2.4	230.9	2.7		39857		
5	11 185	12.16.46.4		A	59.9N	9.5E	1.6							

Figure 6-3. Day 305, 1 November 1985.

(3) Probably unreliable phase velocity determination, but clearly a signal,
(4) Difficult to judge the detection because of low SNR of the beam,
(5) Local high-frequency noise,
(6) Spike detection,
(7) Communication line problems.

Quality of the FK-spectrum (column 22):

(1) Sharp main peak, possible secondary peaks 7 DB below main peak,
(2) Less sharp main peak, or secondary peaks up to 5 DB below main peak,
(3) Very broad main peak, or secondary peaks up to 3 DB below main peak,
(4) Random FK-spectrum.

Characterization of phase association performance (column 23):

(1) Acceptable phase association, one regional event,
(2) High probability of undetected PN phase, otherwise the solution is correct,
(3) SN has been picked instead of LG (which is also detected), otherwise correct,
(4) One regional event, but probably error in phase association,
(5) Interfering phases, more than one event, at least one regional,
(6) Solution rejected, probably no regional event.

The detection list is the basis of the data segmentation. The amount of data saved will be based on the type of event: local, regional, and/or teleseismic as it can be determined. Segmentation rules can be arbitrary at the outset and then refined by analyst interaction and verification of the waveform data. The segmentation rules will be played against the detection list and time segments to be saved are identified. Then any overlaps in this list can be eliminated so that a continuous set of segments, ordered in time, can be presented to subsequent software. A day's worth of time segments are matched to the previously collected continuous data and waveform files are generated in official Center format corresponding to those start and stop times.

6.1.8. Quick Look

When this program is invoked, it causes the last two minutes of data to be extracted from the current buffer which is part of the data collection. These data, in Center standard format, are then passed on to the display software. The program must be invoked from a Tektronix 4014 terminal. The program selects the components to be displayed according to a preselected list. The list currently contains all of the 25 SPZ components, provided they are available from the real-time data collection. Components which are in the delayed section will be displayed according to the most recent delayed time.

The display generated by this program takes a few minutes to complete. It is strictly limited by the speed of the asynchronous serial computer connection, the upper limit of which is 9600 baud. It is possible to increase the speed by the acquisition of a parallel line connection hardware; however, the intent of this program is to give a cursory look at the latest incoming data rather than a strict quality control measure.

Louis Huszar

6.2. PDP-11 SYSTEM SOFTWARE

During the last six months, the Center for Seismic Studies continued to develop the PDP-11/44 system to archive NORESS data. Five major developments included (1) rewriting the NORESS satellite interface program, (2) writing a DR-11W device driver, (3) porting the VAX C library, (4) porting the Ritchie C compiler, and (5) rewriting the TM-11 tape driver. These are described below.

The NORESS data acquisition program was rewritten, following DARPA's request, to modify the interim format for waveform segments.

A PDP version of the VAX DR-11W device driver was implemented to interface to the d/pad software written by SAIC. The driver was integrated into the PDP kernel and has no difficulty keeping up with the 4Kb/sec NORESS data stream.

As the C library has changed greatly between Version 7 and 4.2 Berkeley Software Distribution (BSD), the Center ported the VAX C library to simplify moving various seismological programs between the VAX and PDP machines. The new library also permitted the removal of several system calls and routines from the kernel, since these routines were implemented in the libraries. A side effect of the new library was the significant speed-up of many programs, since the C library was completely reimplemented as part of 4.2 BSD for the VAX. Several of the new VAX system calls were added as well, bringing the PDP much closer to 4.3 standards. At the same time, both the PDP kernel and C library were changed to handle up to 16 8K text overlays, allowing the PDP to run programs needing up to 192K of text space.

One of the major problems with the PDP system has always been the compiler. Both Versions 2.8 and 2.9 had the weak Version 7 compiler, with many known bugs, some of which could crash the system. The Center ported the Ritchie C compiler to eliminate these problems, since it has expanded capabilities and an enhanced optimizer. Features of the compiler also include unsigned bit fields and the ability to handle global structures with identically named elements, as well as much tighter general type checking. The combination of the new C library and the new compiler has been especially powerful. Recompiling the kernel and various application programs revealed several problems in the kernel, the networking, and the NORESS programs.

NORESS needed to have multiple tape drives on a single controller as well as a data rate of 6250 bpi, neither of which the current tape driver could handle. A new driver with these capabilities was installed and appears to be reliable. The only remaining difficulty lies in using large (greater than 1K) block sizes. This is thought to be a bandwidth problem for the UNIBUS itself. The Center is currently porting much of the 4.3 kernel in an

effort to make the PDP's I/O and device drivers compatible with the VAX, and plans to install the VAX TM-11 driver sometime this summer.

Keith Bostic

6.3. OTHER SOFTWARE DEVELOPMENT

A number of additional software efforts have been completed to allow Center users to use other waveform analysis packages such as SAC developed by LLNL, to access and analyze NORESS high frequency array data, and to upgrade the association process for seismic events. These are described below:

6.3.1. WSAC, a Program to Convert a Center Standard Database to LLNL SAC Format

WSAC converts a Center standard waveform database (version 2.7) to a suite of files readable by the LLNL waveform analysis program SAC. A Center waveform database contains one wfdisc file; that is, a file with the name db.wfdisc where db is the database prefix. This file contains information about available waveforms, one record per waveform. Along with the wfdisc file, a Center waveform database contains one or more waveform files with names such as db.1.w, db.2.w, and db.3.w. SAC expects one waveform per file (not required in Center standard waveform files), with the waveform index information contained at the beginning of each file. WSAC reads the wfdisc file and writes out one file for each record in the wfdisc file (one for each waveform). The files created have names such as db01, db02, and db03. The first part of the SAC file name is simply the database prefix.

This prefix is appended with a two digit number from 01 to 99. SAC has a limit of 99 files, so only the first 99 files of any database will be converted. A cross reference listing is printed on the standard output. It gives old and new file names of the waveforms along with station, channel, and time information. Table 6-8 is a cross reference between the Center wfdisc file attributes and the SAC header variables. SAC headers are more robust than Center wfdisc records and as such many defaults (indicated with asterisks) have to be assumed.

Table 6-8		
LLNL	Center or default	explanation
delta	1.0/smprat	samp. interval is inverse of sample rate
* scale	1.0	scale for independent (Y) variable
* b(egin)	time-time	start time relative to start time, 0.0
e(nd)	(nsamp-1) * smprat	end time relative to begin time, length
* internal1	2.0	LLNL software sets it, so CSS does also
nzyear	year	year
nzjday	day	day of year
nzhour	hour	hour
nzmin	min	minutes
nzsec	int(sec)	integral seconds
nzmsec	(sec-int(sec))*1000.	milliseconds
* internal4	6	LLNL software sets it, so CSS does also
* internal5	0	LLNL software sets it, so CSS does also
* internal6	0	LLNL software sets it, so CSS does also
npts	nsamp	number of data points
* iftype	1	1 means time series
* idep	6	6 means displacement trace
* istype	9	9 means time given above is begin time
* lpspol	TRUE	components positive by l.h.r.
* lovrok	TRUE	ok to overwrite file
kstnm	sta	station name
kcmpnm	chan	component name
kdatrd	adate	date data available
kinst	instyp	instrument type

6.3.2. PRSNRS, a Program to Parse NORESS Bulletins

NORESS, the state-of-the-art high frequency seismic array in Norway, is providing the seismic community with large quantities of high quality digital data. In order to manage this data, NORESS also provides a daily log or bulletin of interesting arrivals and events as detected at NORESS. This bulletin is produced by the RONAPP analysis system at NORESS, a sophisticated program that can filter, beamform, detect, and locate events. Reproducing the output of the RONAPP system would be costly. Consequently, to better use the waveforms from NORESS, and to avoid having to duplicate the RONDAPP system processing, the Center undertook the task of reading the NORESS bulletins and converting them to Center standard database files. The NORESS bulletin has matured over the months with the latest versions being much cleaner and easier to read.

Below is an abbreviated excerpt from a recent bulletin, showing two types of the records possible, along with the opening and closing lines "++". As stated above, these bulletins can contain arrivals or events. Event lines have a "5" in the first column. The

first one included here shows that on 18 March 1986 at 14:38:5.4 GMT an event in southern NORWAY was detected at NORESS. This event is based on a P and LG arrival at 14:38:36.4 and 14:39:1, respectively. The first two arrivals in the bulletin are not associated to any event, although the P listed includes a possible latitude and longitude determined from the array parameters.

Abbreviated Excerpt from NORESS Bulletin

++	DATE	TIME	AMP	PER	VEL	AZI	SNR	LAT	LON	ML
2	031886	07.12.05.1	A1_P	1080.5	0.38	35.4	45.0	24.8	12.9N	144.9E	
2	031886	08.21.03.0	C3_LG	492.7	0.14	3.9	218.7	5.1			
2	031886	14.38.36.4	A1_P	535.1	0.26	6.6	203.2	12.9			
4	031886	14.39.01.0	A2_LG	424.1	0.30	3.7	206.6	7.6			
5	031886	14.38.05.4	__A	535	SOUTHERN NORWAY				59.2N	10.0E	1.5
++											

This bulletin, once run through the parser, produces a number of files corresponding to the Center INGRES data base relations.

PRSNRS creates event, origin, arrival, assoc, feature, and counter files and creates wfreq file when invoked with the -w option. Each event in the bulletin generates one record in the event and origin files. Each associated arrival generates a record in the assoc file. Each arrival generates a record in the arrival file and each unassociated P arrival generates a feature record. The following shows the contents of the files produced by the short bulletin.

Parsed Equivalent of the Earlier NORESS Bulletin

Event file:

```
    1     1     1
```

Origin file:

```
 1986077  511540685.400  59.2000  10.0000   0.0000 -1.00 -1.00   1.50 -1
  2  -1  -1  -1  -1 -999.0000      1      1 535  56 _ _          NORESS.1
    NORESS.1_ML    _|ANA=__A|SOUTHERN NORWAY
```

Assoc file:

```
     3      1  -1.000 P        -1.000  -1.000 -999.000 a -1.000 -999.00
_   |DPX=8302|
     4      1  -1.000 LG       -1.000  -1.000 -999.000 a -1.000 -999.00
_   |DPX=8303|
```

Arrival file:

```
 1986077  511513925.100 NRA0  bs + _ P       _     1080.5   0.38 -999.00
-1.0  45.00   3.14  -1.00  -1.00 _ _     1     1    -1     -1 NORESS.1
     |ANA=A1 |SNR=24.8|DPX=8295|

 1986077  511518063.000 NRA0  bs + _ LG      _     492.7   0.14 -999.00
-1.0 218.70  28.51  -1.00  -1.00 _ _     2     2    -1     -1 NORESS.1
     |ANA=C3 |SNR=5.1|DPX=8300|

 1986077  511540716.400 NRA0  bs + _ P       _     535.1   0.26 -999.00
-1.0 203.20  16.85  -1.00  -1.00 _ _     3     3    -1     -1 NORESS.1
     |ANA=A1 |SNR=12.9|DPX=8302|

 1986077  511540741.000 NRA0  bs + _ LG      _     424.1   0.30 -999.00
-1.0 206.60  30.05  -1.00  -1.00 _ _     4     3    -1     -1 NORESS.1
     |ANA=A2 |SNR=7.6|DPX=8303|
```

Feature file:

```
     1 -999.00 -999.00 -999.00 -999.00  -1.00000  -1.00000  -1.00000  -1.0
0000  -1.00000  -1.00000 NORESS.1 ML     _          _      12.9000  144.9000
  0.0000  511513107.552 1986077 210  17
```

Wfreq records, when requested, are generated for each unassociated arrival and for each event. In generating waveform segments for unassociated arrivals, data from 60 seconds before the arrival time to 180 seconds after are saved from the short- and mid-period channels. Large portions of the LP are saved, from 1 hour before the arrival time to 2 hours after the time. For declared events and their associated arrivals, the procedure is more complex. Since P is the earliest arrival, the sp and mp waveform windows start 60 seconds before the P arrival time. The last associated arrival time is used to fix the end of the window. Twice the last arrival time minus the P arrival time is saved along with the 60 seconds before the P arrival time. If this number is smaller than 240, then 240 is used. The LP procedures for events are the same as for unassociated arrivals.

6.3.3. Work on the Content and Format of the NORESS Bulletin

Two major types of data are available from the new NORESS array in Norway. The

first type of data is the waveforms themselves. The second is the bulletin produced by the RONAPP processing system and the analysts reviewing its output. The original format of the NORESS bulletin was mostly complete, but lacked some features that could make it easier to process. The list below shows some records from an early NORESS bulletin. The records that begin with 5 are event records and the records that begins with 2 are arrival records. The early event records show no geographic region numbers or geographic names. Also, on occasion, the ML would overflow its field. The early arrival records also had problems with field widths, with the AMP, SNR, and ML fields all overflowing on occasion.

The AMP, VEL, SNR, and ML fields would also occasionally run into the space that separated them from the previous field. Another problem not apparent from this example is the lack of any clear delimiter indicating when the bulletin starts and when it ends.

Early NORESS Bulletin

	DATE	TIME		AMP	PER	VEL	AZI	SNR	LAT	LON	ML
5	63085	6.10.24.0	NRS						67.5N	33.3E	2.6
2	63085	6.43.31.5	NRS P	220.45	0.50	8.6	121.0	6.1	50.8N	32.0E	
5	61885	22. 6. 4.8	NRS						36.7N	68.9W	***
2	61885	0. 2.38.2	NRS P	135.05	0.36	25.0	0.0	5.3	21.5N	169.2W	***
2	51985	6.18.10.7	NRS P	******	0.79	15.8	16.4	223.2	59.5N	159.3E	0.0
2	52985	20.26.34.4	NRS LG	61.97	0.26	3.6	85.9	******			0.0
2	61985	18.35.19.5	NRS P	******	0.17	9.3	248.2	******			0.0
2	51985	0.49.30.6	NRS P	364.07	0.62	999.9	0.0	5.6	21.5N	169.2W	0.0

In order to correct these problems the changes noted here were made. The next excerpt contains two records from the revised bulletin. Seismic region number and name are now included. Unnecessarily wide field widths are trimmed (such as AMP). The output has ben made more readable by the insertion of leading zeros in the time field when values are less than 10. Added checks have been added into the software to ensure that none of the fields overflow into the space separating the field from the previous field. The seismic data is now delimited by double plus signs before and after, making automatic processing of the bulletin more robust. Lastly, analyst comments have replaced the static NRS field of the old format. These comments indicate if an event is accepted, rejected, or if various other conditions of importance exist.

Revised NORESS Bulletin

++	DATE	TIME		AMP	PER	VEL	AZI	SNR	LAT	LON	ML
5	031886	14.38.05.4	A	535	SOUTHERN NORWAY				59.3N	10.0E	1.5
2	031886	07.12.05.1	A1 P	1069.5	0.38	15.6	45.0	24.8	12.9N	144.9E	
++											

6.3.4. MERGO, a Program to Compare Origin Lists in Space and Time

MERGO is a program that compares event solutions (hereafter called origins) generated by the automatic association program AA that are close in space and time. Currently the AA program tries to identify these "twins" with either user-specified or default values for the allowable time and distance differences. MERGO has two major features in processing an origin list. One feature is that the output of an AA program might have some undetected split events which should actually be merged into one. MERGO would help identify these origins. Secondly, origin lists from different organizations frequently must be compared and merged, and MERGO is well suited for this. The Center database standard allows multiple alternate solutions for an event location. MERGO attempts to group origins it deems the "same" under one event heading.

MERGO has a simple algorithm that looks for origins close in time and then checks to see if the latitude and longitude differ within specified limits. If the origins pass all these tests, they are declared to be the same events. MERGO can create an event file for a given origin file. It can also reassign evids and orids. Both of these options make it a useful tool for database building. MERGO also gives some indication of the depth-difference:time-difference ratio, which can be useful in determining if the two events are likely mislocated versions of the other. If this ratio is between 10 to 20km/sec, then one can assume some depth/time trade off has happened in the location of one of the origins. Included here is an excerpt from the manual page for MERGO.

usage: mergo [-p prefix [-c]] [-t time_tol] [-l lat/lon_tol] [-f] [-o] [-e]
where:
-p prefix specifies the database prefix, if none origin file
is read from the standard in and output is on standard out.
-c specifies reading as well as writing the counter file.
-t time_tol fixes the maximum allowed origin time difference.
-l lat/lon_tol fixes the maximum latitude/longitude difference.
-f indicates that the evid's are to be fixed.
-o indicates that the orid's are to be reassigned.
-e indicates that the evid's are to be reassigned.

Michael Tiberio

6.4. DATABASE DEVELOPMENT

6.4.1. Introduction

The Center for Seismic Studies currently maintains or is preparing many databases. Some of these databases represent a complete set of information of some type and the contents are fixed. Others are undergoing development on a continuous basis as new data of the same type arrive at the Center. An advantage that the Center offers to users is the

ability to retrieve windowed waveform segments containing selected events, selected phases, or selected times upon request. These segments are all in a standard format and contain complete identification and calibration information. Most of the waveform data reside on tape, due to their size, but indices to them are on-line and access to the waveforms is fairly rapid (depending on computer loads).

The database structure at the Center is described in Center Report 85-001 (Center for Seismic Studies, 1985). The database management system can be thought of as a collection of tables and a software system for manipulating them. The data management language is called *QUEL (QUEry Language)*, and it is supported by the *INGRES* relational database management system. Information on this system is contained in the INGRES reference manual and in two general documents available through the Center, *A Tutorial on INGRES* by Robert Epstein, and *Tutorial on the Center for Seismic Studies Databases* by Jon Berger. Additional information can be found in *Accessing Seismic Data at the Center for Seismic Studies* (Lees, 1985) which is available from the Center.

Primary emphasis in the Center program is now shifting toward developing the capability to conduct advanced research. This will be done by using the various databases together with software tools designed to create a friendly interface between the user and the computers used for data retrieval. To accomplish this while providing computing and database management services to other contractors in the DARPA program will require some reordering of priorities, reallocation of available facilities, and development of new software.

6.4.2. Scripts to Access the Databases

As described by Lees (1985), a number of scripts have been written at the Center to help users who are unfamiliar with INGRES find and retrieve data from the databases. For example, a script *helpdb* identifies the types of databases available at the Center (digital waveforms, alphanumeric data, film data). Through a series of queries and user responses, it takes the user to deeper levels of the script, finally accessing particular databases through INGRES to retrieve a list of data available for the event or events of particular interest. *Figure 6-4* is a diagram showing the *helpdb* script and its relationship to integrated and separate databases. Once the user knows which database has pointers to the waveforms or other type of data needed, other scripts are available to create the files to dearchive the data from tape.

6.4.3. Status of Database Development as of 31 December 1985

Within the database management program highest priority during the second half of 1985 was directed toward continuing the collection and archiving of waveform data being received on a daily basis at the Center. High priority during 1986 will be given to collection of continuous data from the NORESS array and new high frequency data from NORSAR, creation of an integrated explosion database, collecting and archiving instrument responses for all waveform data available at the Center, and responding to user requests for data. Lower priority will be given to expanding or creating other databases, and in some cases the Center staff will simply be coordinating the creation of special databases

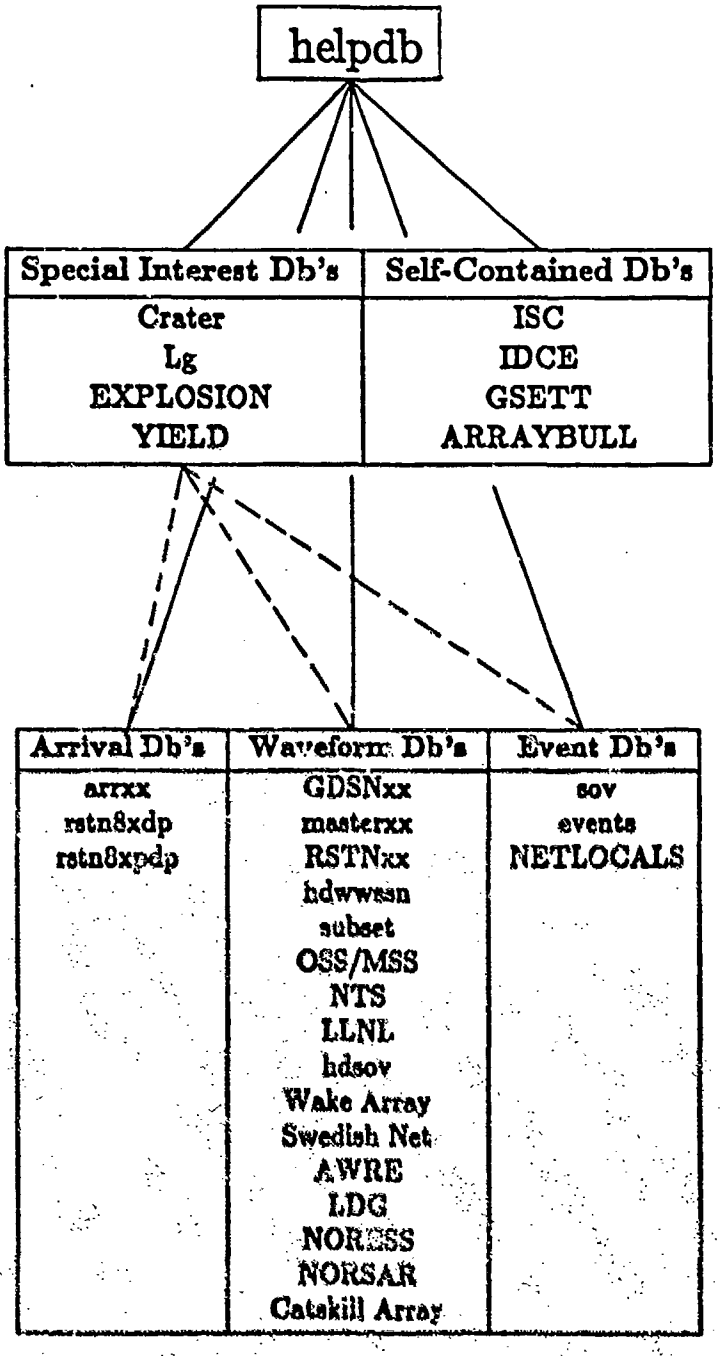


Figure 6-4. Diagram showing relationship of *helpdb* script to integrated and separate databases described in this plan. The *helpdb* script enables the user to access waveforms of other desired information in the databases without prior knowledge of the INGRES database management system.

by contractors outside the Center. The status of official Center databases as of 31 December 1985 was:

GDSN: Complete for the period 1 January 1976 to 12 October 1985.

RSTN: Currently in two archives -- one compiled from SCARS tapes received from Sandia National Laboratories and the other from on-line data received via satellite transmission at the Center. The first of these databases contained 192 days of data for 1982, 364 days for 1983, and 109 days for 1984. The second database contained data for 208 days in 1984 and was complete for 1985.

NORSAR: Contained waveform segments for 11 events.

NORESS: Contained NORESS data for four days, plus waveforms for 143 events recorded by the prototype array during 1983.

hdwwssn: Had hand-digitized WWSSN recordings for 108 events for the period 1963 - 1981.

WAKE: Had Wake Island hydrophone database waveform segments for 2,530 events for the period 8 September 1982 - 15 August 1984.

subset: Contained waveform segments from the Geneva arrays (WMO, BMO, TFO, UBO) for 29 events in 1966.

Catalogs: The "arrival" database contained readings and origin information received from the NEIS, Canadian Seismic Network, United Kingdom, Yellowknife array, and information received over the WMO telecommunication network. It was complete for the period 16 May 1982 - 29 December 1985 at the end of this reporting period. Detector and post-detector lists for the on-line RSTN data were complete for the periods given above for the waveforms. Origins reported in the worldwide epicenter ("wwe") by AFTAC were complete for 15 October 1984 - 15 December 1984. The "events" database (origins from ISC and NEIS) was complete through 1 December 1985. Several regional catalogs have been or are being added to the databases.

Richard Baumstark

6.5. REFERENCES

Center for Seismic Studies (1984). *Center for Seismic Studies Data Base Structure Relations, Attributes and Tape Formats, Version 2.7*, Center Report 85-001, 49 pp.

Lees, A.C. (1985). *Accessing Seismic Data at the Center for Seismic Studies*, Center Report 85-002, 67 pp.