(9)

DESCRIPTION OF AN AUTOMATIC SEISMIC MEASUREMENT PACKAGE: ANALYSIS AND DESIGN GOALS

TECHNICAL REPORT #3

ENSCO, INC.

SAS-TR-81-03
1 September 1981

DESCRIPTION OF AN AUTOMATIC SEISMIC MEASUREMENT PACKAGE: ANALYSIS AND DESIGN GOALS TECHNICAL REPORT #3

Prepared by
Robert Sax

Prepared for

AIR FORCE TECHNICAL APPLICATIONS CENTER
VELA SEISMOLOGICAL CENTER
ALEXANDRIA, VIRGINIA

ENSCO, INC.
Signal Analysis Systems Division
5400 Port Royal Road
Springfield, Virginia 22151-2388

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A132 067	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Description of an Automatic Seismic Measurement Package: Analysis and Design Goals		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER SAS-TR-81-03
7. AUTHOR(s) Robert Sax		8. CONTRACT OR GRANT NUMBER(s) F08606-80-C-0021
9. PERFORMING ORGANIZATION NAME AND ADDRESS ENSCO, Inc. SAS Division Springfield, VA		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS V/0704/B/PMP
11. CONTROLLING OFFICE NAME AND ADDRESS Advanced Research Projects Agency Nuclear Monitoring Research Office Arlington, VA 22209		12. REPORT DATE 1 September 1982
		13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) AIR FORCE TECHNICAL APPLICATIONS CENTER VELA SEISMOLOGICAL CENTER ALEXANDRIA, VA 22314		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

ARPA Order No. 2551

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Automatic Seismic Signal Processors
Automatic Analytic Signal Detection
Ground Motion Measurements
Multi-Variate Source Discriminants

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

This is a system concept study of how the VSC Seismic system could operate as a test bed for developing an optimum automated world-wide seismic network. Many problems are anticipated in developing such a network. These are discussed in a theoretical way and algorithms are suggested which could be tested by the VSC system. These would lead to more optimum performance of an automated seismic network.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by AFTAC/VSC, Patrick Air Force Base, FL 32925, under Contract Number F08606-80-C-0021.

AFTAC Project Number:	V/0704/B/PMP
Project Title:	Routine Automatic Seismic Analysis Package
ARPA Order Number:	4199
Name of Contractor:	ENSCO, Inc.
Contract Number:	F08606-80-C-0021
Effective Date of Contract:	10 June 1980
Contract Expiration Date:	15 February 1982
Project Manager:	Gregory B. Young (703) 321-9000

ENSCO, Inc.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

OVERVIEW AND SUMMARY

This is a system concept study of how the VSC Seismic system could operate as a test bed for developing an optimum automated world-wide seismic network. Many problems are anticipated in developing such a network. These are discussed in a theoretical way and algorithms are suggested which could be tested by the VSC system. These would lead to more optimum performance of an automated seismic network.

The VSC Seismic system is expected to realistically emulate the operation of a seismic surveillance network. This can be done by executing a set of candidate algorithms to perform all of the automatic functions which transform continuous seismic sensor data into desired seismic event information. Operation of the VSC Seismic System is expected to provide a test bed for objectively evaluating various competing automatic algorithms which detect, locate, and identify explosions. It is desirable to impelment a baselined VSC Seismic System with sufficient flexibility to test algorithms needed for optimum operation of a seismic surveillance network. For that purpose, users can apply automatic algorithms to process standard seismic data files and gauge objectively the impact of a proposed algorithm on the operation of a seismic network. By testing these algorithms under such realistic operation conditions, these algorithms under such realistic operating conditions, the best performing algorithms can then be selected for implementation in data centers which monitor test ban treaties.

As its goal, any seismic surveillance system transforms continuous seismic sensor data to a set of located and timed events classified as earthquakes or explosions, with sufficient information to support such a conclusion. An automated system for achieving this goal can be represented topologically as a series of function processes. This is a general representation of the system, and it is independent of the geographical distribution of the sensors and location of the data processing elements of the system. Whether or not the data processing is centralized or distributed throughout the network, a sequence of functions must be carried out to reduce raw sensor data to the desired event information.

At the front-end of the system, where continuous sensor data are accessed, an automatic detector generates a set of apparent arrival times of signals. A location/association process transforms the asynchronous stream of signal detections into an event stream. This provides a preliminary description of the location, origin time, and characteristic of each seismic event. These are the presumed events which are generated by operating the system. Given the event locations, an automatic signal editing process transforms stored continuous sensor data into tentative seismic event records. These records contain signals associated with an event or background noise at the expected arrival of the phase of interest. At this stage, records from arrays, three--component sensors, or single-sensors are reduced to single--channel records of seismic phases. These are composed of compressional, shear, or surface wave phases. Next, a detector, optimized to accurately time each seismic phase, activates processes to measure signals or noise, and extracts a short compressed edit of the detected seismic phases. These signal measurement data and compressed seismic phase edits are subsequently input to an interactive partition of the VSC Seismic System. This performs functions to generate event discriminants, refine the location, and classify the event (e.g., as an earthquake, an explosition, or an unknown event type).

The concept of a sequence of linear programmed function processes, each of which independently access and updates data files, is illustrated by Figure S-1. The decision functions which control the flow of information are shown in diamond shape boxes. The functions to be performed are shown in rectangular boxes. Major data files stored by the system are also shown. Major data files stored by the system are also shown. An asterisk is placed alongside of those functional processes covered by our Automatic Seismic Signal Processing Research.

Inspection of Figure S-1 makes clear that certain tradeoffs are involved in designing the VSC Seismic System. Clearly Figure S-1 does not constitute a design of the system, but merely shows the basic functions to be performed by any seismic surveillance system. It also shows the sequential nature of these function processes and their interaction with data file structures generated by the system. For example, the continuous waveform file might be either centralized or distributed in some manner throughout the network. In any case, the signal editor will have a need to access whatever data are available in order to make signal measurements. The design of the VSC Seismic System obviously will place limits on performance:

- Communications affect the amount of data which can be assessed, the time delay, and the amount of storage required for data accessed.

- The complexity and reliability of the automatic algorithms also affect the amount of raw data storage needed and time delays inherent in processing the data.

- The effectiveness of front-end detection and location affects the amount of interactive processing

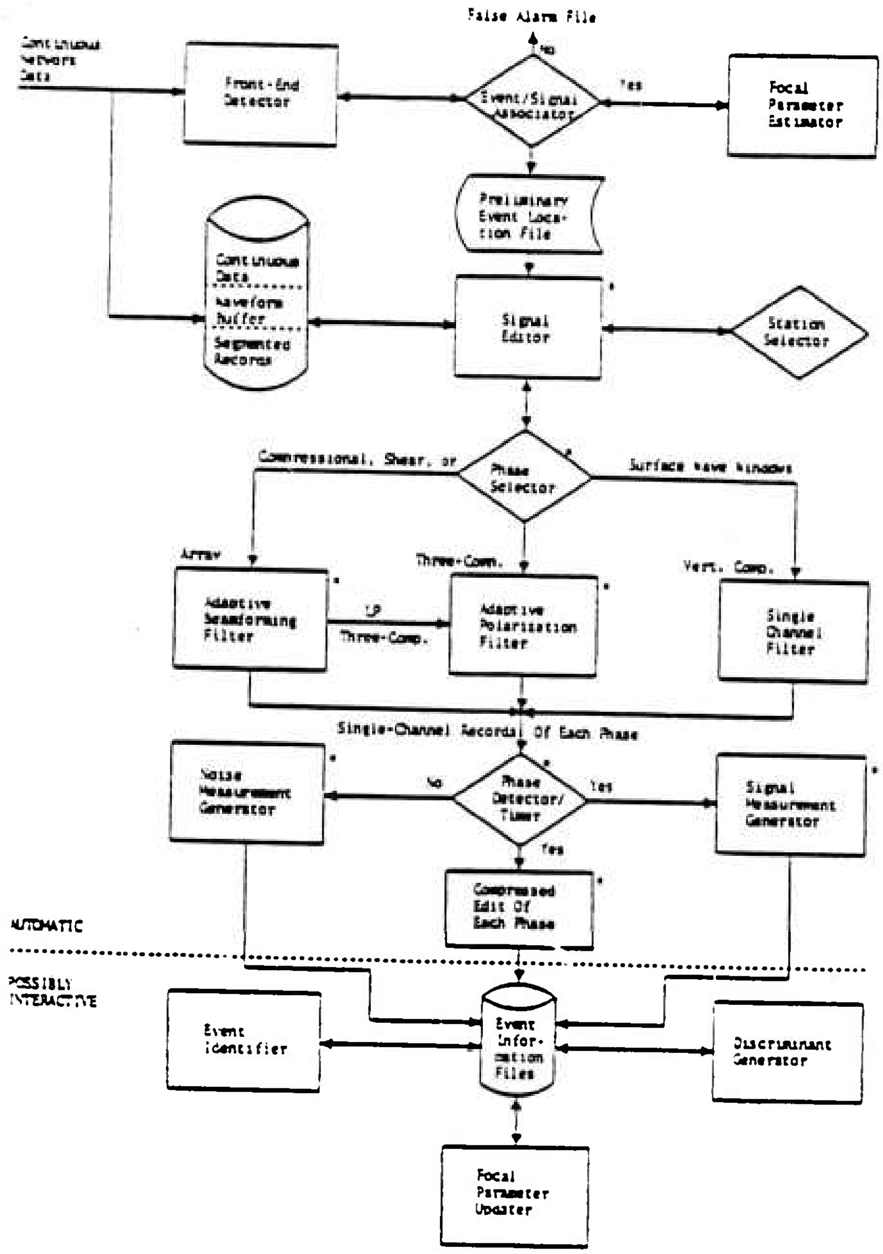


FIGURE S-1
FUNCTIONS PERFORMED BY A VSC SEISMIC SYSTEM

required for quality control. This indirectly affects the amount of raw data which needs to be stored and time delayed.

- The setting of standards on record size, number of phases to be retrieved, and digital sample rates affect the reliability of the signal editing process. It also affects the amount of storage and computer power needed for automatic processing.

- The maximum reduction of data by automatic processing reduces the modest proportions the computer power needed for interactive processing.

- An effective phase detector/timer and compressed edit process greatly reduces the amount of data accessed by interactive processes. This greatly reduces the amount of raw data storage needed and the time delay of the system in reporting results.

Given this system context of functional requirements, we will design an automatic seismic signal processing package which will incorporate some existing, proven algorithms for signal editing, extraction, and measurement and suggest alternative algorithms for future testing. Moreover, we have also redesigned the Unger detector to correct problems which have caused it to miss a large percentage of signals in the past. These algorithms will be designed for modular, independent operation within the VSC Seismic System context. As such, they will be designed for optional execution by users of the VSC Seismic System, or they can be replaced by alternative algorithms for comparative testing. We anticipate that this goal is compatible with the concept of a VSC Seismic System as being a test bed for examining network operation. It is also anticipated that other algorithms put into the baseline of the VSC Seismic System will be similarly structured.

A. DESCRIPTION OF THE BASELINE AUTOMATIC SIGNAL EDITOR (ASE)

The baseline from which the Automatic Signal Editor (ASE) was derived was the system used by ENSCO, Inc., in the Event Identification Experiment. Its purpose and the procedures it utilized to edit seismic signals will be briefly reviewed below. The Automatic Signal Editor, as we define it, is that part of a seismic surveillance system which independently processes and derives measurements from seismic events and updates files of such measurements for use in the interactive (analyst dominated) portion of the seismic processing system.

1. Purpose and Generalized Procedures Associated with the ASE

The purpose of an advanced Automatic Signal Editor is to selectively improve preliminary seismic source estimates, including source location, origin time, depth, magnitude, and discrimination parameters, as well as to reduce the number of unassociated or misassociated seismic phases which are detected. These efforts must be accomplished prior to initiation of the analyst intensive interactive processing phase of analysis to insure a time efficient interactive session. In concept the ASE could be employed as follows.

After an event has been tentatively located and its origin time estimated by an automatic association program, the arrival time of selected seismic phases can be predicted at each available station. A timing tolerance is established for worst case location errors, allowing enough time to observe signals from complex source regions with multiple propagation paths. Sufficient time is also allowed to accurately determine the noise state prior to the arrival of the signal.

Given a set of records containing signals of unknown precise start time and duration, automatic processors are used to (1) select records containing useable seismic data, (2) condition data to minimize the effect of malfunctions, and (3) to utilize an automatic detector algorithm to more precisely time the onset of the seismic signal and determine its duration.

Those portions of the seismic record which contain one or more possible seismic signals from the event are extracted from the record and filed as possible signal waveforms of the known seismic event. These siesmic waveforms can be accessed later by seismic analysts for quality assurance, reprocessing, or post-detection validation of their presumed signal status.

The bottom line of the signal editing process is to measure the character of the seismic signal. The signal measurements are filed for later access by the analyst. They can be used for research and to perform the functions of event discrimination and characterization.

2. Descripton of Seismic Event Editing Procedures

Continuous seismic waveform data are accessed to extract segmented records containing desired signals. These are short-period regional P waves, S waves, short-period surface waves and long-period surface waves.

An automatic detector algorithm is applied to extract more precise time and measure the magnitude of signals from known events. Our experience with Unger's algorithm in the Event Identification Experiment has shown that this procedure currently severely limits the effectiveness of the signal editing process. As a result there is a serious performance gap between the manual retrieval of seismic signals by seismic analysts and that of automatic detecors. Our goal is to overcome this performance gap by redesigning the automatic detector which times and extracts seismic signals from long seismic records.

B. DESCRIPTION OF UNGER'S DETECTOR

1. Theoretical Background

The general model for seismic noise consists of randomly modulated envelope and phase angle functions. The noise statistics underlying the modulation are presumed to be stationary over the duration of the seismic record. Therefore, noise parameters which are derived from noise preceding the signal can be applied over the entire duration of the record.

The general model for seismic signals is a randomly modulated envelope and phase which starts at the K^{th} point and ends at the $(K+D)^{th}$ point. With respect to retrieving seismic signals, the detection problem is to correctly decide where the signal starts, i.e., on the K^{th} point, and how long it lasts, i.e., for a duration of D seconds.

Analytical detectors are based on amplitude and phase angle modulation measurements which are subject to error. The frequency of a signal, measured as the time derivative of the phase angle, is particularly subject to large errors due to the large envelope fluctuations. Considerable effort is needed to obtain more accurate and precise measurements of a signal's frequency so that the frequency, as well as magnitude, can be used for character recognition of weak signals.

2. Unger's Criteria for Timing Signals

Unger greatly simplified the theory of noise by considering noise interference with a coincident signal as a constant level of power subject to random phase modulation. He considered seismic signals as a fixed amplitude level combined with a deterministic phase angle versus time relationship. For example, a turned-on cosine function can be described as a fixed amplitude of one combined with a linear phase versus time relationship over the duration of the signal.

Unger derived a relationship for determining the probability that the envelope of a signal combined with noise interference exceeds the fixed noise envelope level; and that phase angle prediction errors are less than $\pi/2$. He observed that detectors based on this phase angle criteria can certainly detect signals at a level 6 dB lower than by the envelope criteria. However, since he could not derive a deterministic model for the time variation of the phase angle, he could not successfully utilize the phase angle relationship.

Unger utilized the envelope probability relationship to time the onset of seismic signals. He computed the fraction of times a possible signal exceeds the maximum observed noise and applied a threshold to the detection statistic.

3. Design of Unger's Analytic Detector

Unger determined the peak level of noise preceding the portion of the record containing possible seismic signals. He applied an envelope probability threshold of 0.3 to a forward looking time gate of duration 4 seconds. That is, if he found that a maximum fraction of envelope values exceeded the maximum values observed of noise, he would declare a possible signal. Then he searched for the first signal peak at least 2 to 3 dB above the maximum noise to confirm the signal detection and time its onset. Finally he would back up 3/4 of a cycle to precisely time the signal onset.

4. Post-Mortem Evaluation of Unger's Algorithm Applied as an Automatic Signal Editor

In the Event Identification Experiment, Unger's algorithm was used to time the occurrence of signals from known events on long time records. It controlled the decision as to what portion of a short-period record should be extracted, measured, and filed. The detector was part of a larger automated Short-Period Earthquake Editor (SPEED) package which accessed records, corrected for system response and preconditioned the long records of data prior to extracting and measuring signal waveforms. Correspondingly in the Event Identification Experiment, a Long-Period Earthquake Editor (LPEED) performed the same function for long-period surface wave data.

Two aspects of the performance of Unger's detector were excellent. The timing precision of the detected signals was estimated to be approximately $\pm$ 0.1 seconds (about the level which would be expected from seismic analysts). Also, the false alarm fraction of extracted signals was low (about 5%). In this respect it exceeded acceptable performance (a false alarm rate of 10%). Lowering the threshold to detect more signals was attempted but did not significantly change the detection capability of the algorithm.

For acceptable performance, an automatic signal editor should be able to detect most of the signals visible to an analyst. An acceptable level would be a detection probability $P_d \geq 0.9$. Unger's automatic detector applied to signals edited by seismic analysts was not capable of achieving this level of performance at any false alarm rate. Unger's algorithm was unable to detect and time any signals for approximately 40% of the events edited. This indicates a detection probability, $P_d < 0.5$. The detector, in its present form, has a missed signal problem and needs to be revised to attain an acceptable level of signal extraction performance.

Another requirement for acceptable performance is to validate signals extracted by the automatic detector. Network validation could be achieved by extracting signals with extreme magnitude and arrival time errors. In its present form, SPEED does not carry out this network validation function. Consequently, it occassionally detected, timed, and. measured mixed event signals, which although obviously visible to an analyst were not valid because of apparently extreme magnitude and travel time anomalies. We estimated that 5% to 12% of the long records contain mixed event signals. Clearly this type of error must be reduced to negligible proportion since it can result in erroneous event characterization, especially for smaller events.

Other major faults were observed in our use of Unger's detection algorithm in SPEED for the Event Identification Experiment. The algorithm missed about 30% of the signals detected by analysts, which were visible to an analyst but did not exceed the maximum observed noise level. The algorithm missed another 10 to 15% of signals which were obvious to an analyst and clearly well above the maximum observable noise level. Our post-mortem assessment of errors indicated several causes. Unger's method of determining the maximum noise level was not robust. Occasional spikes or glitches in the noise caused his algorithm to set an unreasonably high maximum noise level shich could not be exceeded even by clearly visible signals. Unger's model showed a tendency to miss some clearly visible impulsive signals of short duration. The cause of this effect was the use of a 4 second signal duration for estimating the probability that maximum observed noise level is exceeded. Impulsive signals of less than about 1.3 seconds duration would fail to be sensed by the threshold criteria applied to the estimated probability of exceeding maximum noise. Impulsive signals are not uncommon and it is thus essential that the presumed duration of signals cover a much wider range, e.g., from 0.5 seconds to 10 seconds. Unger's detector missed some complex signals for much the same reason. These extend over a time span greater than 4 seconds and the envelope of these signals are highly variable and skewed toward low level fluctuations. In those cases pulses above the maximum noise level are frequently interspersed with dead spots below that level. This results in a failure of the probability test to stop the detector at the beginning of such complex signals. These are commonly missed or detected with very late start times. Unger's algorithm, as presently implemented in SPEED, warms up on one-minute of noise preceding the window containing signals and continuously updates noise statistics until a signal is detected. This updating procedure reduces the robustness of the algorithm, especially in the case of detecting and timing emergent events. In that case, a signal which is initially less than the maximum noise level rises gradually above that level but not fast enough to pass the probability threshold test. Instead, the maximum observed noise level is gradually raised and effectively shuts off the detector so that it misses emergent signals. Another fault which caused major errors in subsequent event identification was the occasional erroneous extraction of the wrong event. Since event-signals are accessed with four minute time windows and events can be expected at the rate of about one per hour, there is a small but significant chance (about 5 to 10%) that more than one event's signals are accessed by such a long record. Since SPEED has no post-detection validation tests of extracted signals, this effect results in an occasional serious editing problem, especially for small events which are interfered with by larger events. The final fault of SPEED is that it is not fully automatic. Records were visually scanned by us to remove obvious malfunctioning stations. In a fully automated signal editor, this quality control should be built into the signal editor.

As a result of our post-mortem evaluation of SPEED, we are proposing major revisions to obtain a fully automated signal editor which has the potential of performing at a level comparable with that of a seismic analyst. In order to time and measure weak signals at levels below the most probable occurring noise level, we propose applying a multivariate analytic detector capable of sensing the character of such signals, i.e., frequency, bandwidth, duration, etc. as well as sensing positive power fluctuations. The multivariate analytic detector should be designed to detect, time, and determine the duration of multiple arrivals at each station to be sorted out and interpreted by a post-detection network validation strategy. Signal and noise parameters should be measured by robust ordered statistical analysis which is insensitive to occasional spikes and other intermittent malfunctions. No prior assumption should be made of the signal duration. This parameter is highly signal dependent and should be determined effectively and efficiently from measurements of data. There is no need to continuously update noise statistics as is done for front-end detectors; noise parameters should be assessed from noise preceding the signal and assumed to apply over the subsequent duration of the record. Network validation procedures need to be applied to assure that any signal extracted is reasonably consistent with what is expected from the known event. Automated quality control will be performed to remove obvious malfunctioning stations and to condition marginal data subject to intermittent malfunctions. Finally, a reasonably acceptable automatic signal editor should effectively trade-off false alarms for improved detection of weak signals from small events.

C. ANALYSIS OF AN ADVANCED ANALYTIC DETECTOR

1. Generalized Analytic Signal and Noise Model

The ambient noise model will be generalized to include the effects of system malfunctions. Based on Unger's (1978) study of noise envelope fluctuations, normal noise envelope fluctuations are presumed to be consistent with or close to a Gaussian noise model. System malfunctions will appear as severe departures of extreme values from such a model. This generalized noise model provides a basis for building automatic quality control into the automatic signal editor.

The generalized signal model also will have a system malfunction component to avoid contamination of automatic signal measurements by occasional intermittent system malfunctions. The signal model is generalized beyond that of Unger's (which presumes a single signal arrival) to include multiple signal arrivals. These will include multiple signals from complex events (i.e., multiple explosions or earthquakes), secondary propagation phases, regional phases, core phases, mixed event signals and local events, and receiver scattered phases unique to a particular seismic station. The system noise model will include effects caused by spikes and glitches, dead spots and clipping.

2. Multivariate Amplitude and Phase Statistics

We define, on an exacting physical basis, a set of standard multivariate signal amplitude and frequency measurements. These are designed to achieve band-limited whitening of seismic signals of any type which can be reasonably expected. Spectral moment operators are applied for this purpose. Measurement of variates related to the bandwidth of observed signals and measurement of the dominant frequency of observed signals are used for feed-back control of post-bandpass filters to improve extraction of band-limited pre-whitened seismic signals.

3. Basic Ground Motion Measurements

An important factor in implementing Unger's theory of phase detectors is the error and stability problem of computing accurate phase and frequency measurements. The problem is caused by the interdependence between random frequency and amplitude modulations of seismic noise and signals. To solve this problem we will transform the data to an asynchronous time-ordered set of envelope peaks. On that basis, our analysis indicates that we can condition the data to obtain much more accurate and precise measurements of the magnitude of ground motion, frequency, arrival time, and the pulse width associated with each envelope peak. In order to achieve band-limited pre-whitening of any conceivable type of signal, we will apply five spectral moment (difference operators) for that purpose. In this way any plausible seismic source is automatically searched for whitened bands where the source can be accurately measured. For example, if the frequency falls off as f^{-2} above the corner frequency then a second derivative operator with spectrum f^{-2} whitens the source spectrum above the corner frequency.

4. Interpretation of Basic Ground Motion Measurements of Noise; Data Conditioning

From the analysis of Rice (1954), we can detect the presence of strong almost periodic noise interference as significant departures from expected Gaussian noise statistics. A technique was derived for demodulating such noise to separate the desired broadband data from the interfering nearly periodic noise. Where such problems are encountered we will demodulate the desired data from such interfering components.

Based on a relationship observed by Unger relating the frequency of seismic envelope modulations to the standard deviation of phase measurements, we expect to see a large dependence between the magnitude fluctuations of seismic noise or signals and the frequency or spectral bandwidth of such envelope peaks. A theory for correcting noise for observed magnitude and frequency fluctuations was developed. The result of applying such corrections to noise data is expected to significantly reduce the variance of observed magnitude and frequency fluctuations of noise. This, combined with the removal of periodic noise components is expected to significantly enhance the detection of weak signals by independently applying magnitude and frequency measurements.

The application of such corrected independent detection statistics will be applied to the five moment operators designed for band-limited whitening of seismic signals. This will result in the definition of a ten component multivariate measurement of seismic envelope peaks. In general, and especially for weak signals it is not expected that all of the components will be detected. Those detected, however, will be measured. Those not detected will be classified as noise which will be measured as control for subsequent network analysis of signal measurements (i.e., to obtain unbiased estimation of magnitude).

5. Detection of Apparent Signals

A theoretical treatment is given of the problem of reducing basic ground motion measurements to statistically independent magnitude and frequency fluctuations. The input data to such a process are observed magnitude fluctuations corrected to remove any frequency dependence and observed frequency fluctuations corrected to remove any magnitude dependence. A linear rotational transformation operator Q is applied to transform Z statistics of corrected magnitude fluctuations and Z statistics of corrected frequency fluctuations to independent unit variance measures of magnitude fluctuations of noise and frequency fluctuation of noise. Significant deviations of signals from these distributions provide an independent basis for detecting a signal based on its change of frequency or its change of magnitude from values expected from noise.

Pulse width measurements of noise are related to the measured average time interval between noise envelope peaks. Rice (1954) showed that this statistic can be used to estimate the bandwidth of an ideally filtered Gaussian process. By performing this calibration we can use our basic pulse width measurement of envelope peaks to estimate the bandwidth of a Gaussian process whitened by the spectral moment operators. This is extremely valuable information because the observed dominant frequency and the associated bandwidth provide a basis for feedback control of a bandpass filter optimally designed to extract the band-limited whitened Gaussian process. This will be applied to marginal detections of weak signals to confirm their detection status and to extract those signals with minimal seismic noise interference.

6. Detecting Complex Signals

The problem of detecting signals as a single significantly unusual envelope peak is generalized to that of detecting strings of such outlier envelope peaks. The latter envelope peaks would be expected to be replicas of the same signal process which would be expected to yield the same magnitude and frequency fluctuations. Thus by observing repeated anomalous peaks which vary little from each other, we can apply the likelihood detection criteria proposed by Unger (1978) which should enhance the detection of very weak but persistent complex signals.

The detection of signals as strings of envelope peaks provides additional event information. The start time of the signal is taken as the measured start time of the first peak in the cluster of peaks. The end time is taken as the end time of the last envelope peak of the signal process.

Measurements of complex signals will be accomplished by ordered statistical analysis of envelope peak magnitude fluctuations and frequency fluctuations within the time window containing the clustered envelope peaks.

7. Network Strategy for Retriving Signals from Small Events

We expect to be faced with the problem of identifying small events which are masked by seismic noise. By applying fixed constant false alarm rate thresholds to constrain false alarms, we may keep such false alarms to an acceptably low level, but at the same time make it nearly impossible to extract weak signals. In the case of low magnitude signals, we can employ a station variable false alarm rate strategy which raises the expected ratio of retrieved signals to retrieved false alarms to an acceptable level. On the presumption that an identification is required of all events analyzed, it is better to have at least one or a few signals combined with some false alarms than to have only noise measurements to identify the signal. This strategy is made somewhat more sensible by the fact that network validation procedures will be able to eliminate at least some of the false alarms and that network magnitude determinations will compensate to some extent for noisy determinations of the event parameters.

D. SUMMARY AND RECOMMENDATIONS

A concise description will be given for the design of a multivariate analytic detector used to extract short-period seismic signals. From this we will be able to gauge the minimum effort required to implement this srategy for retrieving signal information. From this we expect to develop the concept of a baseline implementation of a multivariate analytic detector which is sufficiently advanced to achieve our goal of editing all possible weak signals from small events. The output of an automatic seismic signal editor is to provide the maximum amount of additional data from known seismic events to improve location of the event and to improve identification of the source.

One of the functions performed by the automatic signal editor is to extract all of the waveforms which can be possibly interpreted as signals from the event. These extractions include long-period surface waves, short-period teleseismic P waves and secondary phases, and regional phases. This information will be automatically inserted into signal measurement files which can be accessed for interactive seismic processing by seismic analysts.

Another function of the automatic signal editor is to reduce basic ground motion measurements to estimates of the event magnitude, the dominant frequency and frequency band, and the complexity of an event and its coda. Some advanced applications of the signal measurements will be to derive source discriminatnts, application of clustering theory to identify anomalous events, and source region calibration to determine precise magnitudes of normal earthquake events.

SECTION I

DESCRIPTION OF THE BASELINE FOR DESIGNING AN AUTOMATIC SIGNAL EDITOR (ASE)

Our background of performing the Automatic Signal Editing (ASE) function was the Event Identification Experiment, Sax, et al. (1979). There we transformed continuous waveform data or long records into compressed extractions of seismic signals and into a set of seismic measurements to be used for event identification. The functions which were performed completely under automatic control were to:

- Access seismic records containing the desired seismic signals from events of known location and magnitude.
- Utilize an automatic detector to detect and accurately time the signals.
- Generate files of compressed seismic signal waveforms.
- Generate files of signal measurements.

The files generated by the ASE were then accessed by an interactive seismic processor which transformed the signal measurements to a set of event measurements characterizing the source of the seismic event. From the event measurements, we generated discriminants and performed adaptive statistical analyses to identify obvious earthquakes and to identify anomalous seismic events such as explosions.

A. PURPOSE AND GENERALIZED PROCEDURES ASSOCIATED WITH THE ASE

Some obvious benefits of speed, efficiency, and objectivity are obtained by automatically reducing the flow of continuous waveform data to the essential data required to perform event identification. Why this rapid reduction of data is a useful goal can best be illustrated by the following scenario of world-wide seismic data collection.

- Events of $m_b > 4$ are expected at the average rate of one per hour.
- A network employed to measure the events includes
  - 10 short-period (SP) vertical component single-sensors
  - 5 nineteen element, SP, vertical component arrays
  - 10 SP, three-component, regional sensors
  - 25 long-period (LP), three-component sensors.
- The input data flow of this network is
  - 10 million words per event of SP data
  - 90 thousand words per event of LP data.

Since most of this data collected by the network is seismic noise, considerable reduction of data is possible by saving only data containing visible signals. To estimate the scale of this reduction consider that an event would typically consist of 12 detected short-period signals per event and 4 detected long-period signals per event. Since signal durations are expected to vary between one and ten seconds, the compressed edit of each signal, allowing sufficient space for coda and noise preceding the signal, can be accomplished with

600 word records of data sampled at 20 words per second. This assumes that an automatic detector can reliably detect and time the seismic signals.

The transformation of continuous waveform data to a set of compressed signal records results in a reduction of data from 10 million words per event to 10 thousand words per event of short-period signal data. A further reduction of data of 10:1 is possible by reducing each detected signal to a set of signal measurements such as the arrival time, measurements of magnitude, complexity, etc. This latter reduction of data to one thousand words per event is a difficult step to take, exclusively, because of the problem of agreeing on data processing standards and the need for waveform data to maintain historical data files. At some future time, at least for events of lower priority, one could anticipate reducing the event data base exclusively to signal measurements.

By reducing continuous waveform data to compressed signal records and signal measurements, it will then become feasible to perform analyst-interactive processing on large data bases. These will either be world-wide or from targeted regions. Our experience with the Event Identification Experiment indicated that the critical element controlling the quality of results obtainable by this automated approach is the performance of the automatic detector used to detect and time seismic signals. Given this goal of developing a capability for fast and reliable event identification, we will concentrate on developing suitable procedures and algorithms for extracting and measuring signals from known events.

The VSC system should provide sufficient noise information from non-detecting stations for networking algorithms such as those for estimating unbiased event magnitudes. Hourly RMS noise measurements are sufficient to provide a reasonable basis for estimating the relative noise at different stations of the network. Such hourly RMS noise figures could be routinely provided as event header information with only 50 words per event. It would be preferable to also provide frequency dependent noise estimates. This could be done on a daily basis. Such network noise information could be most efficiently provided by autocorrelation unction estimates of seismic noise; by 64 lags at 133 words per event.

Providing this important noise information would be less than 200 words per event. It will not add significantly to the amount of storage required in files containing historical seismic signal data.

This reduction of the seismic data flow to sets of detected signals, signal measurements associated with the event, and hourly network noise update information provides all the information needed to rapidly identify earthquakes and to maintain historical seismic event data files required for future event studies. The purpose of this study is to explore the critical procedures and algorithms needed for this automated approach to signal editing and measurement for future application to the VSC system. We will describe some of the existing algorithms used in the Event Identification Experiment with which we found problems. We will describe those problems along with the modifications needed for satisfactory operation of the automated VSC system.

B. DESCRIPTION OF SEISMIC EVENT EDITING PROCEDURES USED FOR EVENT IDENTIFICATION

The automated editor carried out four basic analysis procedures. These are to:

- Access signal time-windows
- Extract the signal waveform
- Measure detected signals
- Measure noise at non-detecting stations.

1. Accessing Signal Time Windows

a. Short-period signals

Short-period signals are accessed from the computed arrival time of events located and timed by the VSC system. The initial location, depth, and timing of the event may lead to substantial errors in timing the arrival of seismic signals. The front-end process for performing this function is shown schematically on Figure I-1. Less than the desired precision in locating the focus of events stems from the following factors.

- Multiple transmission of signals due to complex source, path, or receiver effects
- Large uncertainty of correctly associating the same phase out of sets of multiple arrivals at 4 or more stations
- Occasional mixing of event-signals

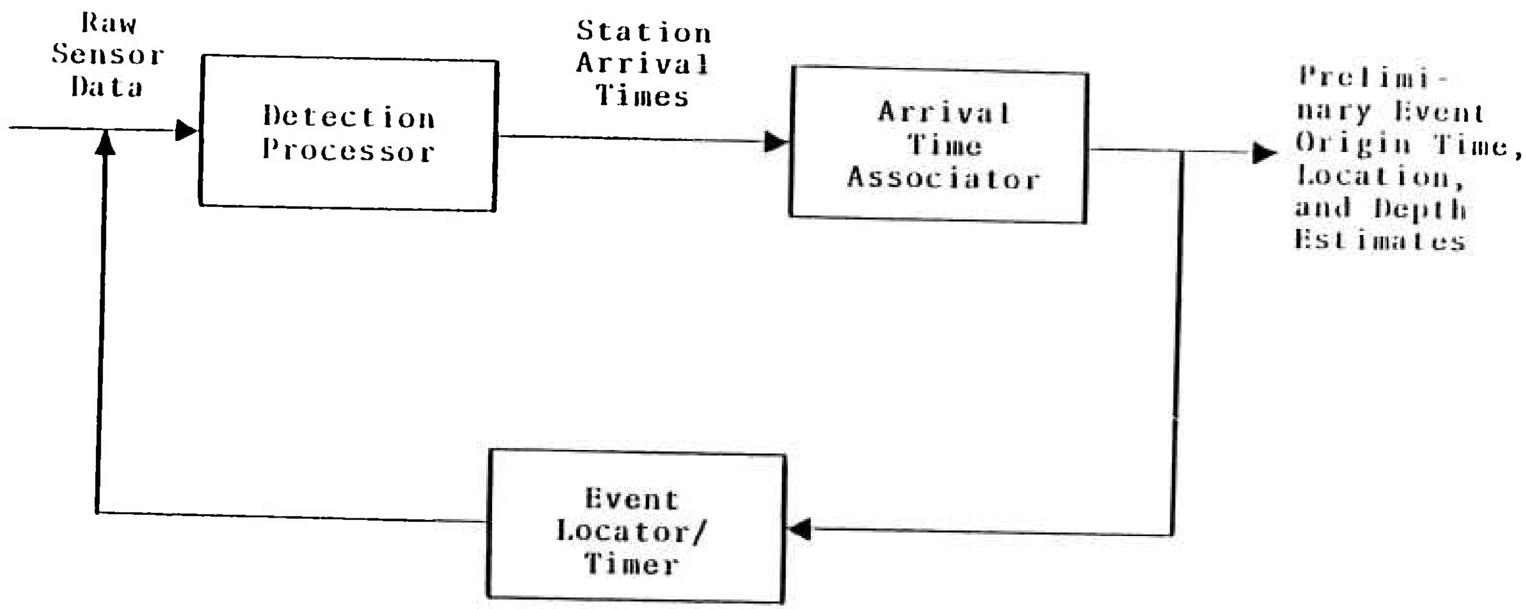


FIGURE I-1
PRELIMINARY EVENT LIST GENERATOR

- Phase association difficulties due to magnitude variation between stations and multiple transmissions
- Noise masking some of the multiple transmissions
- Difficulty of timing emergent and complex phases.

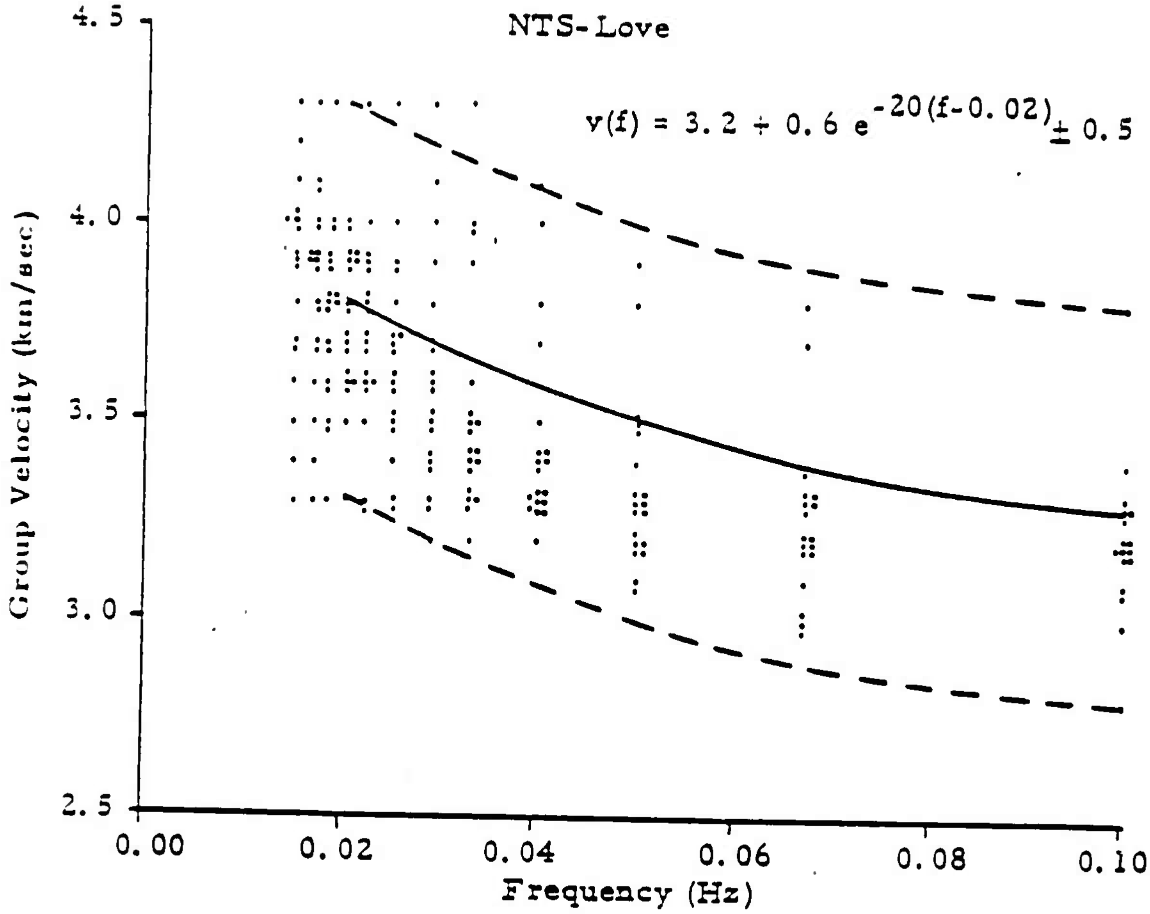
As a result of these uncertainties and difficulties in associating phases and making preliminary focal determinations; also, as a result of the present state-of-the-art of automatic association/location process; occasional large event location and timing errors are to be expected. Therefore, in order to design a robust ASE, it should be made capable of correctly retrieving event-signals, even under such adverse circumstances. Probably at least several percent of the preliminary locations will be in error by as much as 10° epicentral distance. Thus in searching for signals with possibly large location errors, the initially selected time-window should be at least 2 minutes in duration. Also, at least one minute additional is needed to assure sufficient noise preceding the signal. Another one minute interval is needed for coda following the signal. Thus, we recommend that the editor search a 4 minute time window for each desired signal. This was the time interval used by us in the Event Identification Experiment. This larger window prevented most of the serious problems our automated system would have encountered in retrieving signals of mislocated events. On the other hand, searching this larger initial time window increases the chance of encountering mixed events. These odds are estimated to be about 3% to 6%. As a consequence we will require more sophisticated application of post-detection information to correctly select out the valid event-signals. Our performance in automatically retrieving signals for the

Event Identification Experiment suffered from the lack of such post-detection processing to validate each retrieved signal. This situation needs to be improved and will be discussed later under recommended modifications.

b. Long-period signals

Due to the long duration of long-period surface wave signals, we saw no need to search for them with a detector. This is because the predicted signal time interval is small compared to the duration of such signals. Analytical relationships were derived for the broad-region dispersion of surface waves by Unger (1978). These are shown in Figure I-2 for the Asian continent and North American continent. Although some minor difference is indicated, a slightly broader envelope can be used to retrieve the narrowband passed surface wave groups. In the Event Identification Experiment, we determined a positive signal detection status to be 12 dB above the mean noise envelope preceding the predicted signal start time. We found this procedure to be satisfactory in almo-t all cases.

Regarding the retention of long-period surface wave data in permanent read-only files for advanced waveform analysis, we found that the short-period limit for extracting useful signal information was 12 seconds period. For an average 50° epicentral distance, approximately 900 seconds of data are required to sample Love waves and Rayleigh waves. A digital sampling period of 3 seconds most adequately covers the bandwidth required for wave groups of 12 seconds period or longer. Thus, two records of 300 words each are sufficient



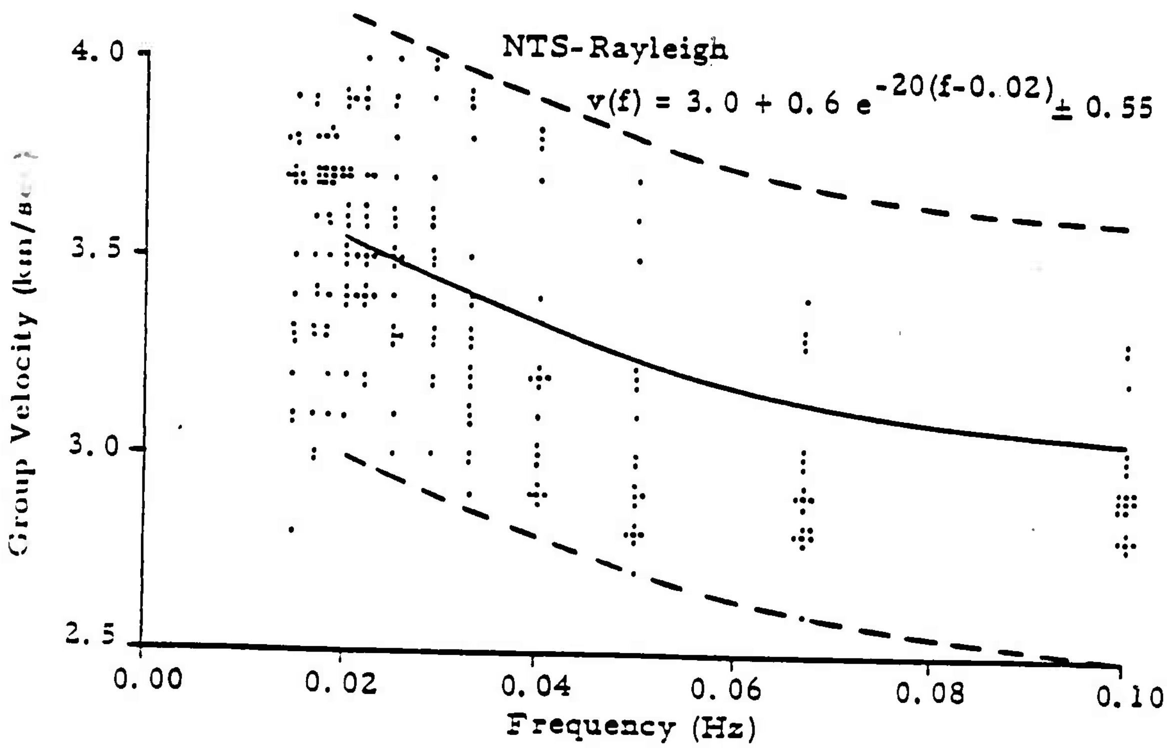
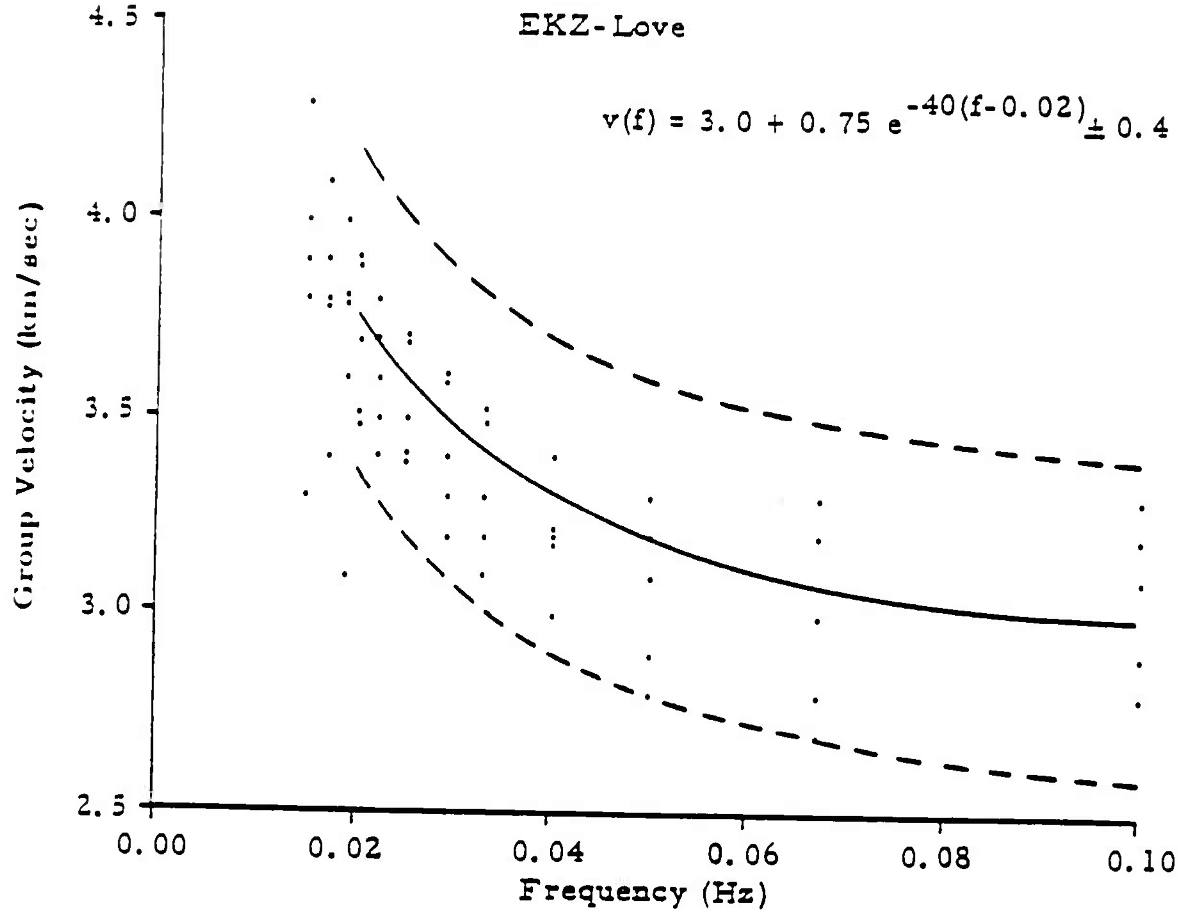


FIGURE I-2 (PART 1)

NORTH AMERICAN CONTINENT DISPERSION FOR WORLD-WIDE STATIONS



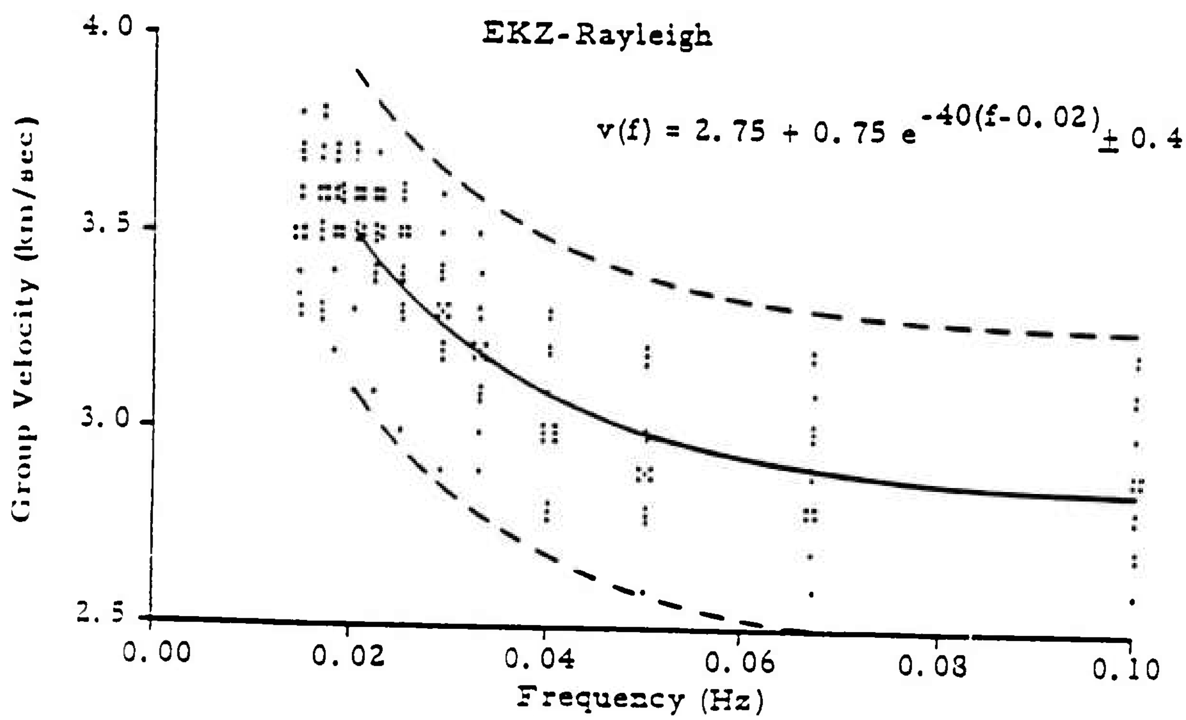


FIGURE I-2 (PART 2)

ASIAN CONTINENT DISPERSION FOR WORLD-WIDE STATIONS to create permanent read-only files for retention of long-period waveforms. Broadband and narrowband filtered noise level measurements of detected signals would be provided as header information provided with signals so that users can determine the detection status of signals. Network noise information can be provided by the procedure described for short-period signals.

Our experience with the Event Identification Experiment indicated that long-period surface wave signals could be successfully extracted from the data stream by deterministic timing of the signals using Unger's broad-region dispersion relationship. The software, LPEED, used for that purpose can probably be transferred with only minor modifications to the VSC system.

By contrast, short-period signals are of short duration compared to the much wider time window initially accessed. It is therefore necessary to apply statistical decision criteria to time and extract these short waveforms from that much longer time window containing noise. This problem is caused by the need to extract weak signals of short but unknown duration from ambient noise by the diverse and complex character of signals to be extracted, by multi-phase transmission of signals, and also by strong interference due to receiver scattering. To cope with these problems, we applied an automatic detector to the task of timing and extracting those short-period waveforms possibly associated with a given event. This will be described in the next subsection.

2. Application of Automatic Detectors to Extract the Short-Period Waveforms of Known Seismic Events

Once an event is located and timed by an automatic association/location process, the next step is to access time windows containing signals from the event. Data within signal time-windows are examined to determine the detection status of the desired signals. Those signal windows with good data but without detectable signals are reduced to a set of noise measurements. Those with bad data are flagged as inoperative. But when signals are detected, the system extracts the signal waveform for present and future quality control and maintenance of data bases. Also detected signals are reduced to a set of signal measurements for the on-going real-time operation of the VSC system. The decision tree required to perform this automatically is shown in Figure I-3. The data reduction accomplished by the decision tree is shown in the left margin of the figure.

In the Identification Experiment, we utilized a detector designed by Unger (1978) to detect and time short-period signals. Unger tested his detector against an analyst's detections finding that ninety percent of the signals detected by both an analyst and the automatic detector were timed within ± 0.5 seconds of the analyst's pick. Slowly emergent signals were picked several seconds late. This compared favorably for timing with the more conventional Z detector of Swindell and Snell (1977). By comparison, the Z detector times signals 1.70 seconds late on an average with a standard deviation of 2.20 seconds. For that reason, Unger's detector was selected for the Event Identification Experiment to automatically time

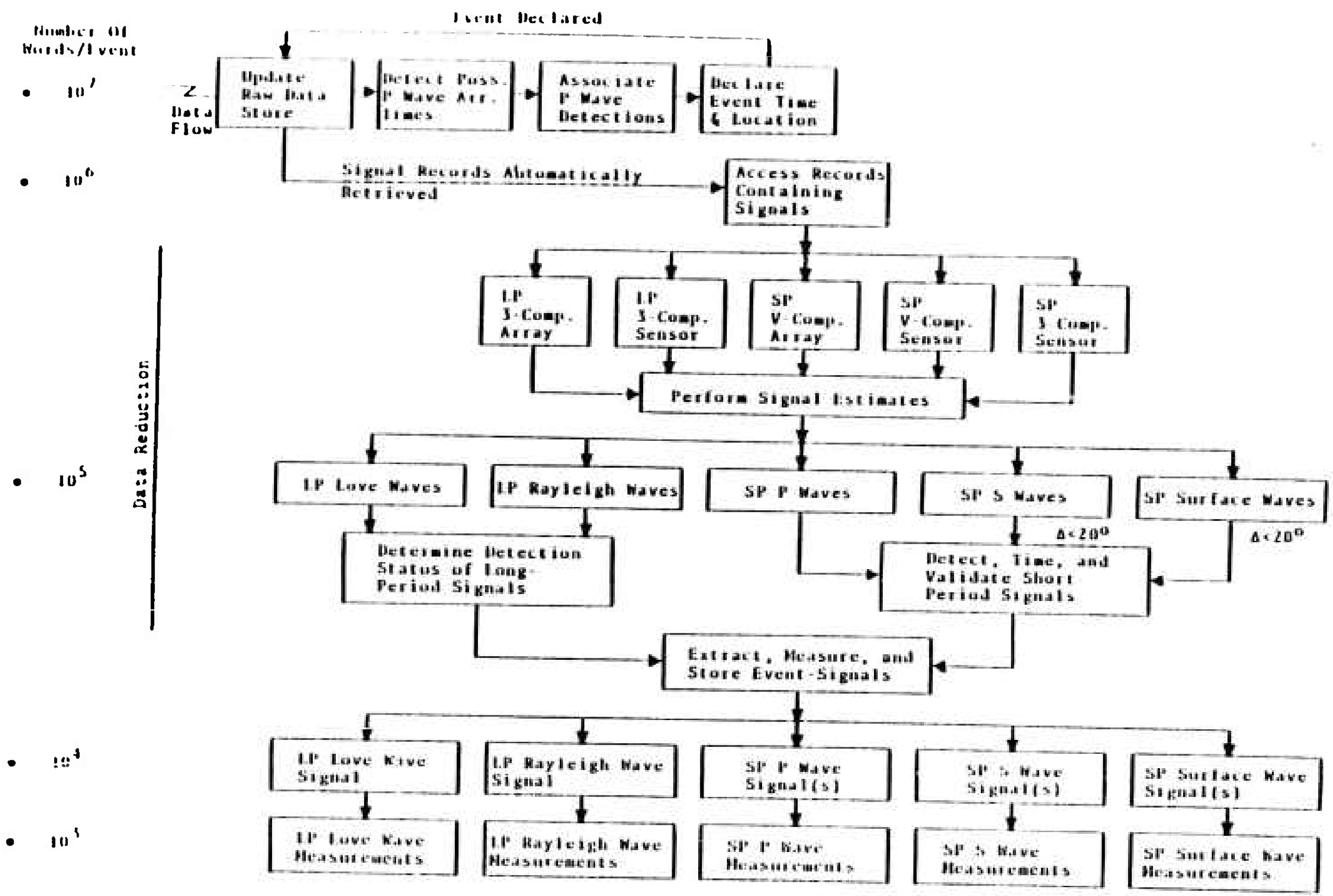


FIGURE I-3

TREE STRUCTURE OF THE AUTOMATIC SIGNAL EDITING PROCEDURE and edit short-period P wave signals from the long four minute records initially accessed. Our experience indicated that where signals were detected, they were accurately timed with only negligible false alarms. We noted however, that impulsive signals much shorter than the selected detection gate and slowly emergent signals were often missed by the detector. In some cases, even very obvious signals were missed because of spikes and glitches occurring in the warm-up noise preceding the signal window. Another serious problem was also observed in using Unger's detector for automatically editing signals. Extremely grave editing errors were caused by erroneously selecting mixed events and multiple phases. This stemmed from the lack of post-detection validation and a too simplistic detection strategy.

We consider problems caused by missed signals and editing blunders to be a very serious flaw in our present software for automatically editing signals. We also believe that these problems are clearly in focus, can be and should be solved. That goal will be a principal objective of this design study. An extensive modification of Unger's detector will be presented in a later section of this report. This new version of an automatic editor should correct the observed faults and will be recommended for future implementation into the VSC system. One important lesson learned from the Event Identification Experiment was that this detector was an important factor limiting our event identification performance.

SECTION II
DESCRIPTION OF UNGER'S ANALYTIC DETECTOR

A. THEORETICAL BACKGROUND

The basic concept of an analytic detector is to represent time data as the real part of a sampled complex exponential function. By this means one can measure signals and noise as modulated amplitude and phase time series. For example, complex seismic sensor noise can be represented as a sequence of random phase and amplitude modulations, as follows.

$$n_j = N_j \text{ Exp } i\phi_j \qquad (j = 1,2,\ldots,J). \qquad \text{(II-1)}$$

Similarly a seismic signal occurs within at least one subset of points on the record as

$$s_j = \begin{cases} S_j \text{ Exp } i\psi_j & (K \le j \le K+D), \text{ and} \\ 0 & \text{elsewhere} \end{cases} \qquad \text{(II-2)}$$

where D is the nominal duration of the seismic signal. The measured seismic data, x_j, is the sum of signal-plus-noise.

$$x_j = N_j \text{ Exp } i\phi_j + S_j \text{ Exp } i\psi_j \; . \qquad \text{(II-3)}$$

The detection problem is to time the start of the signal, K, and to extract and measure the signal. This representation was applied to the detection and measurement of seismic signals by Farnbach (1975) and Unger (1978). One of the earliest formal treatments of this method was given by Dugundji (1958).

The application of this type of phasor representation of signals and noise is practically accomplished by applying a digital filter to observed data. The filter passes signal and noise energy without any change in the amplitude but with a 90° change in the phase angle spectrum. This operator, called a Hilbert transform is described by Bracewell (1965).

This operation can be performed efficiently by means of a Fast Fourier Transform (FFT). It essentially interchanges the real and imaginary parts of the FFT. This is followed by inverting the data spectrum back to the time domain. Alternatively, Quadrature filtering is a much more computationally efficient procedure. With unit amplitude response it shifts each frequency 90°. This more practical operator yields results almost identical to the Hilbert transform.

The time series, y_j, produced by the Hilbert transform or quadrature filter is treated as the imaginary part of the data. The data, x_j, is itself considered to be the real part of the complex phasor record, Z_j, where

$$Z_j = x_j + iy_j. \qquad \text{(II-4)}$$

By this means, the record of data is transformed into an orthogonal sequence of rectangular complex coordinates. These two time series x_j and y_j are then transformed into a polar coordinate representation.

$$Z_j = A_j \ \mathrm{Exp}\ \phi_j \tag{II-5}$$

where

$$A_j = \sqrt{x_j^2 + y_j^2} \quad \text{and} \quad \phi_j = \tan^{-1} \frac{y_j}{x_j} \ .$$

The time series A_j is a measure of the instantaneous envelope of the data; the time series, ϕ_j, the instantaneous phase. The time derivative of ϕ_j is a measure of the instantaneous dominant frequency of seismic data.

In using this method to compute the instantaneous frequency, a serious stability problem was pointed out by Farnbach and Unger; especially when the envelope is small. This can be seen by deriving an expression for the instantaneous frequency as follows.

$$f_j = \frac{1}{A_j^2} \left[x_j + \frac{d}{dt} y_j - y_j \frac{d}{dt} x_j \right] . \tag{II-6}$$

From the above expression, it is seen that estimates of f_j become unstable when A_j becomes very small. It can be further seen that relative variation of f are three times those of the envelope.

$$\frac{\delta f}{f} = -3 \frac{\delta A}{A}$$

Thus, 10% point-to-point independent fluctuations of the envelope could result in a 30% error in frequency estimates. This problem of precise frequency measurement must be solved to effectively utilize phase measurements in the design of optimum analytic detectors. This will be discussed in a later section.

B. UNGER'S CRITERIA FOR TIMING SIGNALS

Unger (1978) constructed a remarkably simple model upon which to base his analytic detector. He analyzed the problem of detecting weak signals as one of detecting effects produced by the interference between signal and noise phasors. The signal in the time window of duration, D, is represented by a fixed signal amplitude level and initial phase. The noise is also represented as a fixed level but with random point-by-point transitions of phase.

$$x_j = N \text{ Exp } i\phi_j + S \text{ Exp } i\omega\Delta_o T_j \quad (K<j<K+D) \qquad \text{(II-7)}$$

$$= N \text{ Exp } i\phi_j\text{, elsewhere on the record } (j=1,2,\ldots,J)$$

The fixed phase of the signal is arbitrarily set to zero. For the purpose of analyzing detections this is without any loss of generality. Although it simply models the problem of detecting a weak signal added to noise, it is also an extreme oversimplification in that the random fluctuations of the envelope are neglected. Also, the phase of short-period seismic signals is generally more complex than a single simple deterministic phase modulated pulse.

Regarding the design of detectors, Unger demonstrated two important points with the above model. The probability of envelope measurements, A_j, exceeding the fixed noise level, N, is as follows. In the signal gate (K<j<K+D),

$$P(A_j > N) = \begin{cases} 1 - \frac{1}{\pi} \cos^{-1} \left(\frac{S}{2N}\right) & (0 < \frac{S}{N} < 2) \\ 1 & \frac{S}{N} > 2 \ . \end{cases} \qquad \text{(II-8)}$$

Thus, if the signal level is twice the noise level A_j can be expected to always exceed the noise level.

Even more interestingly, if the signal phase changes are uniform as indicated in equation (II-7) or indeed can be predicted by any deterministic model, then the signal can be detected as a stationary initial phase condition. Unger gives the probability of an observed stationary phase condition in terms of measured phase fluctuations, $\Delta\phi$, occurring within prescribed limits.

$$P\left(-\frac{\pi}{2} < \Delta\phi_j < \frac{\pi}{2}\right) = \begin{cases} 1 - \frac{1}{\pi} \cos^{-1} \left(\frac{S}{N}\right) & (0 < \frac{S}{N} < 1) \\ 1 & \frac{S}{N} < 1 \ . \end{cases} \qquad \text{(II-9)}$$

The expected probability of meeting the above detection conditions under noise is 0.5 for both equations (II-8) and (II-9). Under these equal likelihood of detection conditions, detection gain advantage of 6 dB is obtained by such a phase detector, provided that some deterministic phase model can be obtained for short-period seismic signals. Unger (1978) attempted this with a quadratic phase versus time model; but found that seismic signals commonly demonstrated random phase fluctuations not too dissimilar from seismic noise. As a result, Unger applied equation (II-8) as the model for design, his analytic detector which will be described in the next subsection.

C. THE DESIGN OF UNGER'S ANALYTIC DETECTOR

As a baseline for utilizing an analytic detector in an automated edit process, we will describe the design of Unger's detector and its application as an automatic signal editor in the Event Identification Experiment. There, the detector package was inserted in a supervisory routine (SPEED) which directs automatic control of signal timing and extraction to a detector and subsequently to a set of signal measurement algorithms.

First the short-period record is filtered with a broadband filter to subdue high frequency effects such as spikes and aliasing and low frequency effects such as data offsets and drift. This was done without distorting signals by placing the filter cutoff points at 0.3 Hz and 9.9 Hz for 20 Hz sampled data and at 0.3 Hz and 4.9 Hz for 10 Hz sampled data. This prefiltering was followed by application of the Unger (1978) detector.

The procedure of detecting and timing the onset of short-period (SP) signals is as follows (Figure II-1). First, over a specified warm-up period (e.g., 40 seconds), the peak noise envelope, $|\vec{n}|_{max}$, is established. This peak envelope is cosine tapered over subsequent waveform points, with a

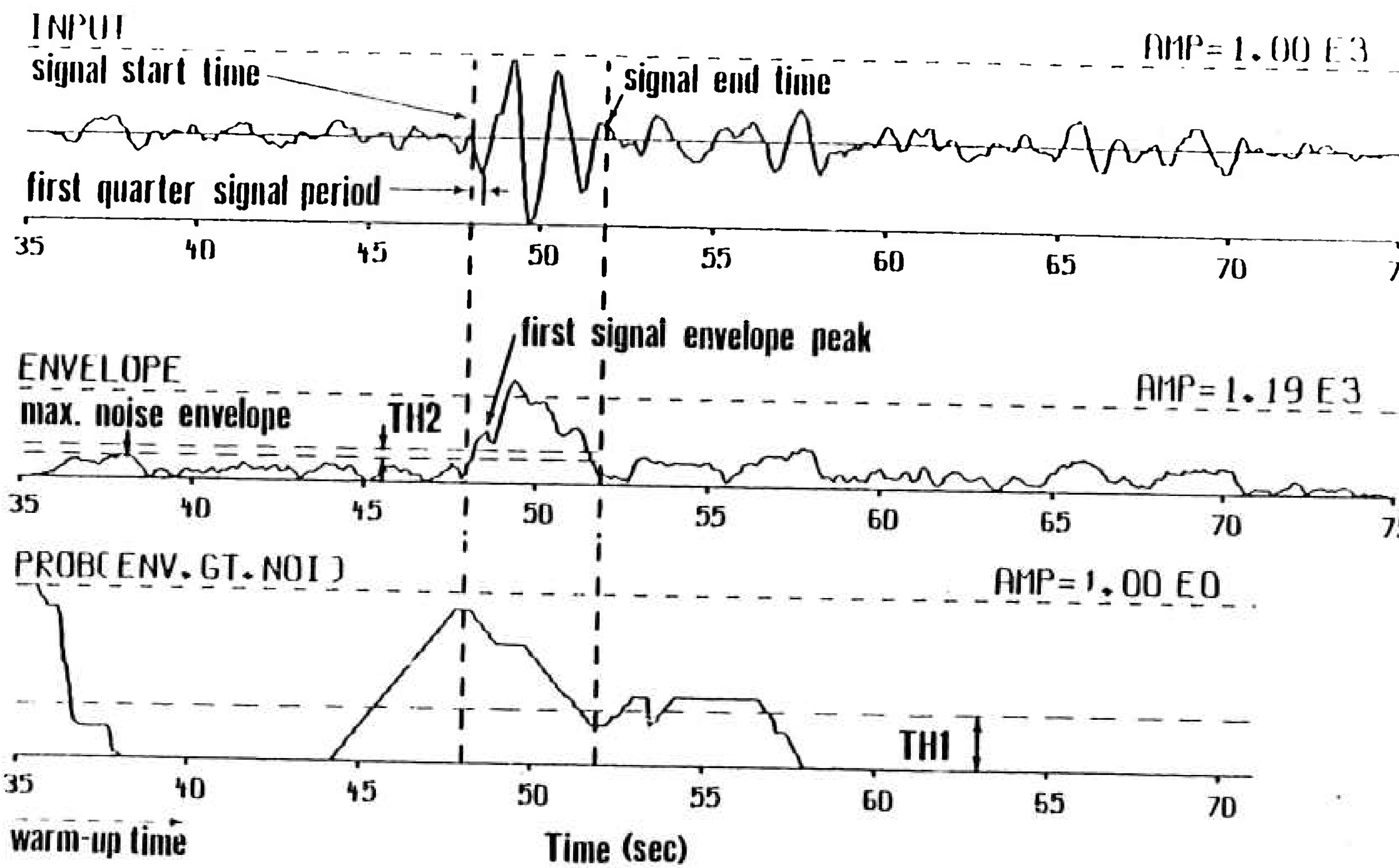


FIGURE II-1

SP SIGNAL DETECTION AND TIMING

specified time constant (e.g., with a 60-second time constant, the original peak value is halved at 30 seconds and equals zero at 60 seconds). An envelope value exceeding the tapered peak value established a new noise peak, unless a signal detection is declared; in that case no noise peak update takes place until the signal is declared to be terminated.

A signal detection is called whenever, in a forward looking (leading) time window of specified length (e.g., 4 seconds), the probability that the envelope is greater than the tapered peak noise envelope, $P(|\vec{r}_s(t)|>|\vec{n}|_{max})$ exceeds a spedified threshokd, TH1 (e.g., TH1 = 0.3). When this probability reaches its maximum the algorithm starts looking for the first signal envelope peak. When the ratio of first signal envelope peak and tapered noise envelope peak exceeds a second specified threshold, the SNR threshold TH2 (e.g., TH2 = 2 to 3 dB), the signal detection is confirmed and a frequency-dependent stepback is performed to determine the signal onset time.

The stepback procedure (Figure II-2) is based on the observation that in most cases the first signal envelope peak (at t_4) occurs within one signal period, and frequently at approximately 3/4 period, after the signal onset (at t_0). In a high-SNR waveform the signal onset time is most accurately found by detecting the first maximum or minimum of the signal's instantaneous value (at t_3), and stepping back 1/4 period (=0.25/instantaneous frequency at t_3). For low-SNR waveforms the first quarter period may be obscured by noise; in that case we step back 3/4 mean period (=0.75/mean frequency at t_4) from the first signal envelope peak at t_4. The mean frequency is the closed-form derivative of the phase regression polynomial

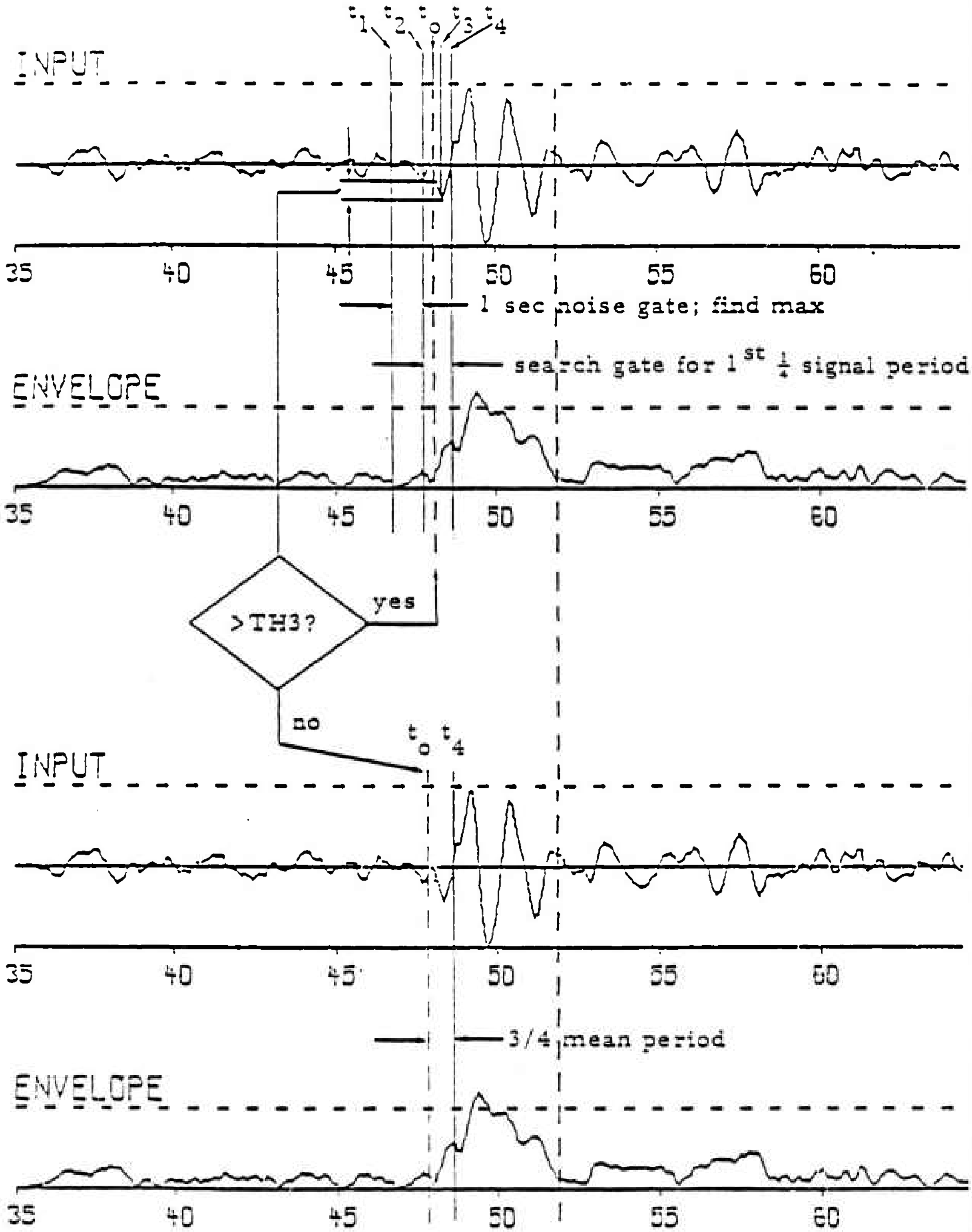


FIGURE II-2

STEPBACK PROCEDURE IN SP SIGNAL TIMING

evaluated at time t_4. The search for the first quarter period is started at t_2, i.e., at 0.8 mean period before t_4; the first quarter period is detected when its maximum or minimum exceeds, by a third threshold, TH3 (e.g., TH3 = 1 dB), the immediately preceding noise in the one-second time interval (t_1, t_2).

If the second threshold (the SNR threshold) is not satisfied, the detection is annulled and the noise peak value is updated with what at first was believed to be the signal envelope peak. Thereafter, the noise peak is updated as usual, until the next supposed signal detection, etc.

The signal end time is found as the moment of the first envelope minimum occurring either after $P(|\vec{r}_s(t)|>|\vec{n}|_{max})$ falls below its threshold, or after the signal duration exceeds a specified maximum, whichever is first. If this envelope is updated with this value, the noise peak updating and signal detection resume as normal. In principle this procedure enables the detection and timing of later phases and other signals in the coda.

From the previously described model, several important advantages can be cited. For signals less than 6 dB over noise, peaks of estimates of the $P(A_j>N)$ correspond to maxima of $S/N<2$. Thus, these indicators of weak signals satisfy the maximum likelihood condition independently of the statistical noise distribution. For signals greater than 6 dB over the noise, the detector saturates. Thus, the detector provides a robust means of detecting and initially timing the larger signals. In principal, this detector should optimally time the arrival of short-period P waves. It should be robust, also, because in estimating the probability of the envelope exceeding noise, it counts the number of times that the signal exceeds noise in a leading time window containing the signal. If properly implemented, it is potentially insensitive to unusually large and occasionally erratic envelope fluctuations in the noise or signal windows. By careful design, it should be insensitive to spikes, glitches, and other large amplitude errors.

D. POST-MORTEM EVALUATION OF UNGER'S ALGORITHM AS AN AUTOMATIC SIGNAL EDITOR

Our analysis of errors in applying Unger's automatic detector as a short-period signal editor lead us to suggest the following requirements for acceptable performance.

- Maintain a specified acceptable false alarm rate in retrieving signals associated with a known event.
- Detect, time, and validate almost as many signals as a seismic analyst.
- Separate signals of a known event from mixed signals of other events and other seismic phases.

Of these minimum requirements, we can only claim that the first one was met by our Event Identification Experiment experience. Therefore, the above requirements for automatic editing necessitate a re-design of the analytic detector algorithm. The following brief post-mortem evaluation is provided as background for the re-design.

Unger's analytic detector produced negligible false alarms and almost all of the validated detected signals were accurately timed including many cases of barely visible signals. Thus, the first requirement of an acceptable false alarm rate and part of the second requirement of accurate timing was satisfied by Unger's analytic detector.

Missed signals caused a very serious problem in identifying weak events. The detector algorithm operated only on envelope measurements of the seismic data, whereas, an analyst implicitly operates on a much more diverse set of signal characteristics. Given a reasonable false alarm rate we estimate that Unger's detector could only detect at a level of 70% of a typical analyst detection capability. For this reason, we recommend that detection of signals be based on a more diverse set of multivariate amplitude and frequency measurements of ground motion.

We observed another reason why some signals obviously visible to an analyst were missed by the automatic detector. Most seismic detector algorithms were optimized as an unknown complex transient added to stationary Gaussian noise. On that basis an envelope threshold is set to detect signals and control false alarms. Given such a model, such an optimum detector should be stable and the operating characteristics would then be determinable. In practice often such idealized assumptions are not true. For example very large amplitude deviations occasionally occur from environmental effects such as storms or from electronic malfunctions. Because Unger's algorithm gauges the maximum envelope of observed noise to determine the nominal noise level such large glitches or spikes can serve to effectively shut-off the detector. In such cases only very large events can be detected. Only signals much larger than the noise glitch would be detected. This happened to us enough times in the Event Identification Experiment to recommend modifying Unger's procedure of gauging the interfering noise level and replacing it by a more robust procedure.

Even when the noise data were of excellent quality we encountered numerous cases of missed visible signals. This was one of many problems caused by the diverse nature of seismic signals, in this case by emergent seismic signals. Unger's detector is designed to be optimum when signal starts at some point K and maintains a fixed level for duration D and then shuts off or gradually decays. This assumption is not at all optimum for detecting emergent seismic signals. In that case, Unger's maximum noise level estimate will be sequentially updated when the emergent signal rises above the noise but remains below the detection threshold of a signal. As a result, the threshold is continuously raised. A gradually emerging signal is then either missed altogether or detected with a very large time delay. This is another example where Unger's noise level updating procedure is not sufficiently robust to detect emergent signals. It is recommended that the presently applied noise level updating procedure be replaced by a more robust noise estimation procedure which is less sensitive to glitches and spikes and to interference effects produced at the beginning of emergent signals.

In this regard, we point out that Unger's algorithm was originally designed as a front-end continuous waveform

SECTION III
ANALYSIS OF AN ADVANCED ANALYTIC DETECTOR

In the preceding section, our post-mortem evaluation of Unger's detector indicated excellent performance in timing signals. That is for signals which both the automatic detector and analyst detected. However, in the case of emergent complex seismic events, accurate timing presented a problem which needs to be solved. But the real problem with the detector is the missed signal problem.

The automatic detector misses some large signals obvious to an analyst and many more small signals which can be detected by a seismic analyst. Some of the more obvious automatic detector misses could have been avoided by minor modifications to the detector.

The severity of the problem was such that for about 40% of the events, no signals were detected by the automatic detector. In those cases, discriminants had to be derived by networking noise estimates. This was to put upper limits on magnitude measurements which were derived to compute discriminants. Although this approach produced good results, it became obvious that signals missed by the automatic detector were primal in limiting event identification performance. Clearly, more than minor modifications to Unger's detector would be required to achieve a satisfactory level of performance.

Of the signals examined in the Event Identification Experiment, it is estimated that 10% of them which were missed by the automatic detector were obviously large signals easily detected by an analyst. These signals probably could have been detected automatically by relatively minor modifications of Unger's detector. Suggested modifications are as follows:

- Use ordered statistics to gauge noise thereby avoiding influence of occasional large glitches and spikes.
- Combine Unger's detector with a conventional Z detector to detect impulsive earthquakes and explosions.
- To avoid missing or late timing of emergent signals, count strings of ascending envelope measurements as above threshold if terminated by at least one envelope value above the threshold.

Even with these changes in Unger's detector algorithm we would still expect to miss about 20% to 30% of the events which could be detected by an experienced seismic analyst. This follows from the fact that Unger's algorithm detects on sustained signal power over some nominal signal duration, D. As such it is no different than most seismic detectors which attempt to detect signals solely by their enhanced broadband power. At reasonable false alarm rates, such detectors cannot be expected to detect more than 70% to 80% of the low signal-to-noise rate signals detected by seismic analysts, Swindell and Snell (1977).

To bridge this gap of automatically detecting weak signals with the something approaching the efficiency of a seismic analyst, we will analyze the design of a new multivariate analytic detector. This will be discussed in the following subsections.

A. GENERALIZED ANALYTIC SIGNAL AND NOISE MODEL

The theoretical background for the design of analytic detectors, described in Section II, is generalized to encompass the diverse types of seismic signals expected from earthquakes and explosions. These include multiple transmissions due to a complex source or to different propagation paths, variable length duration of signals, multiple pulse content of signals, and the fading coda scattering associated with signals.

Ambient noise preceding the time window searched for signals is generalized to possibly include spikes and glitches due to intermittent system malfunctions, environmental effects at the receiver site, and local seismicity.

1. Ambient Noise Model

Each record accessed by the automatic signal editor will contain at least two minutes of short-period data to be searched for signals. One half to one minute of noise data will be provided to gauge the noise statistical distribution. These noise statistics will be presumed to be fixed over the balance of the record possibly containing the desired signal. Equation (II-1) describes the complex seismic sensor noise at the station. This ambient noise model is generalized to include spikes, glitches, and local events. These will occur as short bursts of energy which complicate the signal detection decision function. The noise model is generalized to include large non-stationary energy fluctuations in the noise. For this we obtain

$$n_j = (N_j + \sum_{\ell=k_1}^{k_N} G_\ell \delta_{j-\ell}) \text{ Exp } i\phi_j \quad (j=1,2,\ldots,J) \qquad \text{(III-1)}$$

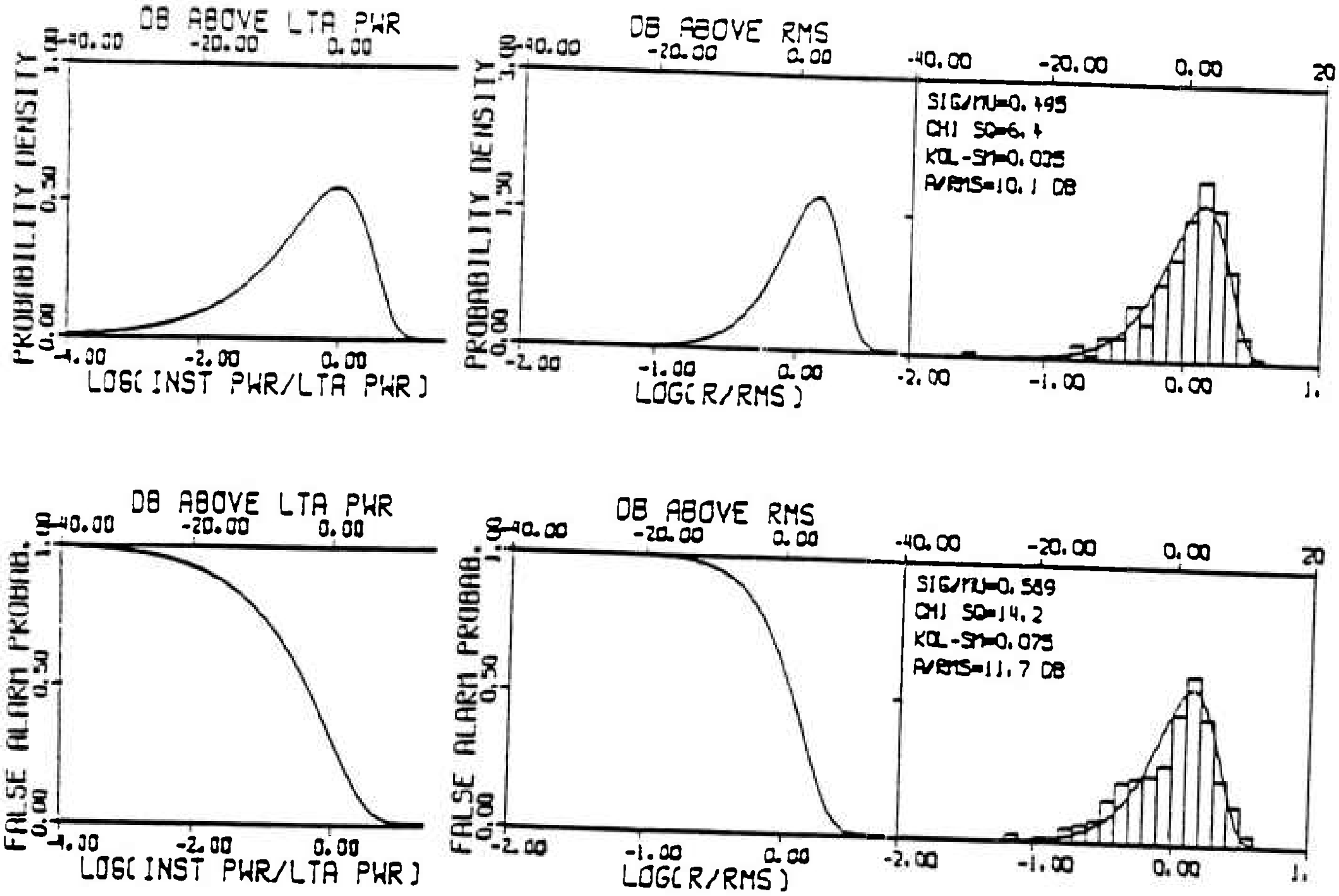
where

$\delta_m=1$ for m=0 or zero elsewhere; and

$k_1, k_2, \ldots, k_N$ are N random energy spike occurrences.

Unger (1978) performed a study of the distribution of N_j for seismic noise. Normalizing the noise fluctuations by dividing by the standard deviation, he interpreted the seismic noise trace as a stationary Gaussian distribution. In that case, the envelope of that noise would be a Rayleigh distribution. Since measurements of signal envelope are primarily used to derive magnitudes, he also determined the distribution $w = \log_{10} N_j$; the noise magnitude distribution.

Based on the analysis of one hour of seismic noise at the Korean Seismic Research Station (KSRS) broken down into 102.4 second data segments, Figure III-1 shows Unger's results in fitting measured noise magnitude distributions to the distribution derived from a stationary Gaussian noise assumption. Unger tested the significance of the fit of measured noise to a Rayleigh distribution. At the 5% significance level 80% of the measured distributions were consistent with the Gaussian



a. Gaussian log-power

b. Gaussian log-envelope

c. Observed log-envelope compared to Gaussian

THEORETICAL AND OBSERVED NOISE DISTRIBUTIONS

FIGURE III-1

noise hypothesis. Of the 20% of rejections, the Gaussian hypothesis slightly underestimated the number of large magnitude fluctuations. Thus, these rejected samples appeared as almost Gaussian distributed.

Based on the Gaussian noise assumption, certain facts can be ascertained about magnitude measurements of seismic noise. Some of these were given by Unger (1978) others were derived from graphical analysis of his results. The most probable occurrence of magnitude is at the $\log_{10}$(RMS) of seismic noise. The standard deviation of large noise fluctuations above the $\log_{10}$(RMS) of noise is approximately 0.20 magnitude units. Measurements at seven Seismic Research Observatory (SRO) stations (Strauss and Weltman, 1977; Weltman, et al., 1979) indicated that peak one-second noise amplitude measurements had estimated magnitude standard deviations between 0.17 and 0.23; and four stations, between 0.13 and 0.16.

The shape of the distribution of seismic magnitudes based on Unger's observations as well as on the Gaussian noise assumption is highly skewed. Only one third of the envelope measurements are expected above the RMS noise (the most probable noise magnitude). Of the two thirds of the envelope measurements below the RMS noise level, the RMS of these negative magnitude fluctuations is approximately 0.5 magnitude units. On this basis it is estimated that in 20% of a time window containing the seismic noise envelope measurements are lower by more than a half magnitude than the noise RMS level.

At intervals of time where such low level noise envelopes are encountered, a skilled analyst can possibly recognize and detect signals by intermittent changes in the character, such as by changes of frequency and the duration of such changes. Perhaps this might explain, in part, the 20% of signals which can be detected by skilled analysts; but which are always missed by detector algorithms based solely on the signal energy level. It suggests that weak signals might sometimes be detected at noisy stations at levels as much as one half magnitude higher than the expected signal magnitude. If further suggests that the algorithm for automatically editing and measuring signals be based on more than energy level criteria alone.

As for our automated signal editing procedure, the generalization of the noise model of equation (III-1) can be applied to the following tasks.

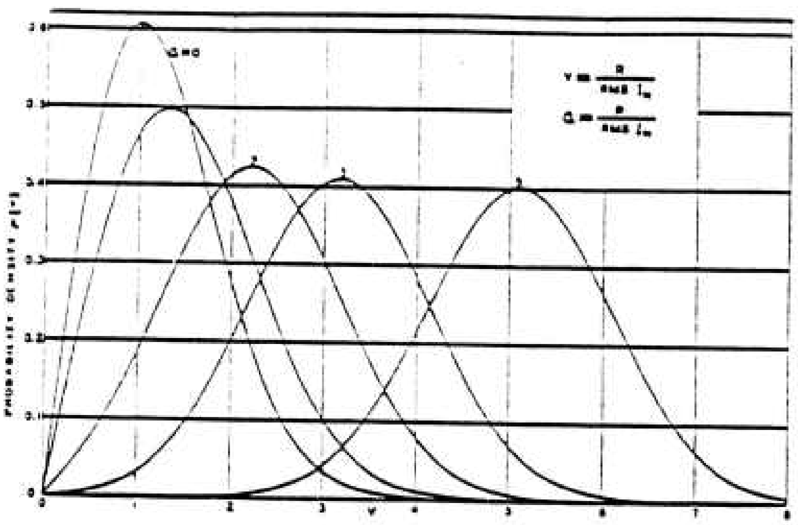
- Design robust noise parameter estimation techniques which are insensitive to occurrences of spikes, glitches, and local events.
- Automatically detect and flag malfunctioning stations.

The first step in this analysis is to access the one minute window containing seismic noise. This noise data precedes the three minute window possibly containing signal and coda. As is usual in robust estimation procedures, noise measurements are analyzed by application of ordered statistics. This is to avoid influencing the result by heavily weighting large deviations due to spikes, glitches, and local seismic events, as shown on equation (III-1). Envelope measurements of the noise are sorted from smallest to largest. By taking the logarithm to base ten of the noise envelope measurements, the envelope measurements are transformed into a set of ordered noise magnitude measurements. Applying Unger's derived distribution for log-envelopes, the following percentiles are determined by interpolating the ordered set of measured noise magnitudes.

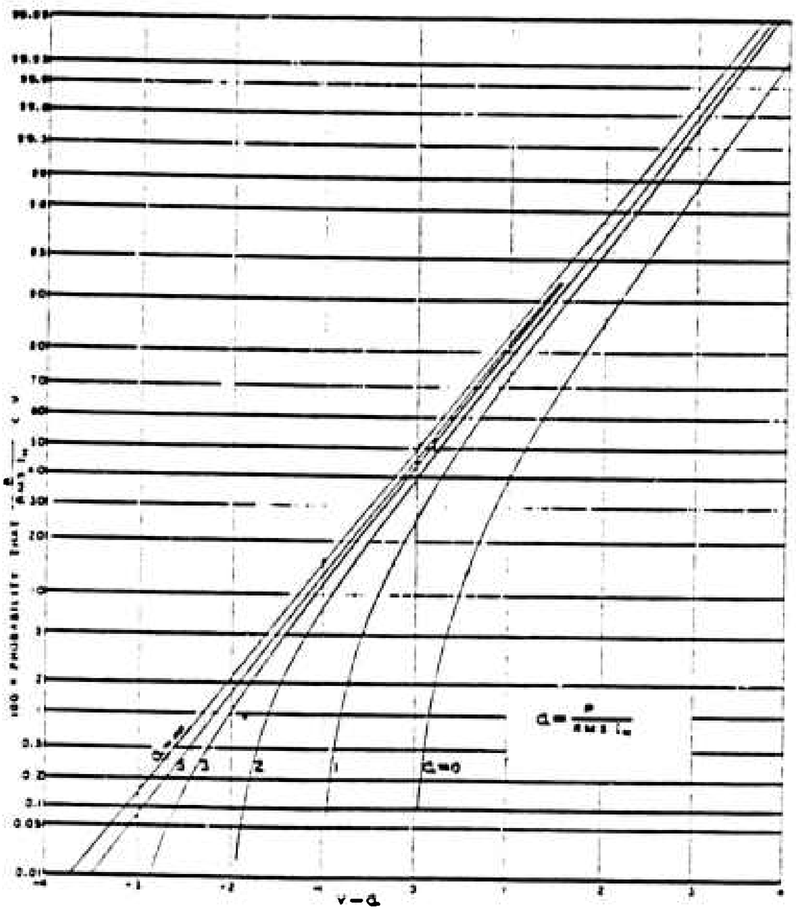
- The 66 percentile level of smaller magnitudes corresponds to $\log_{10}$(RMS) of the noise trace and corresponds to the most probable occurring envelope magnitude.
- The 88 percentile level minus the 66 percentile corresponds to the standard deviation of positive noise magnitude fluctuations, expected to be approximately 0.20 magnitude units.
- The 66 percentile minus the 20 percentile level corresponds to the standard deviation of negative noise magnitude fluctuations, expected to be approximately 0.50 magnitude units for Gaussian noise.

The expected values of the positive and negative fluctuations of noise magnitude were observed by Unger at KSRS, and correspond to what is expected for ideal bandpass filtered Gaussian noise. By contrast, noise from ocean generated microseisms is characterized by a single sharp spectral peak between 0.15 and 0.35 Hz. Rice (1954) analyzed this situation of a periodic process added to random Gaussian noise. If the periodic process is large, the noise distribution of envelope measurements shifts from the Rayleigh distribution (the envelope distribution of Gaussian noise) to a normal distribution with standard deviation of the RMS of the additive Gaussian noise. The mean of the normal distribution is centered at the amplitude of the periodic process. This is shown in Figure III-2. Thus, unless the influence of the microseismic peak is minimizes by filtering or some other means, Unger's application of the Guassian noise model would have to be modified. In the case of a large nearly periodic component in the seismic noise, noise magnitude fluctuations would be nearly symmetrical for positive and negative deviations and almost normally distributed. For a spectral peak at three times the RMS Gaussian noise, the standard deviation of envelope magnitudes would be about 0.15 magnitude units. The SRO noise magnitude measurements suggest that at least some seismic stations would be close to this alternative noise model. In those cases, techniques would need to be applied to minimize the effect of sharp spectral peaks occurring in the noise to optimize detection performance.

A systematic procedure for applying ordered statistics to magnitude measurements is given by Sax, et al. (1979) and is illustrated in Figure III-3. Based on Unger's envelope study, the percentiles applied to noise data which is nearly Gaussian are given above. By applying this procedure, magnitude measurements of seismic signals and noise are reduced approximately to homogeneous normal statistics. The discrimination operating results, after using this normalization procedure shown on Figure III-3, could be closely predicted by application of normal error theory. For this reason, it should be applied to the problem of detecting signals, where the distribution of noise and signal magnitudes may be highly skewed.



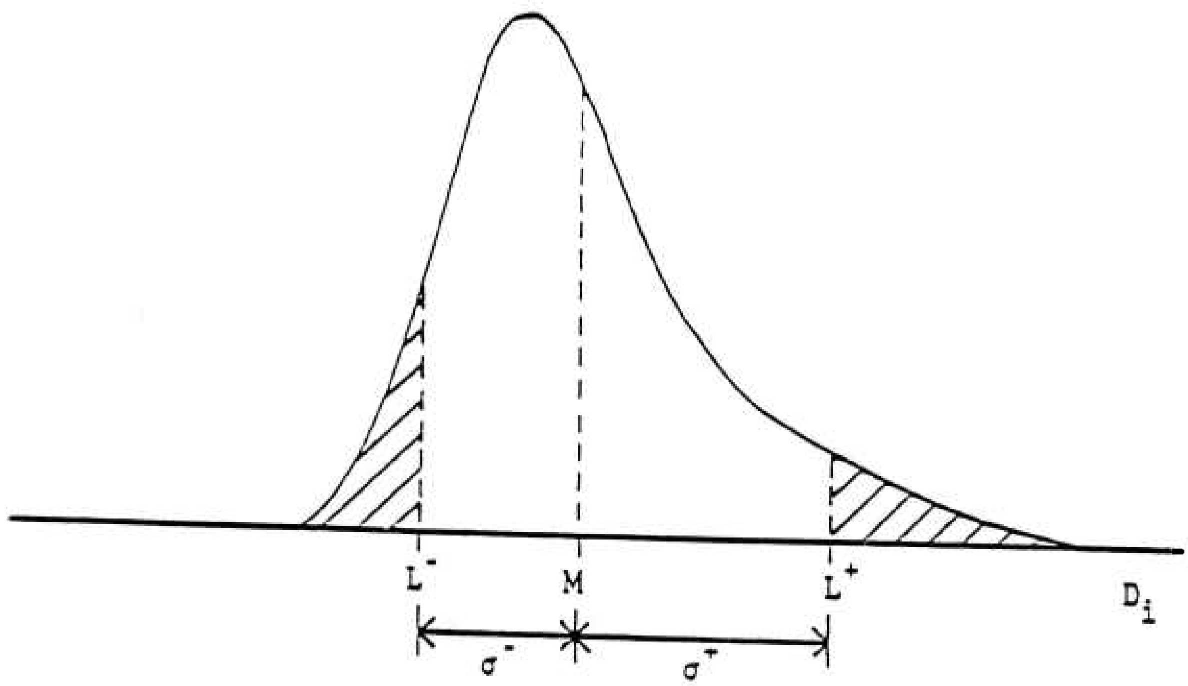
·Probability density of envelope R of $I(t) = P \cos pt + I_N$



Distribution function of envelope R of $I(t) = P \cos pt + I_N$

FIGURE III-2

A PERIODIC PROCESS WITH ADDITIVE GUASSIAN NOISE



- LEVELS OF VARIATION

  L_i^+: 85% of $D_i \le L^+$ 85% Level

  M_i: 50% of $D_i \le M$ 50% Level

  L_i^-: 15% of $D_i \le L^-$ 15% Level

- PARAMETERS FOR NORMALIZATION

  MEDIAN : M_i

  POSITIVE DEVIATION: $\sigma_i^+ = |L_i^+ - M_i|$

  NEGATIVE DEVIATION: $\sigma_i^- = |L_i^- - M_i|$

- NORMALIZATION OF D_i TO APPROXIMATE UNIT NORMAL STATISTICS

$$Z_i = \frac{D_i - M_i}{\sigma_i} \quad \text{where} \quad \sigma_i = \begin{cases} \sigma_i^+ & \text{if } D_i \ge M \\ \sigma_i^- & \text{if } D_i < M \end{cases}$$

FIGURE III-3

NORMALIZATION OF SKEWED DETECTION STATISTICS

This ordered statistical analysis can be used to obtain the most probable occurrence of noise measurements and the standard deviation of positive or negative fluctuations. This reduces deviations from the most probable noise occurrence to unit normal statistics. In singal windows a set of significant deviations from a unit normal distribution can be used to detect and time signals.

Unusual deviations can also detect and diagnose automatically those stations which are malfunctioning. Note that unusually large envelope fluctuations, G_k, in equation (III-1) can be seen as unusually large noise magnitude deviations, but will not influence the noise parameters used to detect signals.

In validating a properly functioning station, the median, positive, and negative standard deviations can be checked against historical data. If these parameters fall within acceptable limits, e.g., falls within a 99% acceptance region, the station noise data are further checked for large intermittent errors, G_k. If the average of observations above the one standard deviation limits is more than two standard deviations, then those measurements closest to the outlier average are counted as noise spikes or glitches, G_k. Positive G_k's may indicate spikes or clipped data. Negative G_k's may indicate zero returns or intermittent dead spots. If the number of G_k's exceed an acceptable value, the station is classified as a malfunction.

As part of the quality control built into the automatic editing detector, the noise window is searched for sparsely occurring spikes, G_k. If more than a specified number of spikes occur above three standard deviations, these are interpreted as sparse occurrences of spikes. In that case, the amplitude and phase measurements of the record are smoothed (similar to 'liftering' in cepstral analysis) to minimize the effect of sparse occurrences of G_k. This will usually have small distortional effect on signals since their duration will be at least one half second or greater. At the same time it will minimize the influence of sparsely occurring spikes or much smaller duration.

The above quality control criteria is merely a preliminary specification of the procedures needed to validate properly functioning stations and to condition data collected by marginally functioning stations. Obviously, these criteria need to be tested on stations exhibiting various types of malfunctions. Such tests will result in modified optimized procedures for validating stations and conditioning accessed signal records. In the Event Identification Experiment this quality control was done by human analysts. In the VSC system, we anticipate that the station editing quality control will need to be done automatically in order to quickly reduce the volume of seismic data to manageable proportion.

2. Generalized Signal Model

Having passed through noise quality control, signal data are processed to detect and time seismic signals. In order to obtain valid signal measurements, such detections need to be modeled with sufficient generality to verify that measurements will be made on proper P wave signals. The generalized signal model must therefore include effects produced by multiple phase reception of the signal, scattered coda characteristics, and instrument malfunctions.

The multiple signals accounted for by the model should include the following.

- Complex source signals and depth-phases, each propagating world-wide as first-motion P waves
- Surface- and core-reflected phases propagating world-wide as later phases of each P wave signal
- Long-distance propagated PKP and PKIKP core phases
- Localized multiple transmissions and receiver-scattered phases not correlated at other stations of the network
- Mixed signals from another event.

This model describing signal phases which might be detected in signal time windows represents a complicated situation. Therefore, the signal model is changed to handle this situation.

The simpler single phase signal model of equation (II-2) is generalized to that of multiple phase transmissions from a complex source. As in the noise model, we also include G_ℓ to represent intermittent system noise in the accessed signal window. The generalized signal model is given by equation (III-2).

$$s_j = \left(S_j - \sum_{\ell=r_1}^{r_s} G_\ell s_{j-\ell}\right) \operatorname{Exp} i\,\Psi_j \qquad (K_{mn} \leq j \leq K_{mn} + D)$$

$$0 \qquad\qquad \text{elsewhere} \qquad\qquad \text{(III-2)}$$

where

$m = 1, \ldots, M;\ n = 1, \ldots, N.$

The model is generalized to MN signal source-transmissions propagated to each seismic station. This model indicates a complex source generating a sequence M signals over a time interval of one or two minutes. Each signal is propagated as a primary P wave (n=1) followed by a sequence of later arriving secondary phases (n=2,...,N), provided that phases such as PP, PcP, etc. occur within the time frame of the accessed record. Thus, the sequence of start-times, K_{mn}, can be viewed in a network sense as a sequence of origin times, each augmented by the propagation time delay of P waves or secondary phases.

Complex explosion sources are realized by detonating a sequence of explosions at the source. Complex earthquakes are commonly observed from some seismic regions. Sax (1979) observed such sequences, which from array processing were apparently propagating as source delayed P waves in accordance with the above model. The duration of signals, D, is also viewed as a source-dependent parameter which varies nominally between one half to ten seconds. Small, high stress and high stress drop sources would tend toward the low end of the duration range as simple and highly impulsive events. Complex signal waveforms, associated with large low stress drop events from highly heterogeneous source regions, would tend toward the high end of the duration range. The design of the detector must be sufficiently general to cope with these realistic but more complex signal editing situations. The situation is further complicated in that signal magnitudes associated with a complex source may also vary considerably, i.e., the magnitude of later source emissions may be larger than earlier ones.

This obviously complicates the detection validation problem of correctly associating independent edits at one station to those at other stations of the network. This problem of validating station detections is considerably simplified by using the criteria of a maximum acceptable travel time anomaly between the arrival of an elemental source phase at any two stations of the network. It is important to validate signal edits not only to assure propagation characteristics consistent with the given event location but also provides a sound relatively easy basis for automatically shifting the event focus to a position which minimizes observed travel time anomalies.

In addition to utilizing this criteria of propagation consistency to validate automatically edited signals, we will also utilize magnitude consistency. The difference in magnitude of network associated phases should be maintained within a maximum acceptable tolerance. Although source magnitude consistency is a weaker criteria than propagation consistency because of the variability of magnitudes between stations, it is very important to utilize it to avoid large blunders caused by false associations of an event with large receiver-scattered phases or mixed signals.

To complete the analytical signal model we include a term to encompass other apparent signals which can be detected at a station. These signals do not correlate across the network and therefore cannot be validated by means of propagation consistency criteria. In some cases, signal measurements of these uncorrelated phases could be very misleading.

These uncorrelated signals from the source can in some cases be associated with the desired event. There could be one or two delayed secondary images of the source caused by rapid upper mantle increases in the propagation velocity. Such phases can be large but would be seen only in a narrow distance band from the source and therefore would not be generally correlated across the network. Also, strong signals could be produced from an inhomogeneous source medium as higher order multipole components. These would only be seen in narrow azimuth bands and also not generally be correlated across the network.

Other uncorrelated signals are not directly related to the source and can be caused by large envelope fluctuations due to receiver scattering by interfering signals from some other known event, singular observations of an unknown event, i.e., local or regional, or noise false alarms. In the latter case, we may have two events with overlapping records at some stations. It is important that these signals from mixed events be properly sorted out by our post detection analysis.

As for representing these detectable signals which occur singularly and cannot be verified as propagation consistent, we will generalize the signal model of equation (III-2) as follows.

$$s_j = (S_j + \sum_{\ell=r_1}^{r_s} G_\ell \delta_{j-\ell}) \text{ Exp } i \Psi_j + U_v \text{ Exp } i \Omega_v \qquad \text{(III-3)}$$

where

$$(K_{mn} \leq j \leq K_{mn} + D), \; (K_n \leq v \leq K_n + D').$$

The start times K_n are of false signals. These cannot be verified either as propagation consistent or as source magnitude consistent. They will generally be rejected as noise. One of the advantages of this model is that the decision threshold can be set much lower to detect weak signals. By applying post detection analysis, false alarms from ambient noise and other undesireable detections can be weeded out of the automated signal editing process. The purpose of representing false signals of type U_v in equation (III-3), is to establish criteria for identifying such signals by the lack of network propagation consistency, source magnitude consistency, or other criteria and to keep such false detections from contaminating the signal measurements to be used for event discrimination.

B. MULTIVARIATE AMPLITUDE AND PHASE STATISTICS

Multivariate statistics are needed to effectively retrieve small signals below the RMS level of seismic noise. This need stems from the often observed gap between the signal retrieval capability of a seismic analyst compared to that of an automatic algorithm which is based solely on power fluctuations.

In using Unger's algorithm to time, measure, and retrieve signals in the Event Identification Experiment, we failed to retrieve one or more signals for about 40% of the events analyzed. In several percent of the cases, we not only missed retrieving signals of a desired event but committed the more serious error of retrieving much larger signals from some other unknown event occurring on the same seismic record.

About 25% of these missed signals and most of the serious editing blunders could have been avoided by modifying Unger's algorithm; in effect using more robust techniques to implement the algorithm and performing post-detection analysis to validate signals. Yet, at reasonable false alarm rates, approximately 30% of the signals detected by seismic analysts simply cannot be retrieved by automatic detector algorithms based solely on large power fluctuations. Unfortunately, most, if not all conventional automatic seismic signal detection algorithms fall into this category. For this reason, we decided that simply making small modifications of Unger's algorithm would not result in a sufficiently improved level of signal retrieval performance. To achieve that, his decision algorithm would need to be generalized to operate on the basis of multivariate amplitude and phase measurements.

The goal of utilizing a multivariate automatic signal editor is to retrieve signals of known events with the effectiveness of a seismic analyst. As previously discussed, magnitude measurements of seismic noise are highly skewed. Noise magnitudes computed as the logarithm of envelope measurements, as is done by Unger's algorithm, occur most frequently at the RMS level of the seismic noise. The standard deviation about the most probable occurring noise magnitude is 0.5 magnitude units for negative deviations; and 0.2 magnitude units for positive deviations.

As a result of this skewness, small seismic signals up to 0.5 magnitude less than RMS noise can be expected to be larger than the noise in 10 to 20% of the signal windows examined. This provides an analyst, skilled in recognizing signals by their waveform character, a reasonable opportunity of detecting at least one or two weak signals from a network. We expect that by discerning change of signal frequency as well as amplitude over the expected duration of a possible signal, the analyst can extract, time, and measure signals well below the threshold of an automatic power detector. Such small signals would always be missed by a power detector.

Given the task of automatically retrieving signals of known events, we will generalize the detector so it can recognize signals based on a multivariate set of observations of frequency and amplitude. By doing this, there is at least some possibility of attaining the level of signal retrieval performance expected from a seismic analyst. This is expected to yield as much as 0.5 magnitude units of enhanced magnitude capability of extracting small signals.

1. Definition of a Standard Set of Multivariate Seismic Amplitude and Phase Measurements

The basic problem of defining a set of standard multivariate measurements is to encompass the physical characteristics of earthquake or explosion seismic sources. The first step in achieving this is to transform seismic measurements to a multivariate set of ground motion measurements. In principal, this theoretically whitens the source between two source dependent frequencies. By sensing the range of peak frequencies of a possible signal, an optimum filter can be automatically designed to extract the signal.

The multivariate set of ground motion measurement will be based on the system response removal function now being implemented into the VSC seismic system. The seismic record will be transformed into measurements of the integral of the ground displacement, ground displacement, velocity, acceleration, and the derivative of the acceleration. From Randall (1973), the asymptotic behavior of earthquake source models in terms of normalized frequency, $x=f/f_c$, where f_c is the corner frequency, are given as follows for the displacement amplitude spectrum in the form $\Omega(\omega)=\Omega(0)F(x)$.

a. A function derived by Keilis-Borok (1959) equated to the low- and high-frequency asymptotes intersecting at the corner frequency,

$$F(x) = 1 \;, \quad (0 \leq x \leq 1)$$
$$\quad\;\; = x^{-2}, \quad (x>1).$$

b. The function considered by Brune (1970),

$$F(x) = (1+x^2)^{-1} .$$

c. The spectral shape for either of the source models of Randall (1966) and Archambeau (1968),

$$F(x) = 3^{\frac{1}{2}} x^{-3} \left[\sin(3^{\frac{1}{2}}x) - 3^{\frac{1}{2}} x \cos\ (3^{\frac{1}{2}}x)\right] .$$

d. A function given by Randall (1973) defined asymptotically like function 1 but assuming a w^{-3} high frequency behavior,

$$F(x) = 1 \qquad (0 \leq x \leq 1)$$
$$= x^{-3} .$$

e. Two models given by Aki (1967) were derived by Haskell's method from the spatial and temporal correlation of the velocity of a fault dislocation,

$$F(x) = (1+ax^2)^{-\frac{1}{2}} (1+x^2)^{-\frac{1}{2}}, \quad w^2 \text{ model}$$

$$F(x) = (1+ax^2) \quad (1+x^2) \quad , \quad w^3 \text{ model.}$$

f. Mueller (1969) derived the displacement amplitude spectrum of explosion sources from the Latter, et al. (1959) model,

$$F(x) = x^1 (x^6+a_1x^4+a_2x^2+a_3)^{-\frac{1}{2}} .$$

Randall (1973) showed that expressions derived to compute seismic energy and characteristic stress are independent of assumptions as to source model. Thus, all of the models previously listed provide a reasonable basis for designing broadband ground motion filters to extract the spectral characteristics of seismic signals. These are shown schematically in Figure III-4.

The models illustrated in Figure III-4 show that moments of the signal spectrum are generally whitened over a portion of the frequency band controlled by the source. In practice, this effect sometimes will be masked by path absorption and receiver scattering requiring source region-station calibration.

The k^{th} spectral moment of the normalized displacement amplitude distribution as a function of frequency, F(x), is given by

$$\mu_k = \int_0^\infty x^k F(x)\, dx \quad .$$

Time domain estimates of μ_k can be obtained by taking the k^{th} derivative of the seismic record, detecting and timing the signal, and measuring the magnitude of the signal. By measuring the frequency corresponding to envelope peaks, an estimate can be obtained of the frequency band of the k^{th} spectral moment.

The time domain operators which whiten the signal models in Figure III-4 are the k^{th} derivatives where $k=-1,0,1,2,3$. Frequency and magnitude measurements of these derivatives should in principle completely specify the source characteristics of the seismic event.

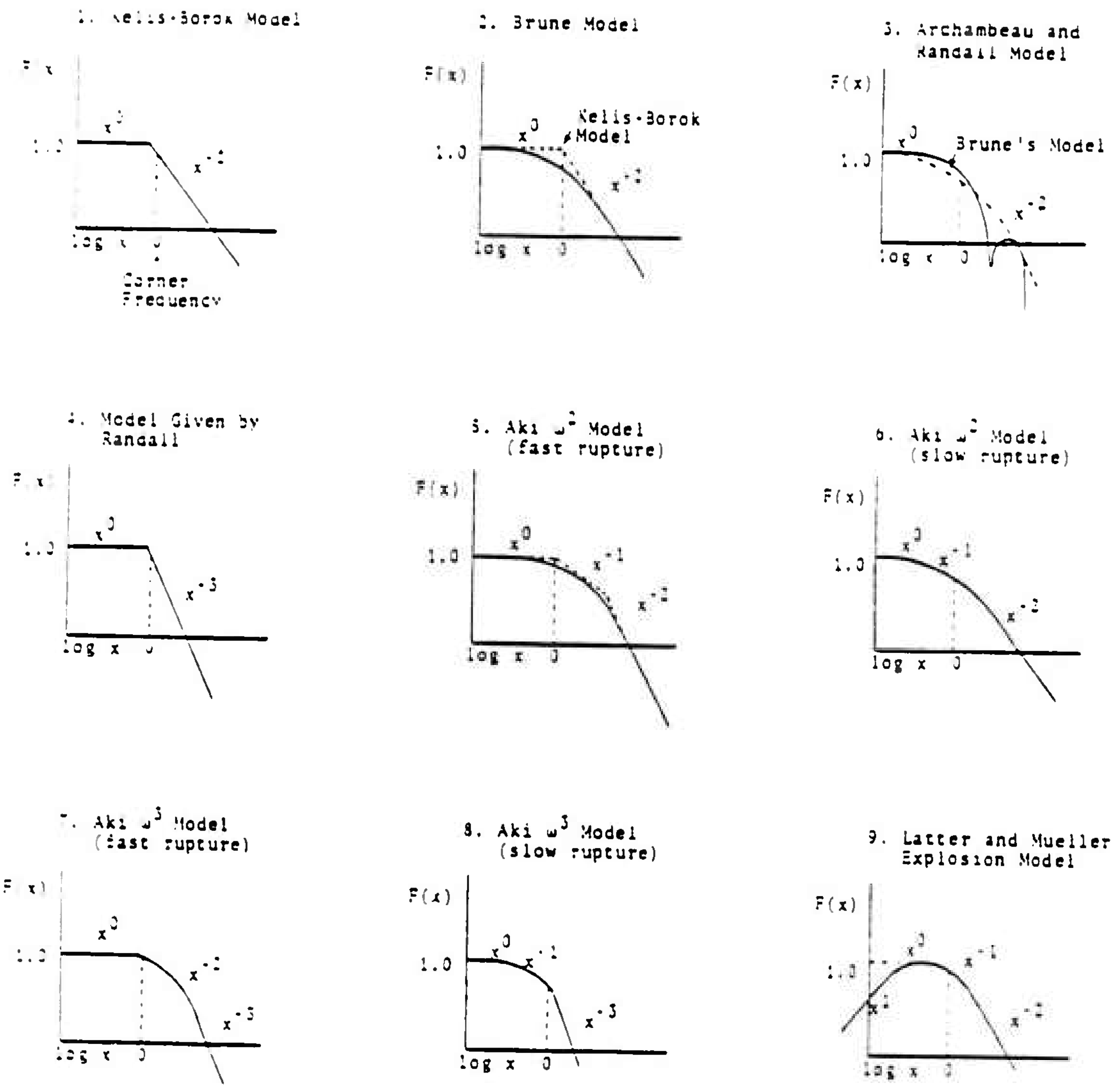


FIGURE III-4
ILLUSTRATION OF SEISMIC SOURCE MODELS

The set of k^{th} derivatives (-1,0,1,2,3) of the displacement ground motion, each reduced to measurements of magnitude and frequency, will result in ten multivariate statistics representing each event. To utilize these measurements, magnitude and frequency measurements of each derivative need to be calibrated as a function of event magnitude since the corner frequencies are magnitude dependent. Also, the calibration is needed to correct for path absorption and site characteristics.

A difficult problem is anticipated in separating the whitened signal bands from seismic noise; especially for the ground displacement (0^{th} derivative) and the integral of the ground displacement (-1^{th} serivative). As pointed out in the discussion of noise models in Section II, the existance of nearly periodic peaks mixed with nearly white Gaussian noise will tend to change the skewed statistical distribution of noise magnitudes to a more symmetrical normal distribution of much lower standard deviation. If the observed frequency of noise peaks of the k^{th} derivative trace are outside the anticipated signal frequency band, e.g., such as microseisms peaking at 0.2 Hz. Filtering to extract the nearly periodic noise component and applying amplitude and phase demodulation techniques to minimize the interference of microseisms, it should be possible to improve measurements of weak signals.

C. BASIC GROUND MOTION MEASUREMENTS

In Section II, we discussed errors associated with the use of the analytic technique of Unger and Farnbach; especially the problem of making precise estimates of the frequency of signals. In that case, the error increases without limit as the envelope approaches zero. More generally, if the envelope changes over the time interval of the frequency estimate, then the frequency estimate is biased or will lack precision.

For this reason, Unger's point-by-point technique of measuring signals as the logarithm of the envelope contained in a 4 second window is modified as follows.

- Seismic records are transformed into sequences of measureable envelope peaks
- Data associated with each envelope peak are
  - magnitude of ground motion
  - frequency
  - arrival time.

If a seismic peak is encountered, we will have timed the envelope minimum preceding the peak and that following the peak. By symmetry considerations, it can be shown from Rice (1954) that most probable occurrences of envelope minima of a Gaussian time series are approximately 0.5 of the RMS of the time series; of envelope maxima, 1.5 times the RMS.

To distinguish signal envelope peaks, the envelope level of the maxima must be 1.5 to 3 times the adjacent minimum values. After applying this criteria, the magnitude of the seismic signal peaks will be computed as the common logarithm plus a transmission B-factor to correct for propagation from a source of known location.

The problem of determining accurate frequency measurements is minimized by avoiding determinations at times where there are large negative magnitude fluctuations. The frequency measurements are optimized by determining the frequency at maxima of the envelope peaks. Figure III-5 shows an envelope peak occurring between the i^{th} and $(i+1)^{th}$ data point on a seismic record.

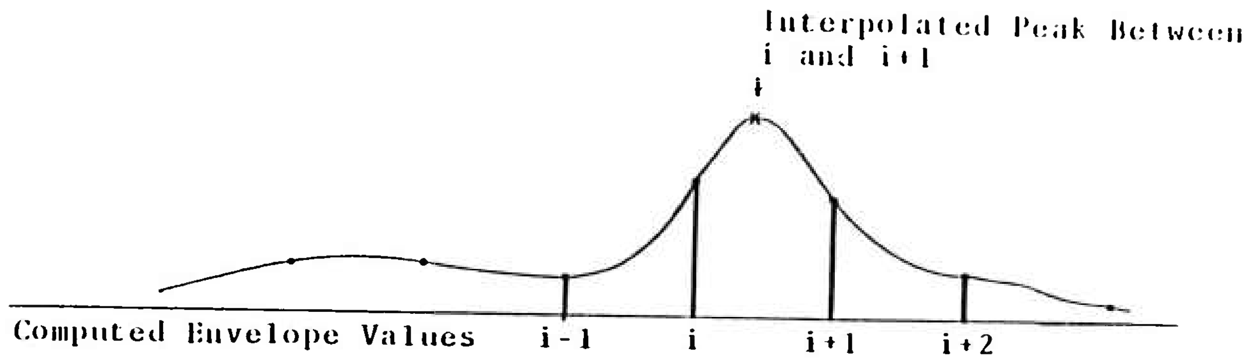
By interpolation, the envelope peak shown in Figure III-5 can be modeled by a cosine modulation.

$$E(t) = E_o \cos w_\ell t \ .$$

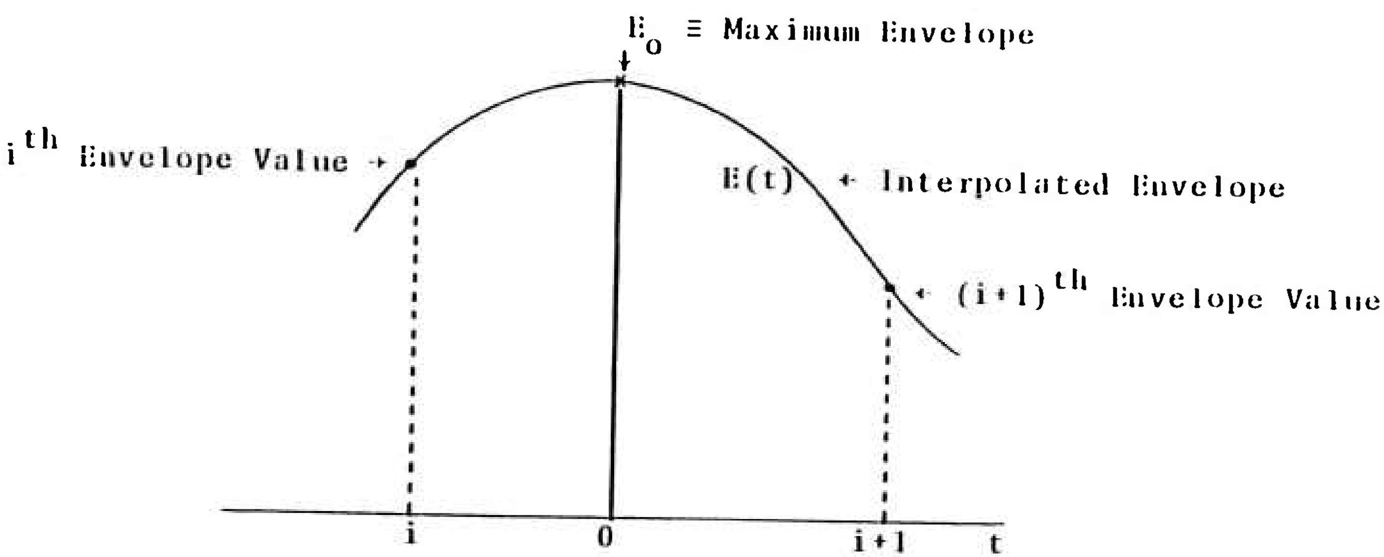
This adequately represents the time variation of the envelope in the immediate neighborhood of the interpolated maximum at point, x, shown in Figure III-5. Similarly by interpolation, the seismic data, x(t), is interpolated between points i and i+1;

$$X(t) = E(t) \sin (w_o t + \phi_o)$$

$$= E_o \cos w_\ell t \left[\cos \phi_o \sin w_o t + \sin \phi_o \cos w_o t\right]$$



a. Envelope Trace of Seismic Record



b. Interpolated Time Function Between i^{th} and $(i+1)^{th}$ Peak

FIGURE III-5

UNBIASED SIGNAL FREQUENCY ESTIMATION

$$= \tfrac{1}{2}E_o \{\cos\phi_o [\sin(w_o+w_\ell)t + \sin(w_o-w_\ell)t]$$

$$+ \sin\phi_o [\cos(w_o-w_\ell)t + \cos(w-w_\ell)t]\}.$$

We see that even under these ideal conditions of measuring frequency at the envelope peak, time varying phase angle estimates of the frequency could be seriously in error due to envelope modulations of angular frequency, w_ℓ. Nonetheless, at the interpolated maximum envelope peak an unbiased frequency estimate can be obtained by combining envelope and phase angle measurements with the time derivative of the seismic trace, as

$$x'(o) = \frac{dx(t)}{dt}\Big|_{t=0} = w_o E_o \cos\phi_o .$$

Since the Hilbert transform or quadrature filtered seismic trace, $y(t)$, is $x(t)$ phase shifted 90^o; the time derivative is

$$y'(o) = \frac{dy(t)}{dt}\Big|_{t=0} = -w_o E_o \sin\phi_o .$$

Thus, an unbiased estimate of the frequency is obtained,

$$f_o = \frac{\sqrt{x'(o)^2 + y'(o)^2}}{2\pi E_o} , \qquad \text{(III-4)}$$

at the interpolated maximum envelope peak. This requires time derivatives of the seismic trace and its Hilbert transform and measurements of the envelope at that point. By applying the preceding analysis, it can be shown that by modeling envelope modulations at inflection points as a constant envelope value E_o plus a sinusoidal modulation $\Delta E \sin w_\ell t$, the slope at such a point is given as

$$x'(o) = \left.\frac{dx}{dt}\right|_{t=0} = E_o w_o \cos\phi + \Delta E_o w_\ell \sin\phi_o \ .$$

Since $\Delta E_o w$ is the derivative of the envelope function at t=0, we see that computing frequencies at points other than the peak of the envelope function yields bias due to the slope of the envelope function. Thus, equation (III-4) should be used to compute frequency at interpolated envelope peaks to avoid bias due to envelope modulation.

The problem of phase modulation caused by random interference of other frequencies is another source of error. Near peaks and troughs of the random phase angle modulation, the correct frequency will be given by equation (III-4). Near zero crossings, maximum positive or negative deviations from the correct frequency are caused by phase modulations. Unfortunately, if we only measure frequency at envelope peaks, we are exposed to this source of error. But, if we initially estimate the frequency by equation (III-4), data conditioning can be carried out to minimize this source of error.

The time series containing the envelope peak, can be interpolated and filtered by a bandpass filter of sufficient bandwidth to pass the peak without distortion. This is given by Rice (1954) both for an ideal (square) bandpass filter and an equivalent bandwidth Gaussian bandpass filter. Such filtering about the initially estimated frequency will reduce phase modulation errors.

Further reduction of phase modulation errors is possible by averaging the frequency measurements over a time span covering that portion of the time series exceeding the most probable occurring envelope from some hypothetical Gaussian input giving rise to the observed envelope peak. This can be determined approximately and robustly by measuring the frequency at points about the envelope peak which are above the median envelope level. To avoid errors due to envelope modulation, the filtered data x(t) and the Hilbert transform y(t) are transformed to minimize the effect of envelope modulation.

$$u(t) = x(t)(E_o/E(t))$$

$$v(t) = y(t)(E_o/E(t)) \ .$$

The time series u(t) and v(t) are substituted for x(t) and y(t) in equation (III-4). Since envelope modulation effects are reduced by this process, the frequency can be measured at all points in the neighborhood of the envelope peak above the median of the envelope function defining the peak. By applying this procedure to measure the frequency associated with each observed peak envelope we expect to obtain more precise estimates of frequency than those previously obtained by Unger and others. Our frequency estimates will be more precise and minimally biased by minimizing errors caused by envelope and phase modulation.

By measuring seismic ground motion with the above technique, we expect to obtain accurate magnitudes corresponding to observed envelope peaks. In addition, estimates of the dominant frequency of ground motion will be associated with each observed envelope peak. The arrival time of each peak will be taken as the time of the first envelope minimum preceding the observed envelope peak.

The input to such a basic ground motion measure package would be the n^{th} derivative of ground displacement ($n=-1,0,1,2,3$); the output of a sequence of envelope peaks, each specified by

- Magnitude
- Frequency
- Arrival time.

Seismic records will be converted sequentially into five derivative ground motion time series. In turn, each of these derivative ground motion traces will be transformed into a corresponding set of the basic ground motion measurements described above.

D. INTERPRETATION OF BASIC GROUND MOTION MEASUREMENTS OF NOISE; DATA CONDITIONING

The first step in associating basic ground motion measurements with signals from a known event is to distinguish possible signals from magnitude peaks associated with seismic noise. Each record accessed after locating and timing a seismic event will contain approximately one or two minutes of seismic noise data starting about 2 or 3 minutes before the expected signal first arrival time and ending about 1 minute before expected signal arrival time. The purpose of applying a 1 minute guard gate is to avoid mixing signal data with that which is to be interpreted as noise. Since ordered statistics will be used to derive the noise parameters needed to detect signals, mixing of a small fraction of signal information with noise should not seriously affect the determination of noise parameters. The advantage of the ordered statistics approach is their robustness. By neglecting extreme large and small deviations effects produced by intermittent malfunctions such as spikes, signals mixed in noise, etc., will be minimized.

Following the Gaussian noise model described in Subsection A, the magnitude level bounding 20% of the smallest magnitudes, 66% and 88% are used to derive the most probable occurring noise magnitude, as well as the standard deviation of positive and negative magnitude deviations. Under the ideal Guassian assumption, the negative standard deviation is expected to be about 2.5 times larger than the positive deviation and the most probable envelope peak magnitude is expected to occur at 1.5 times the RMS of the noise. If, however, relatively broadband Gaussian noise is mixed with an extremely narrowband (almost periodic) component (Rice, 1954), the ratio of negative to positive standard deviation is close to one. Should this condition be detected, and should a narrowband of frequencies be indicated by the ordering of frequency statistics (e.g., 70% of the observed frequencies between the octave band 0.2 and 0.4 Hz), then the data must be conditioned to remove this noise source.

Data conditioning applied to demodulate useful seismic data from data containing an interfering nearly periodic noise component is illustrated by Figure III-6. The equations for separating a desired broadband process, $\Delta x(t)$, from the mixed data $x_\mu(t)$ are as follows.

By inspection of the illustration and the definition of variables in Figure III-6, we have the following equations for demodulation of the small broadband data component $\Delta x(t)$.

$$\delta(t) = \pi - \gamma(t) - \Theta_f(t) .$$

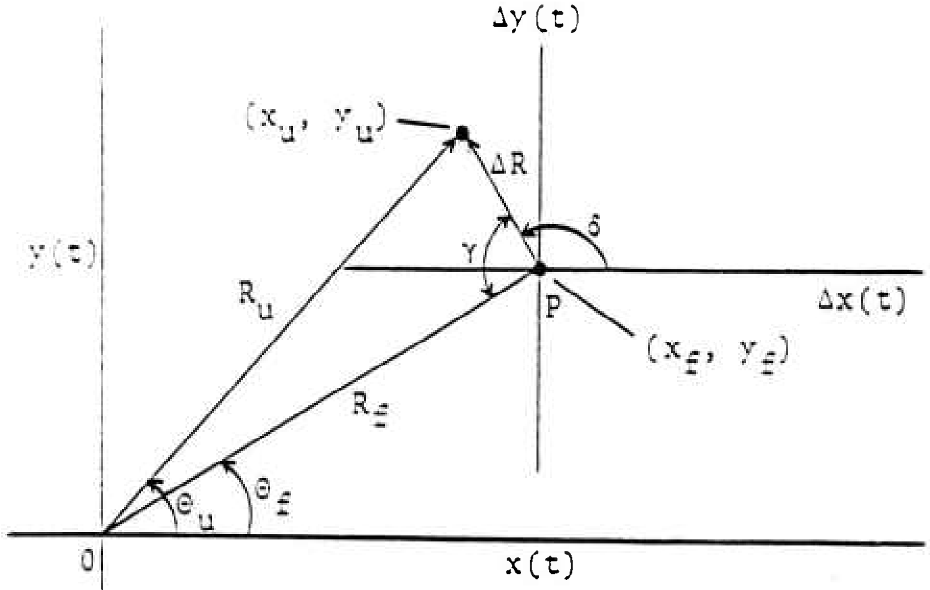
By law of cosines,

$$\Delta R(t) = \left[R_\mu(t) + R_f(t) - 2R_\mu(t)\ R_f(t)\ \mathrm{Cos}(\Delta\Theta) \right]^{\frac{1}{2}}$$

where

$$\Delta\Theta(t) = \Theta_\mu(t) - \Theta_f(t)$$

$$\cos\gamma(t) = \frac{\Delta R(t)^2 + R_f(t)^2 - R_\mu(t)^2}{2R_f(t)\ \Delta R(t)} .$$



<u>Complex seismic data as seen in x-y plane from 0</u>

x_u Unfiltered data (mixed process)

y_u Hilbert transformed unfiltered data

R_u Envelope of unfiltered data

Θ_u Phase angle of unfiltered data

Similarly (x_f, y_f, R_f, and Θ_f) are complex seismic data points from the filtered narrowband process

ΔR Envelope of demodulated data (effect of interfering narrowband component removed)

δ Phase angle of demodulated data

$\Delta x(t)$ Demodulated broadband data obtained by removing the strong interfering spectral component $x_f(t)$ from unfiltered data $x_u(t)$

FIGURE III-6

SEPARATION OF A BROADBAND PROCESS FROM THE EFFECT OF A STRONG INTERFERING NEARLY PERIODIC COMPONENT

The desired demodulated broadband data are given as

$$\Delta x(t) = \Delta R(t) \cos\phi(t) \ . \qquad \text{(III-5)}$$

Simply getting rid of the strong almost periodic component by linear filtering is often insufficient because the modulation of the small broadband noise component effectively broadens the bandwidth of the interfering narrowband process. This makes it difficult to see signals at frequencies near the interfering spectral peak. This can be seen intuitively by examining Figure III-6. The broadband data containing much higher frequencies moves around point P many cycles while the strong low frequency interference component moves point P very slowly. This has the effect of modulating the envelope R_f, between $R_f-\Delta R$ and $R_f+\Delta R$. It also modulates the phase ,Θ, between $\Theta_f \pm \tan^{-1}(\Delta R/R_f)$. These broaden the bandwidth of the spectral peak. A broader band filter is required to remove this periodic component.

Having thus conditioned the data, we are in a position of generating and interpreting noise statistics. At least as a reasonable approximation, the noise can be interpreted as a Gaussian process.

On that basis, we use ordered statistics to robustly determine the most probable occurring noise magnitudes. This is done separately for each of the n^{th} derivative seismic traces.

Although the distributions are skewed, we use a method of normalization which divides positive deviations by a positive standard deviation; negative deviations by a negative standard deviation. For the purpose of multivariate analysis, this ultimately reduces all of the signal detection variables to a homogeneous set of unit normal noise statistics.

For each n^{th} derivative ground displacement seismogram, we start with a set of three Basic Ground Motion measurements. These are the arrival time of an envelope peak, the magnitude, of the envelope maxima, and the dominant frequency.

From ordered statistical analysis of the noise data preceding the window containing signals, we generate statistics of magnitude and frequency fluctuations.

One of the main purposes of using a multivariate detector to time and measure signals is to improve our ability to automatically time and measure small signals. These presumably could be detected as statistically significant deviations of the frequency or magnitude from values expected of seismic noise. These would be expected to persist for a duration of 0.5 to 10 seconds (the expected duration of a seismic phase).

For that purpose, we separately determine relationships for positive and negative deviations from the most probable occurring noise magnitude. Let ΔM and f_i be the normalized magnitude fluctuation and corresponding frequency of the i^{th} observed noise envelope peak. Equation (III-6) shows the equations for determining a statistical noise prediction model for positive magnitude fluctuations.

$$\begin{array}{c} \text{n observations} \\ \text{of noise peaks} \end{array} \begin{bmatrix} 1 & 1 & 1 \\ \underline{1} & \underline{\Delta M^{+}} & \underline{\log f^{+}} \\ \downarrow & \downarrow & \downarrow \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} \\ \underline{m} \\ \\ \end{bmatrix} \qquad \text{(III-6)}$$

where $\underline{1}$ is a vector of one's; $\underline{M^{+}}$, positive magnitude deviations; and $\underline{f^{+}}$, the associated frequency of positive magnitude deviations.

The purpose of (III-6) is to derive a systematic relationship between positive noise magnitude deviations and their associated frequency. The relationship is used to predict the most probable magnitude of noise, $\underline{m}$, a constant vector. Its value is equal to the most probable noise magnitude derived from ordered statistical analysis of the observed noise magnitudes.

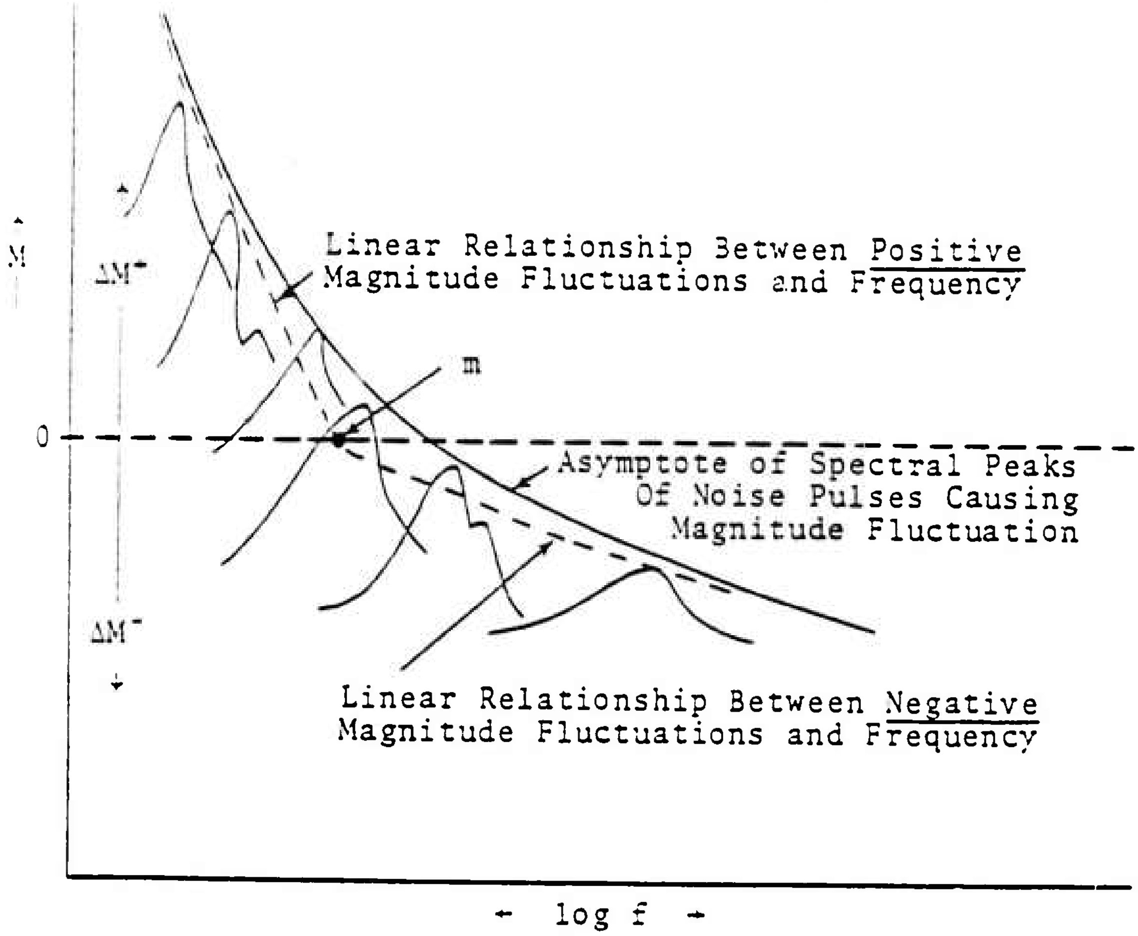
Similarly equation (III-7), determines a noise prediction error operator $\underline{b}$ for negative magnitude fluctuations which utilizes the frequency of observed noise peaks to minimize their deviation from the most probable observed noise.

$$\begin{bmatrix} 1 & 1 & 1 \\ \underline{1} & \underline{\Delta M^{-}} & \underline{\log f^{-}} \\ \downarrow & \downarrow & \downarrow \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 1 \\ \underline{m} \\ \downarrow \end{bmatrix} \qquad \text{(III-7)}$$

The physical rationale for this approach is seen in Figure III-7. Unger (1977) observed a linear relationship between dominant frequency and phase; or equivalently, the frequency standard deviation or bandwidth of seismic signals. This is shown schematically in the hypothetical relationship in Figure III-7. Note that negative magnitude fluctuations are postulated as broader bandwidth (higher variance) pulse spectra. The possible association of higher magnitude fluctuations as more coherent lower frequency noise pulses of equivalent seismic energy seems to be a plausible concept for the noise prediction error model of equations (III-6) and (III-7).

Given that the noise prediction error model works as expected, it will significantly reduce the variance of noise magnitude fluctuations. This results in more reliable extraction of signals of known seismic events. Even small magnitude fluctuations caused by weak signals can then be more reliably associated with known events, timed, and measured. The application of equations (III-6) and (III-7) is expected to result in sufficient reduction of the standard deviation in the magnitude of noise peaks, that small signal pulses greater than two standard deviations of the noise peaks will reliably be detected by Unger's algorithm discussed in Section II.

An important part of our signal extraction strategy is to use frequency measurements as a dependent variable for detecting weak signals which are as much as a half magnitude below the noise level. We do this by applying ordered statistics to measurements of the frequency of noise peaks.



m = Most frequently observed magnitude of noise

M = Noise magnitude

ΔM^{+} = M-m if Sgn (M-m) $\geq$ 0

ΔM^{-} = M-m if Sgn (M-m) < 0

log f = Log of dominant frequency of envelope peaks

FIGURE III-7

HYPOTHETICAL MODEL SHOWING MAGNITUDE FLUCTUATIONS OF NOISE: HIGHER FREQUENCY ENVELOPE PEAKS ASSOCIATED WITH BROADER BANDWIDTH AND LOWER AMPLITUDE; LOWER, WITH NARROW BANDWIDTH AND HIGHER AMPLITUDE

For that purpose, we separately determine relationships for positive and negative deviations of the frequency of noise peaks. Since we do not have a statistical model of frequency fluctuations to draw on, we assume that the distribution is skewed. The median frequency, F, is taken as the most probable occurring frequency of noise envelope peaks. Positive and negative standard deviations are then at the 85th and 15th percentile of observed frequencies of noise peaks ordered from lowest to highest frequency. Let Δf_i be the observed frequency fluctuation and M_i the associated magnitude. Following the methodology resulting in equation (III-6), we have the corresponding equations for observed positive frequency fluctuations.

$$\begin{bmatrix} 1 & 1 & 1 \\ \underline{1} & \underline{\Delta f^+} & \underline{M^+} \\ \downarrow & \downarrow & \downarrow \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} \\ F \\ \\ \end{bmatrix} \qquad \text{(III-8)}$$

Following the methodology of equation (III-7), we have a corresponding equation for observed negative frequency fluctuations.

$$\begin{bmatrix} 1 & 1 & 1 \\ \underline{1} & \underline{\Delta f^-} & \underline{M^-} \\ \downarrow & \downarrow & \downarrow \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} \\ F \\ \\ \end{bmatrix} \qquad \text{(III-9)}$$

The purpose of (III-8) and (III-9) is to derive a systematic relationship between frequency fluctuations and the magnitude measurement associated with each observed frequency fluctuation. The physical motivation was shown in Figure III-7. This procedure is expected to minimize the variance of frequency fluctuations of noise peaks. This technique makes it possible to detect weaker signals.

For some signals we expect that it will be preferable to time, extract, and measure signals based on significant magnitude fluctuations. In other cases, we expect it will be perferable to use frequency measurements as the dependent variable for this purpose. Unger (1978) showed that phase angle criteria is at least 6 dB more sensitive in detecting weak signals than is amplitude. Since frequency is a phase modulation measurement, it should be a more sensitive criteria for detecting those signals of significantly different frequency than is noise. For example, it is commonly observed that regional Lg phases can be most easily recognized by abrupt frequency changes. For these reasons, we will apply both the magnitude and the frequency criteria to time, extract, and measure weak seismic signals.

The operators a, b, c, and d, in equations (III-6) through (III-9) are derived from the analysis of noise data provided at the front end of each record. These are optimum operators for removing the effects of frequency fluctuations from magnitude measurements and magnitude fluctuations from frequency measurements. Positive and negative fluctuations are separately treated. Letting the matrix operator G or H, represent noise observations and the vectors g or h, one of the desired operators for removing the effect of fluctuations; we can determine a least squares solution as follows. For estimating magnitudes,

$$\underline{g} = (G^T G)^{-1} G^T \underline{m} \quad , \tag{III-10}$$

for estimating frequency,

$$\underline{h} = (H^T H)^{-1} H^T \underline{F} \quad . \tag{III-11}$$

As was done for scaling discriminants in the Event Identification Experiment, more robust determination of $\underline{g}$ and $\underline{h}$ can be obtained from small samples subject to occasional large errors. This is done simply by altering the rules for vector dot multiplication by replacing $\underline{x} \cdot \underline{y}$ with N times the median of $\{x_i y_i\}$, where $i=1,2,\ldots,N$. Since matrix multiplication involves a set of vector dot products, robust determinations involve a set of sort operations for median determinations. It is noted that if statistical deviations are normally distributed and N is very large, robust determinations of the operators are equivalent to least squares estimates.

The purpose of applying the operators to measurements of envelope peaks is to obtain minimum noise variance estimates of magnitude, m, and frequency, F. These, obtained from each of the n^{th} derivative ground motion measurements, are designated m_n, F_n. The weak signal extraction problem gets down to distinguishing the ten element multivariate vector $(m_{-1}, F_{-1}, m_o, F_o, \ldots, m_3, F_3)$. It is perhaps worth noting that m_3 versus F_3 has been used as an effective discriminant (referred to as the third moment method) between earthquakes and explosions. It is not unreasonable to expect this multivariate signal extraction process to distinguish various types of events as well as distinguish signals from noise.

E. MEASUREMENTS ON ENVELOPE PEAKS; POST-FILTERING

1. Statistically Independent Estimates

In Subsection D, we derived correction operators g and h for obtaining m and F. m is the magnitude fluctuation associated with an envelope peak corrected for its apparent frequency dependence; F the frequency fluctuation corrected for its magnitude dependence. The corrections were designed to minimize the variability of noise fluctuations. The purpose of this correction was to maximize the sensitivity for automatically detecting, timing, and measuring weak signals occurring in some unknown position of the accessed time windows of a known event. Any statistically significant deviations of m and F from values expected for noise can then be interpreted as possible onsets of seismic signals.

The first step in analyzing noise measurements of m and F is to normalize m and F to unit normal variates by application of ordered statistics. By subtracting the most probable occurrence from each variate and dividing by the appropriate positive or negative standard deviation, m and F are reduced to approximate unit normal statistical variates Z(m) and Z(F). The observations of Z(m) and Z(F) are operated upon by rotational transformation operator, Q, which transforms them to

Z'(m) and Z'(F) which are statistically independent. This operation is shown by equation (III-12).

$$\begin{array}{c c}\text{Observation Matrix from Noise Sample} & \begin{bmatrix} 1 & 1 \\ \underline{Z}(m) & \underline{Z}(F) \\ \downarrow & \downarrow \end{bmatrix}\end{array} \begin{bmatrix} Q \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ Z'(m) & Z'(F) \\ \downarrow & \downarrow \end{bmatrix} \qquad \text{(III-12)}$$

where Q is a 2x2 matrix. By taking the transpose of the observation matrix and applying it to both sides of equation (III-12) we obtain a least squares determination of Q .

$$RQ = \begin{bmatrix} \mathrm{var}\{Z(m)\} & \mathrm{cov}\{Z(m), Z(F)\} \\ \mathrm{cov}\{Z(m), Z(F)\} & \mathrm{var}\{Z(F)\} \end{bmatrix} \begin{bmatrix} Q \end{bmatrix} = \begin{bmatrix} 1.0 & 0 \\ 0 & 1.0 \end{bmatrix} \qquad \text{(III-13)}$$

The solution for Q is simply obtained as the inverse of the covariance matrix, R, of magnitude and frequency fluctuations. Q, by definition, and when applied to observations Z(m) and Z(F), produces Z'(m) and Z'(F) which are statistically independent variables for detecting signals by significant magnitude or frequency fluctuations from noise. Since m and F have been pre-conditioned to remove any obvious dependencies, it is expected that R^{-1} will be a very stable operator. The rotation operator Q is applied to remove any apparent dependency between m and F so that the assumption of independent Gaussian statistics is valid.

As desired, we now have formulated two independent statistics for detection signals. $Z'(m)$ is an approximate unit normal statistic for detecting signals by their unusually large magnitude fluctuations from noise. $Z'(F)$ detects small signals based on unusually large changes in frequency from the most probable frequency occurrence of noise.

2. Bandwidth of Envelope Peaks and Post-Filtering of Signals

Our rationale in separately searching for signals in n^{th} derivative ground motion traces of a record containing signals from a known event was illustrated schematically by Figure III-4. The basic idea was to whiten signals between pairs of corner frequencies governing the spectrum of a signal. For example, a Brune model source can be considered to be a low bandpass white source of ground displacement (0^{th} derivative) up to near the corner frequency and a high bandpass white source of ground acceleration (2^{nd} derivative) above the corner frequency. Suppose such a source was measured with seismic data which are band limited between f_1 and f_2. Then the bandwidth of ground displacement would be expected to be Δf (depth) = $f_c - f_1$, where f_c is the corner frequency; of ground acceleration, $\Delta f(accel) = f_2 - f_c$. By separately analyzing all five motion derivatives between -1 and 3 we expect to obtain physically meaningful band-limited white representation of the source in different portions of the spectral band. By measuring one or more detected envelope peaks; their magnitude, frequency and bandwidth; we expect to obtain useful source-related information.

Rice (1954) analyzed the problem of determining the rate at which envelope peaks are passed by a band limited white Guassian input. If the average time between envelope peaks of noise is $\Delta\tau$ seconds, then the effective bandwidth of the noise process, Δf, is derived by Rice as

$$\Delta f \Delta \tau = 1.56 \ . \tag{III-14}$$

Thus, if we observe noise envelope peaks occurring on an average interval of $\Delta\tau_n$ seconds, then the effective bandwidth of the noise process can be approximated as $1.56/\Delta\tau_n$. Since Unger (1978) demonstrated that seismic noise envelope measurements can be reasonably modeled as a Gaussian process, this should be a reasonable assumption for samples of noise of long duration.

A more interesting parameter, which can be applied to observations of individual envelope peaks is the relationship between the measured pulse width of envelope peaks. Our basic envelope peak measurements include the start time as the first significant minimum preceding the envelope peak and the end time as the first significant minimum following the envelope peak. Defining the measured pulse width of noise, T_n, as the average pulse width of noise envelope peaks, then we can determine a constant, a, which relates measured pulse width to $\Delta\tau_n$ as a T_n. This yields a relationship between effective bandwidth and measured pulse width as

$$\Delta T = a\Delta\tau \ . \tag{III-15}$$

The purpose of measuring the effective bandwidth of an envelope peak or sequence of envelope peaks is to post-filter weak signals. The filter will enhance the signal-to-noise ratio of weak signals as part of the signal extraction process and to extend and to enhance the quality of subsequent signal measurements. Figure III-8 shows examples of post-filters applied to extracting weak signals. Δf is the full bandwidth of a filter designed to pass weak signals. For the trapezoidal filter, Δf, is obtained from equation (III-14). For the Gaussian filter, f, it was derived by Rice (1954) and is shown by equation (III-16).

$$\Delta f \Delta \tau = 0.67 \quad . \tag{III-16}$$

Rice's derivation was on the basis of the bandwidth passing the same average amount of power as an ideal bandpass filter.

The approach toward applying post-filtering is to initially set a high false alarm rate threshold to detect weak signals. If such weak signals are detected, post-filtering will be applied. This will be followed by applying a more exacting threshold permitting acceptable network operating characteristics for extracting signals.

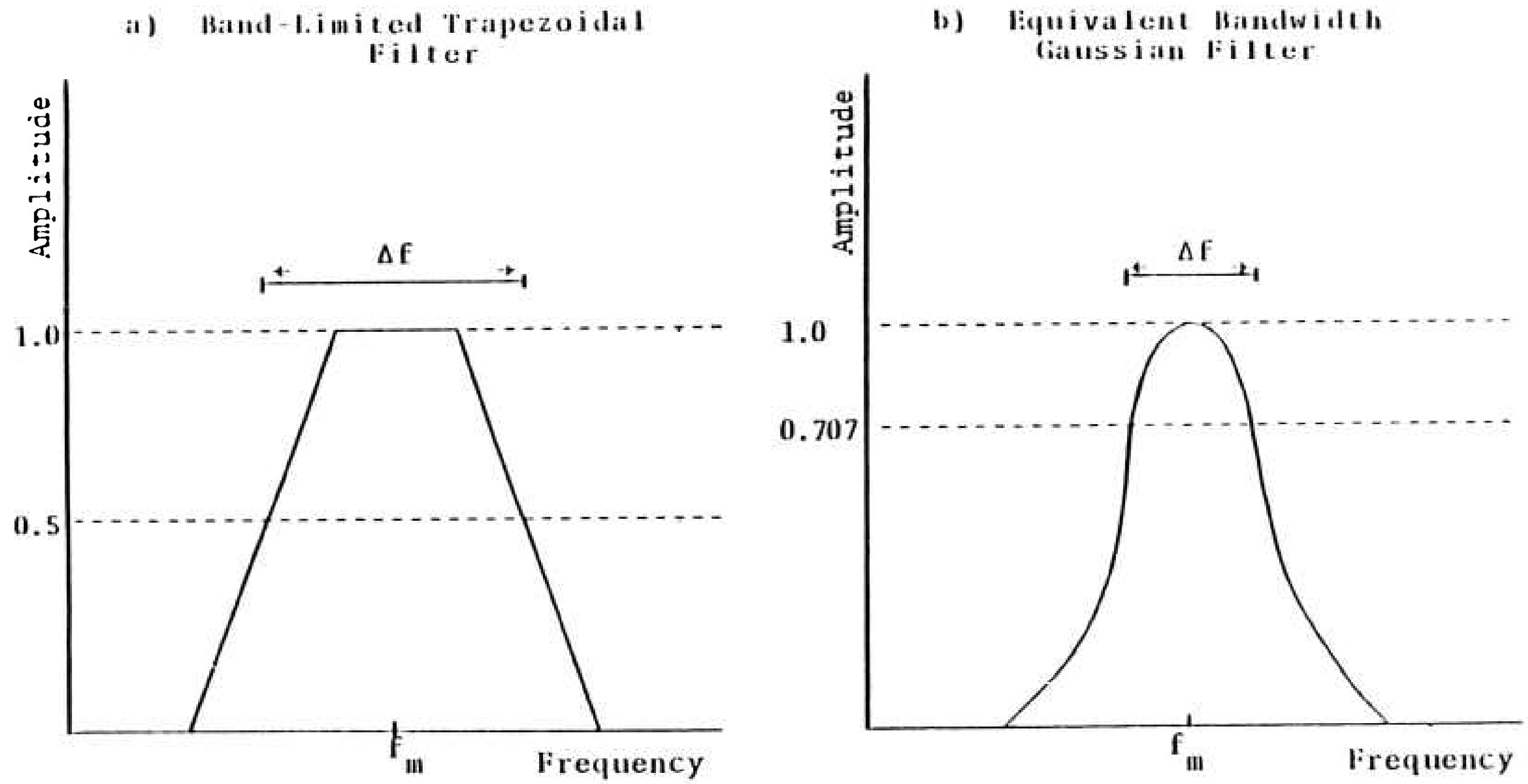


FIGURE III-8

EXAMPLES OF POST-FILTERS FOR EXTRACTING WEAK SIGNALS

F. DETECTING COMPLEX SIGNALS

1. Random Signal Characteristics

Due to strength heterogeneity of the source region and other complicating factors, the signal does not always propagate as a single envelope peak. Such a source can be more generally characterized as a randomly distributed source which can be modeled as a cluster of delayed transmissions over the signal duration, D. Over a longer time scale the source density decays and the signal gradually fades out.

Observations by Aki (1969) of long-time terms of coda well below the signal level indicate that such coda can be explained by source and receiver back-scattering at frequencies dependent only on travel-time from the source. This effect can be viewed as an illumination of the medium near the earth's surface, all along the path from the seismic source, and due to multiple scattering of waves observed well after the signal arrival and at much lower levels than the signal. The time-frequency scattering function observed near the source by Aki indicates that the effect fades rapidly due to spreading, dispersion, and absorption. There he observed a fading amplitude with the dominant period shifting rapidly to longer periods with time elapsed from the event origin time. The same dominant periods are seen at a simultaneous time delay from the origin time of the source, at different receiver sites, at varying distance from the source.

A completely different coda model is relevant to the detection of signals. It stems from the concept of a seismic source distributed randomly in time and space. Here our concern is with determining a B-factor correction applicable to fading coda measurements. We view the signal as a uniformly distributed random source modeled as shot noise over the signal duration, D. This will be characterized as a cluster of envelope peaks of approximately the same magnitude and of similar spectral characteristics. For times greater than D, we continue to see delayed replicas of the signal, the only difference being that the magnitude decreases after the signal duration, D, has elapsed and then fades gradually with time. This differs from Aki's model in that it pertains to short-term time delays at initial magnitudes close to that of the signal. This particular model was used to simulate earthquakes by Shoup and Sax (1974). It realistically predicted the time-span of visible signals over a wide range of magnitudes. The fading shot noise coda model B-factor correction is given by equation (III-17).

$$\Delta B_c(\tau) = a \log (\tau/D) + c, \quad (\tau>d) \tag{III-17}$$

where τ is the elapsed time from the start of the signal. This is in contrast to Aki's model, where he computes decay with time as the elapsed time from the origin time of the event. This model is considered valid only for short-term decay of signals where Aki's model is valid for the long-term decay occurring well after the arrival of the signal. The initial evidence for the coda B-factor correction of equation (III-17) was obtained from an attempt to calculate consistent magnitude estimates from measurements of frequency-wave number peaks.

An example is shown on Table III-1, where data were extracted from measurements of spectral wave number peaks detected by the program FKCOMB by Smart (1972). There, the peaks were measured with varying time windows of 3.2, 6.4, and 12.8 seconds. The data were treated as a stationary random process with each spectral component averaged over the respective time window. The effect observed was consistent fading of the average amplitude as progressively longer time-windows were analyzed. Note on the bottom of Table III-1 that the B-factor increases by 0.1 for each doubling of the time window used to sample the earthquake.

Consistent with observing a number of events, the relationship of equation (III-17) is a simple coda decay model reflecting these results. The parameters a and c in equation (III-17) were estimated at a = 0.33 and c = 0.15. This implies a model where the coda amplitude, $A_c(t)$, drops to about 0.7 of the signal amplitude after duration, D, and decays with the dimensionless factor (τ/D) raised to the -0.33 power. This is shown by equation (III-18).

$$A_c(t) = 0.7\ (\tau/D)^{-0.33}, \qquad (\tau > D). \qquad \text{(III-18)}$$

This decay model was observed to be consistent with modified B-factors computed empirically for determining magnitudes from spectral wave number peaks. It is considered applicable for the decay immediately following a signal of duration, D.

TABLE III-1

TIME WINDOW EVALUATION

P-Wave, Back Azim = 114.5, Δ = 44.8°, f-K Peaks

Start 163/1926 56.4, End 163/1927 10.3, 1971

NOA M_B 5.0

LASA M_B 5.1

Sub-Array F-4, LASA

PERIOD (SEC)

B.AZIM.	3.2	2.3	1.8	1.5	1.3	1.1	0.9	0.8	0.7	0.6
3.2	111	---	---	110	---	108	110	116	---	---
6.4	98	---	---	111	114	118	112	114	116	110
12.8	---	115	101	101	115	---	113	112	---	111
dt/dΔ										
3.2	8.43	---	---	9.20	---	6.40	7.04	7.66	---	---
6.4	7.46	---	---	7.50	7.33	10.6	7.25	7.40	6.23	6.40
12.8	---	5.14	7.46	6.92	7.34	---	7.60	7.14	---	6.12
F-Stat.										
3.2	84	---	---	294	---	157	194	96	---	---
6.4	77	---	---	249	189	88	160	87	194	242
12.8	---	105	225	230	193	---	108	99	---	152
MAG'										
3.2	5.1	---	---	5.2	---	4.9	5.2	5.2	---	---
6.4	5.0	---	---	5.1	5.1	4.7	5.2	5.1	4.8	4.9
12.8	---	5.0	5.0	5.0	4.9	---	5.2	5.2	---	4.8

TIME WINDOW (SEC)

$$MAG' = \text{Log}\,(P^{1/2}/T) + \begin{cases} 2.8 & \text{(3.2 Sec. Time Window)} \\ 2.9 & \text{(6.4 Sec. Time Window)} \\ 3.0 & \text{(12.8 Sec. Time Window)} \end{cases}$$

It was further observed with FKCOMB that spectral wavenumbers in the coda following a signal peak indicated frequency peaks, azimuths, and dT/dΔs consistent with those observed at the initial signal peak. This provides some support for the shot noise model to describe the initial strong motion characteristic of earthquake coda.

C. Onset Time and Duration of a Propagated Phase

Part of the random characteristics of signals is taken care of by the assumption of the sudden onset of clustered signal peaks persisting for duration D, followed by a decaying coda. For more completeness we need to also consider the emergence time of a signal onset. Since we are detecting signals on segmented records, there is no need to continuously update the noise statistics. Therefore, there is no problem of missing emergent events; only a problem of accurately timing the beginning of such events. A criteria for the model of an emergent signal is a sequence of monotonically increasing envelope peaks. Also, a model for the end of a signal duration D is a sequence of monotonically decreasing signal peaks. Thus, at the onset of a possible signal, the start of the time window is held stationary as monotonically ascending peaks are sensed. If the maximum peak of such a sequence passes signal detection threshold criteria, then the starting point of the emergent signal is correctly ascertained. Conversely, the end of the signal window is similarly determined by similarly operating on a monotonically decreasing sequence of envelope peaks. If the minimum envelope peak of such a sequence passes signal rejection criteria, then the end of the signal window is accurately ascertained. In short, by treating monotonic envelope sequences as single maximum or minimum envelope peaks, signal windows, including those of emergent events, should be accurately timed.

3. Measurement of Propagated Phases

After detecting a propagated phase as a cluster of envelope peaks, we determine a time window of onset time, τ, and duration D. The detection status and the magnitude, frequency and bandwidth of each spectral moment is determined for each detected signal. Network processing consists of retrieval of this information for all detected phases on the records associated with the preliminary event location under consideration. Absorption corrections can be performed at some of those stations where signals are detected at anomalously low frequencies. Unless these absorption corrections are applied and are sufficient to obtain frequency band convergence, the use of uncorrected data could reduce the earthquake/explosion discrimination power of the data. Next, time anomaly and magnitude consistency criteria can be applied to refine the process of associating propagating phases and updating the location of the event. Finally, the measurements of signals and noise can be analyzed to determine unbiased estimates of the magnitude of the event and event discriminants.

G. NETWORK STRATEGY FOR RETRIVEING SIGNALS OF SMALL EVENTS

We expect to be faced with the problem of identifying small events which are masked by seismic noise. By applying fixed constant false alarm rate thresholds to constrain false alarms, we may keep such false alarms to an acceptably low level, but at the same time make it nearly impossible to extract weak signals. In the case of low magnitude signals, we can employ a station Variable False Alarm Rate (VFAR) strategy which raises the expected ratio of retrieved signals to retrieved false alarms to an acceptable level. On the presumption that an identification is required for all events analyzed, it is better to have at least one or a few signals combined with some false alarms than to have only noise measurements to identify the signal. This strategy is made somewhat more sensible by the fact that network validation procedures will be able to eliminate at least some of the false alarms and that network magnitude determinations will compensate to some extent for noisy determinations of the event parameters.

This type of post-automatic association network level signal analysis effectively reduces independent signal measurements to more accurate event measurements. It provides a network view to check the consistency of associated signals. Signals with reasonable identifiable source and propagation characteristics can be passed on while other 'false' signal measurements can be eliminated. Array and polarization measurements as well as travel times can be used in this context to identify and weed out obvious local events and to identify core phases outside the normal teleseismic range, later teleseismic phases, and regional phases.

In this phase of applying network strategies, we define our objective as that of detecting small signals; possibly even below the most frequently occurring noise level or by applying multivariate waveform recognition procedures. It is expected that the normally high automatic detector false alarm rate will be reduced by such multivariate analysis of these small signals. Furthermore, network strategies applied to these measurements such as measurement of travel time anomalies and magnitude deviations will further reduce such false alarms to a level which will allow accurate phase association and location. As previously pointed out, VFAR threshold strategies can further reduce false alarms and improve network multivariate magnitude measurements for event discrimination.

SECTION IV
CONCLUSIONS AND RECOMMENDATIONS

A concise description was given for the design of a multivariate analytic detector to be used as an editor of short-period seismic signals. The output of such an automatic seismic signal editor could provide additional data to improve the reliability and efficiency of event identification.

One of the functions performed by the automatic signal editor is to extract all of the waveforms which can be possibly interpreted as signals from the event. These extractions include long-period surface waves, short-period teleseismic P waves, secondary phases, and regional phases. This information could be automatically inserted into signal measurement files and accessed for additional interactive seismic processing by seismic analysts.

Another function of the automatic signal editor is to reduce basic ground motion measurements to multivariate source discriminants such as event ground motion or spectral moment magnitudes, including their dominant frequency, bandwidth and complexity. These signal measurements can be clustered to identify anomalous events and to calibrate normal earthquakes from various source regions. These spectral moment magnitudes can be interpreted physically, after appropriate attenuation corrections, as conventional magnitude measurements (log A/(Dominant Period)) of ground motion; i.e., displacement potential, displacement, velocity, acceleration and jerk. These are band-limited whitened source data in appropriate roll-off portions of the signal spectrum. Their use avoids many problems, including source size-scaling associated with the arbitrary application of narrow bandpass filters to compute source discriminants.

SECTION V
REFERENCES

Aki, K., Analysis of the Seismic Coda of Local Earthquakes as Scattered Waves, J. Geophys Res., 74, 615-631, 1969.

Aki, K., The Scaling Law of the Earthquake Source Time Function, Geophys. J. D. Astr. Loc, 31, 3-25, 1972.

Archambeau, C. B., General Theory of Elastodynamic Source Fields, Rev. Geophys., 6, 241-288, 1968.

Bracewell, R.N., The Fourier Transform and Its Applications, McGraw-Hill, New York, 1965.

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

Dugundji, J., Envelopes and Pre-Envelopes of Real Waveforms, Trans. IRE, Vol. IT-4, 53-57, 1958.

Farnback, J.S., The Complex Envelope in Seismic Signal Analysis, Bull Seism. Soc. Am., 65, 951-962, 1975.

Keilis-Borok, V.I., On Estimation of the Displacement in an Earthquake Source and Source Dimensions, Ann. Geofis. (Rome), 12, 205-214, 1959.

Latter, A. L., E. A. Martinelli, and E. Teller, Seismic Scaling Law for Underground Explosions, Phys. Fluids, 2, 280-282, 1959.

Mueller, R. A., Seismic Energy Efficiency of Underground Nuclear Detonations, Bull Seis. Soc. Am., 59, 2311-2323, 1969.

Randall, M. J., Seismic Radiation From a Sudden Phase Transistion, J. Geophys Res., 71, 5297-5302, 1966.

Randall, M. J., The Spectral Theory of Seismic Sources, Bull Seis. Soc. Am., 63, 1133-1144, 1973.

Rice, S. O., Mathematical Analysis of Random Noise, in Selected Papers in Noise and Stochiastic Processes, N. Wax, ed., Dover, Pub., New York, 1954.

Sax, R. L., A. G. Bell, and D. L. Dietz, Event Identification Experiment: Priority II Data Set, SAR (01)-TR-79-06, ENSCO, Inc., Alexandria, VA, 1979.

Shoup, E. M. and R. L. Sax, Simulation of a World-Wide Seismic Surveillance Network, ALEX (01)-TR-74-13, Texas Instrument, Inc., Alexandria, VA, 1974.

Smart, E., FKCOMB, A Fast General-Purpose Array Processor, SAAC Report No. 9, Teledyne Geoteck, Alexandria, VA, 1972.

Swindell, L. and R. Snell, Station Processor Automatic Signal Detection System, Phase I, Final Report, AFFAC PA/6103/B/ETR, Texas Instruments, Inc. 1977.

Strauss, A. C. and L. C. Weltman, Continuation of the Seismic Research Observatories Evaluation, ALEX(01)-TR-77-02, Texas Instruments, Inc., Alexandria, VA, 1977.

Unger, R., Automatic Detection, Timing, and Prelininary Discrimination of Seismic Signals with the Instantaneous Amplitude, Phase, and Frequency, ALEX (01) - TR-77-04, Texas Instruments, Alexandria, VA, 1978.

Weltman, L. C., H. Hsiao, and R. Oliver, An Evaluation of the Seismic Research Observations: Final Report, SAR(01) - TR-79-04, ENSCO, Inc., Alexandria, VA, 1979.