LEVEL II

SDAC-TR-77-14

A THREE-COMPONENT, SINGLE-STATION, MAXIMUM-LIKELIHOOD SURFACE WAVE PROCESSOR

EUGENE SMART
Seismic Data Analysis Center
Teledyne Geotech, 312 Montgomery Street, Alexandria Virginia 22314

21 JULY 1978

Sponsored by
The Defense Advanced Research Projects Agency (DARPA)
ARPA Order No. 2551

Monitored By
AFTAC/VSC
312 Montgomery Street, Alexandria, Virginia 22314

Disclaimer: Neither the Defense Advanced Research Projects Agency nor the Air Force Technical Applications Center will be responsible for information contained herein which has been supplied by other organizations or contractors, and this document is subject to later revision as may be necessary. The views and conclusions presented are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency, the Air Force Technical Applications Center, or the US Government.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: SDAC-TR-77-14	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): A THREE-COMPONENT, SINGLE-STATION, MAXIMUM-LIKELIHOOD SURFACE WAVE PROCESSOR	5. TYPE OF REPORT & PERIOD COVERED
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Eugene Smart	8. CONTRACT OR GRANT NUMBER(s): F08606-77-C-0014, ARPA Order-2551
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Teledyne Geotech, 314 Montgomery Street, Alexandria, Virginia 22314	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61p.
11. CONTROLLING OFFICE NAME AND ADDRESS: Defense Advanced Research Projects Agency, Nuclear Monitoring Research Office, 1400 Wilson Blvd., Arlington, Virginia 22209	12. REPORT DATE: 21 July 1978
	13. NUMBER OF PAGES: 62
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): VELA Seismological Center, 312 Montgomery Street, Alexandria, Virginia 22314	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

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

18. SUPPLEMENTARY NOTES

Author's Report Date 12/20/77

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

Particle Motion
Maximum Likelihood
Surface Waves

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

An entirely analytic single-station surface-wave processor is developed that analyzes long-period seismic records two orders of magnitude faster than real time on an IBM 360/44 computer. In processing four days of continuous synthetic data the algorithm detected 85% of the 170 signals of S/N = ½, with false alarm rate of one per day. A novel detection theory is introduced that exploits the consistency of the azimuthal estimates associated with a sequence of detections.

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

Unclassified 408 258

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

A THREE-COMPONENT, SINGLE-STATION, MAXIMUM-LIKELIHOOD SURFACE WAVE PROCESSOR

SEISMIC DATA ANALYSIS CENTER REPORT NO.: SDAC-TR-77-14

AFTAC Project Authorization No.: VELA T/7709/B/ETR

Project Title: Seismic Data Analysis Center

ARPA Order No.: 2551

Name of Contractor: TELEDYNE GEOTECH

Contract No.: F08606-77-C-0014

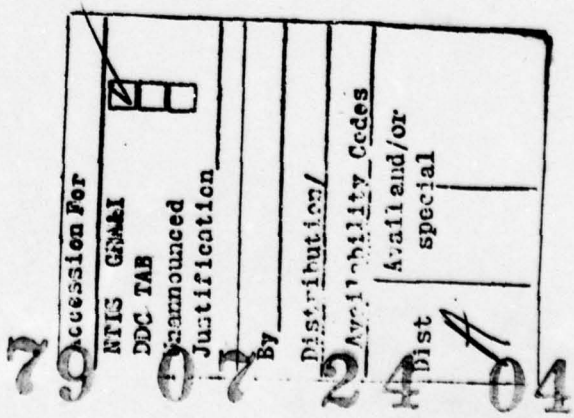
Date of Contract: 01 October 1976

Amount of Contract: $3,005,183

Contract Expiration Date: 30 September 1977

Project Manager: Robert R. Blandford
(703) 836-3882

P. O. Box 334, Alexandria, Virginia 22314



ABSTRACT

An entirely analytic single-station surface-wave processor is developed that analyzes long-period seismic records two orders of magnitude faster than real time on an IBM 360/44.computer. In processing four days of continuous synthetic data the algorithm detected 85% of the 170 signals of S/N = ½, with false alarm rate of one per day. A novel detection theory is introduced that exploits the consistency of the azimuthal estimates associated with a sequence of detections.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES (Con't)

INTRODUCTION

Though a number of techniques for processing surface-waves have appeared, they all lack comprehensive capabilities. Among these capabilities are adaptive processing (Barnard and O'Brien, 1974), match filtering (Capon, et al., 1969), and bandpass filtering and polarization filtering (Simons, 1968; Choy and McCamy, 1973; Sobel and von Seggern, 1975). The deficiencies of these techniques has left a need for a fast, adaptable, general purpose seismic surface-wave processor to serve as detector, signal estimator and azimuth estimator for routine searching and analyzing of long-period, single station three-component data.

This report presents a processor designed to meet that need. It incorporates the most salient qualities of the other techniques. The essential features are that (von Seggern, 1977):

a) no assumption need to be made about the arrival azimuth.

b) the process is entirely linear, which well preserves amplitude information and at the same time ensures stability.

c) it is very fast since it is entirely analytic, running 100 times real time on the IBM 360/44.

d) a statistical confidence measure of signal presence is produced.

e) false alarm rates can be estimated from noise properties.

In addition, this processor is a more sensitive detector, though retaining robustness because it is linear.

Barnard, T. E. and L. J. O'Brien, 1974. An evaluation of adaptive-beamforming techniques applied to seismic data, ALEX (01)-TR-74-08, Texas Instruments, Inc., Dallas, Texas.

Capon, J., R. J. Greenfield, and R. T. Lacoss, 1969. Long period signal processing results for the Large Aperture Seismic Array, Geophysics, 34, 305-329.

Simons, R., 1968. PHILTRE-A surface wave particle motion discrimination process, Bull. Seism. Soc. Am., 58, 629-637.

Choy, G. and K. McCamy, 1973. Enhancement of long-period signals by time varying adaptive filters, J. Geophys. Res., 78, 3505-3511.

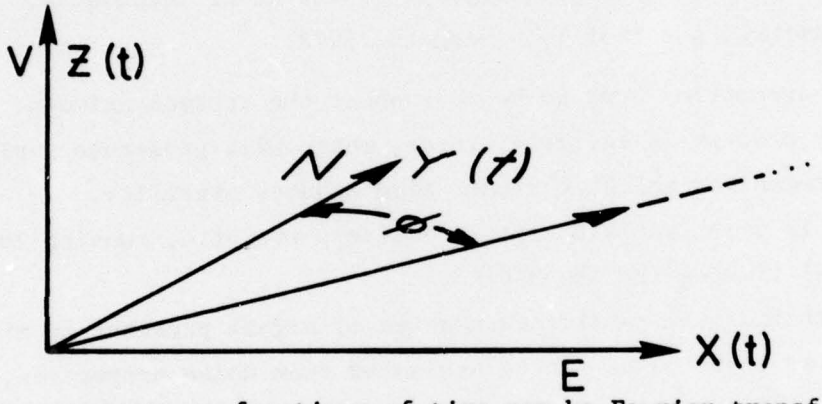
Sobel, P. A., and D. H. von Seggern, 1975. Performance of the Philtre Processor at Low Signal-to-Noise Ratios, SDAC-TR-75-6, Teledyne Geotech, Alexandria, Virginia.

von Seggern, D. H., 1977. Methods of automating routine analysis tasks in preparing a global seismic bulletin, SDAC-TR-77-13, Teledyne Geotech, Alexandria, Virginia.

DEVELOPMENT OF SIGNAL MODEL ERROR EXPRESSION

Searching three-component single station records for possible surface waves may be done by modeling a surface wave and computing the difference between the model and the data, and then optimizing the model to minimize that difference. As the following development shows this optimization can be performed analytically; no computer search is required.

To model and optimize the surface wave, consider its Fourier coefficients at a single frequency and then compare them to the coefficients of the data. Suppose the earth motion has been measured along vertical, north, and east axes, as in this illustration.



The three components as functions of time may be Fourier transformed:

$$Z(t) \Longrightarrow z_n + iz'_n$$

$$Y(t) \Longrightarrow y_n + iy'_n$$

$$X(t) \Longrightarrow x_n + ix'_n$$

where the terms at the right are the Fourier coefficients, at the nth frequency, of the vertical, north, and east components, respectively, of the recorded data.

Hypothesize a Love-Rayleigh signal traveling along the azimuth Ø
The transform of the vertical component may be represented by

$$r_n + ir'_n$$

The horizontal component of the Rayleigh motion is 90 degrees ahead of the vertical and it has an amplitude ε_n times that of the vertical, where ε_n is the ellipticity ratio of the orbital particle motion.

So its transform is

$$\varepsilon_n(r_n + ir_n')e^{i\pi/2} = -\varepsilon_n r_n' + i\varepsilon_n r_n$$

The transform of the Love wave may be represented by

$$\ell_n + i\ell_n'$$

The north and east components of the hypothesized signal are, then,

Rayleigh

North: $\cos\phi(-\varepsilon_n r_n' + i\varepsilon_n r_n)$

East: $\sin\phi(-\varepsilon_n r_n' + i\varepsilon_n r_n)$

Love

North: $\cos(\phi + \pi/2)\ (\ell_n + i\ell_n') = -\sin\phi(\ell_n + i\ell_n')$

East: $\sin(\phi + \pi/2)\ (\ell_n + i\ell_n') = \cos\phi(\ell_n + i\ell_n')$

The Love particle motion is now orthogonal to the Rayleigh, as required.

Taking

$$\alpha \equiv \cos\phi$$

$$\beta \equiv \sin\phi$$

measure the degree of fit -- or the difference -- between the data and the hypothesized signal. It is

$$\begin{pmatrix} E(z_n) \\ E(z_n') \\ E(y_n) \\ E(y_n') \\ E(x_n) \\ E(x_n') \end{pmatrix} = \begin{pmatrix} z_n \\ z_n' \\ y_n \\ y_n' \\ x_n \\ x_n' \end{pmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -\alpha\varepsilon_n & -\beta & 0 \\ \alpha\varepsilon_n & 0 & 0 & -\beta \\ 0 & -\beta\varepsilon_n & \alpha & 0 \\ \beta\varepsilon_n & 0 & 0 & \alpha \end{bmatrix} \bullet \begin{pmatrix} r_n \\ r_n' \\ \ell_n \\ \ell_n' \end{pmatrix}$$

where the terms on the left are, collectively, the data less the signal model. The variance of those terms is the error, E.

If the azimuth is allowed to vary with frequency, then E may be minimized frequency-by-frequency. Thus, at a given frequency

$$E = (r - z)^2 + (r' - z')^2 + (-\beta\varepsilon r' + \alpha\ell - x)^2 + (\beta\varepsilon r + \alpha\ell' - x')^2 + (-\alpha\varepsilon r' - \beta\ell - y)^2 + (\alpha\varepsilon r - \beta\ell' - y')^2$$

Rather than writing partial differential equations with respect to r, r', ℓ, ℓ', ε, and ϕ to minimize E_n, note that there are six unknown and six data in E_n. Therefore, E_n minimized will equal zero, and each of the six terms in E_n, which are squared, will equal zero, thus leaving six equations in six unknowns. So

$$r = z$$
$$r' = z'$$

and also

$$-\beta\varepsilon r' + \alpha\ell = x \qquad \beta\varepsilon r + \alpha\ell' = x'$$
$$-\alpha\varepsilon r' - \beta\ell = y \qquad \alpha\varepsilon r - \beta\ell' = y'$$

From these equations, by elimination,

$$\ell = \alpha x - \beta y$$
$$\ell' = \alpha x' - \beta y'$$
$$\alpha = \varepsilon(xz + x'z')/(xy' - x'y)$$
$$\beta = \varepsilon(yz + y'z')/(xy' - x'y)$$
$$\varepsilon = \pm(xy' - x'y)\left[(xz + x'z')^2 + (yz + y'z')^2\right]^{-\frac{1}{2}}$$

Since

$$\varepsilon > 0$$

the $\pm$ ambiguity is resolved and the azimuth, ϕ, is thus fixed by

$$\alpha = \cos\phi$$
$$\beta = \sin\phi$$

This demonstrates that any given three-component surface wave record may be uniquely transformed into the complex spectra of the apparent Love and vertical Rayleigh components, plus the ellipticity ratio, $\varepsilon(f)$, and the azimuth, $\phi(f)$, both as functions of frequency. Since the ellipticity ratio is a function of station and frequency, not source, such measurements of $\varepsilon(f)$ for a suite of events at a given station can be stacked and averaged to yield a stable estimate. Once determined, the ellipticity ratio ceases to be an unknown in the model and it becomes a known input parameter to the processor, giving one degree of freedom per frequency to the solution for the signal model. Knowledge of $\varepsilon(f)$ would, of course, not only permit optimum operation of this surface wave signal processor, but it would also make it possible to infer the structure beneath the station.

A second degree of freedom per frequency is achieved from the hypothesis that the azimuth is a constant, at least over each of several bands of frequency in the range of interest, if not over the entire spectrum. We proceed with the development then, treating the ellipticity ratio ε_n as a known input parameter, and the azimuth as a single unknown to be solved for over a frequency band of arbitrary width. The several spectra of the surface wave model are solved frequency-by-frequency.

Estimation of Signal Spectra

Taking partial derivatives of E, with respect to the various spectra, ($\mathcal{r}_n$, etc.), and setting those derivatives equal to zero to minimize the error, yields

$$\mathcal{r}_n = (1 + \varepsilon_n^2)^{-1}[z_n + \varepsilon_n(\alpha y'_n + \beta x'_n)]$$

$$\mathcal{r}'_n = (1 + \varepsilon_n^2)^{-1}[z'_n - \varepsilon_n(\alpha y_n + \beta x_n)]$$

$$\ell_n = \alpha x_n - \beta y_n$$

$$\ell'_n = \alpha x'_n - \beta y'_n.$$

If these terms are substituted into the error expression, then the error is minimized in relation to the spectra of the model, for any arbitrarily chosen azimuth $\emptyset$. Note that now the error is equal to the variance of the data minus the variance of the model.

The minimized error, still a function of azimuth at this point, is now written

$$E = S - 2\alpha A - 2\beta B - 2\alpha\beta C - \alpha^2 D - \beta^2 G - H$$

where

$$S \equiv \sum_{n=k}^{K} z_n^2 + z'^2_n + y_n^2 + y'^2_n + x_n^2 + x'^2_n$$

$$A \equiv \sum_{n=k}^{K} (y'_n z_n - y_n z'_n)\varepsilon_n(1 + \varepsilon_n^2)^{-1}$$

$$B \equiv \sum_{n=k}^{K} (x'_n z_n - x_n z'_n)\varepsilon_n(1 + \varepsilon_n^2)^{-1}$$

$$C \equiv -\sum_{n=k}^{K} (x_n y_n + x'_n y'_n)(1 + \varepsilon_n^2)^{-1}$$

$$D \equiv \sum_{n=k}^{K} (x_n^2 + x'^2_n)(1 + \varepsilon_n^2)^{-1}$$

$$G \equiv \sum_{n=k}^{K} (y_n^2 + y'^2_n)(1 + \varepsilon_n^2)^{-1}$$

$$H \equiv \sum_{n=k}^{K} \left[(z_n^2 + z'^2_n) + \varepsilon_n^2 (y_n^2 + y'^2_n + x_n^2 + x'^2_n)\right](1 + \varepsilon_n^2)^{-1}$$

Estimation of Signal Azimuth

To minimize the error with azimuth, the partial derivative of E with respect to ϕ may now be taken and set to zero, noting that since

$$\alpha = \cos\phi \text{ and } \beta = \sin\phi$$

$$\frac{\partial\alpha}{\partial\phi} = -\beta \quad \text{and} \quad \frac{\partial\beta}{\partial\phi} = \alpha$$

So

$$\frac{\partial E}{\partial\phi} = 2\beta A - 2\alpha B - 2(\alpha^2 - \beta^2)C + 2\alpha\beta (D - G)$$

Taking $P \equiv D - G$ and setting the derivative to zero

$$\alpha B - \beta A + (\alpha^2 - \beta^2)C - \alpha\beta P = 0$$

$$2C\alpha^2 + B\alpha - C = \beta(A + \alpha P)$$

$$[2C\alpha^2 + B\alpha - C]^2 = \beta^2(A + \alpha P)^2 = (1 - \alpha^2)(A + \alpha P)^2$$

$$\alpha^4(4C^2 + P^2) + \alpha^3(4BC + 2AP) + \alpha^2(A^2 + B^2 - 4C^2 - P^2)$$

$$+ \alpha[-2(AP + BC)] + (C^2 - A^2) = 0$$

A quartic equation is now formed in which the unknown is α, the cosine of azimuth. There is a general analytic solution for quartic equations, and so α is determined analytically. The sine of the azimuth, β, may be similarly solved for. These two equations each have multiple roots, but the extraneous roots are eliminated by restraints that α and β must be real and the $\alpha^2 + \beta^2$ must equal 1. Thus, the optimum azimuth is determined analytically leaving a unique solution. The resulting ϕ is the best azimuthal estimate over the band of frequencies included in the summation from n=k to K in the error expression. (Note: No 180° ambiguity exists in the azimuth because the model, as observation and theory of the fundamental mode demand, requires the horizontal component of Rayleigh wave particle motion to lead the vertical component in the direction of propagation.)

The solutions for optimized α and β can be substituted into the expressions for the spectra , r_n, r'_n, ℓ_n, and ℓ'_n, and into the error expression, which is thereby minimized. Because it can be solved for analytically, without a computer search, this estimated surface wave can be found quickly. The user is free to solve for the best estimate over any frequency band, or over any combination of bands, as the variable limits k and K on the summation in E indicate.

Testing Signal Presence

From these results a convenient F-statistic can be drawn to measure the likelihood that a given time window of data contains a surface wave signal -- although buried in noise -- and that the modeled signal does not result from an accidental combination of noise. The F-statistic for this model is proportional to the ratio of the variance of the modeled surface wave, to the variance of the error, i.e., E.

$$F = \frac{\sum_{n=k}^{K} (1+\varepsilon_n^2)(r_n^2 + r'^2_n) + \ell_n^2 + \ell'^2_n}{E} \left\{ \frac{\frac{1}{4(K+1-k)}}{\frac{1}{(6-4)(K+1-k)}} \right\}$$

It is a ratio of two variances, each divided by the respective number of degrees of freedom. The variance of the model has $4(K + 1 - k)$ degrees of freedom, 4 per frequency. The error expression E is a difference of variances: the variance of the data, with $6(K + 1 - k)$ degrees of freedom, minus the variance of the model. Thus, E has $(6 - 4)(K + 1 - k)$ degrees of freedom. The ratio of degrees of freedom, in parentheses at the right, reduces to $\frac{1}{2}$.

The well studied and widely understood F-statistic provides a convenient index of whether or not a signal is present. If the model and the data have standard normal distribution, then the statistic has the F distribution. However, even if the distributions are not known the F-statistic can be applied empirically to the detection decision problem. As E is minimized, the F-statistic is maximized thus maximizing the measured likelihood of signal presence. So, within the imposed restraints the optimized model represents the best case for the presence of a surface wave in the data.

APPLICATION OF THIS PROCESSOR TO DETECTION

Synthetic data of precisely known content were used to test this single-station, maximum-likelihood surface wave detector. Using time windows of 1024 points (17 minutes and 4 seconds), 2 days of synthetic noise (170 windows) were processed. Then, 2 days of synthetic surface waves, buried in synthetic noise, of signal-to-noise ratio 0.5, were similarily processed. (Throughout this report signal-to-noise ratio means the ratio of the variances.) The noise spectrum was designed to have the same shape and center frequency as the signal to provide a worst-possible-case trial. Figures 1 through 6 illustrate the synthetic data used to test the detection processor.

Figures 1-4 show the amplitude and group velocity spectra of the Love and Rayleigh components of the synthetic surface wave. The group velocity spectra, after Brune (1969), are for a mid-continent crust. The amplitude spectrum is taken from the response of the LASA LP instrumentation, approximately. These amplitude and group velocity spectra, Fourier transformed into the time domain, are shown in Figure 5. The vertical, north, and east components of these signals, buried in noise (S/N = ½), are shown in Figure 6. Note that time domain Figures 5 through 10 are all to the same arbitrary amplitude scale.

Because this detection algorithm estimates the azimuth and complex spectra of the desired surface wave, the estimate may be transformed for display into the time domain. The best estimate on the data of Figure 6 is shown in Figure 7. Compare this estimated signal with the original synthetic surface wave and then with the same surface wave mixed with synthetic noise.

The next figures demonstrate that the signal estimates of the surface wave detector result from a noise rejection process and not simply from input records rotated to a suitable source azimuth. Consider the record of Figure 8 that shows the same synthetic surface wave generated above, only buried in noise such that the signal-to-noise ratio is now 0.1; the signal was made to arrive from 126°. Figure 9 displays the signal estimate computed by the detector. The detector estimated the back azimuth at 114°.

Brune, J. N., 1969. Surface waves and crustal structure, The Earth's Crust and Upper Mantle, Geophysical Monograph 13: AGU, Washington, D.C.

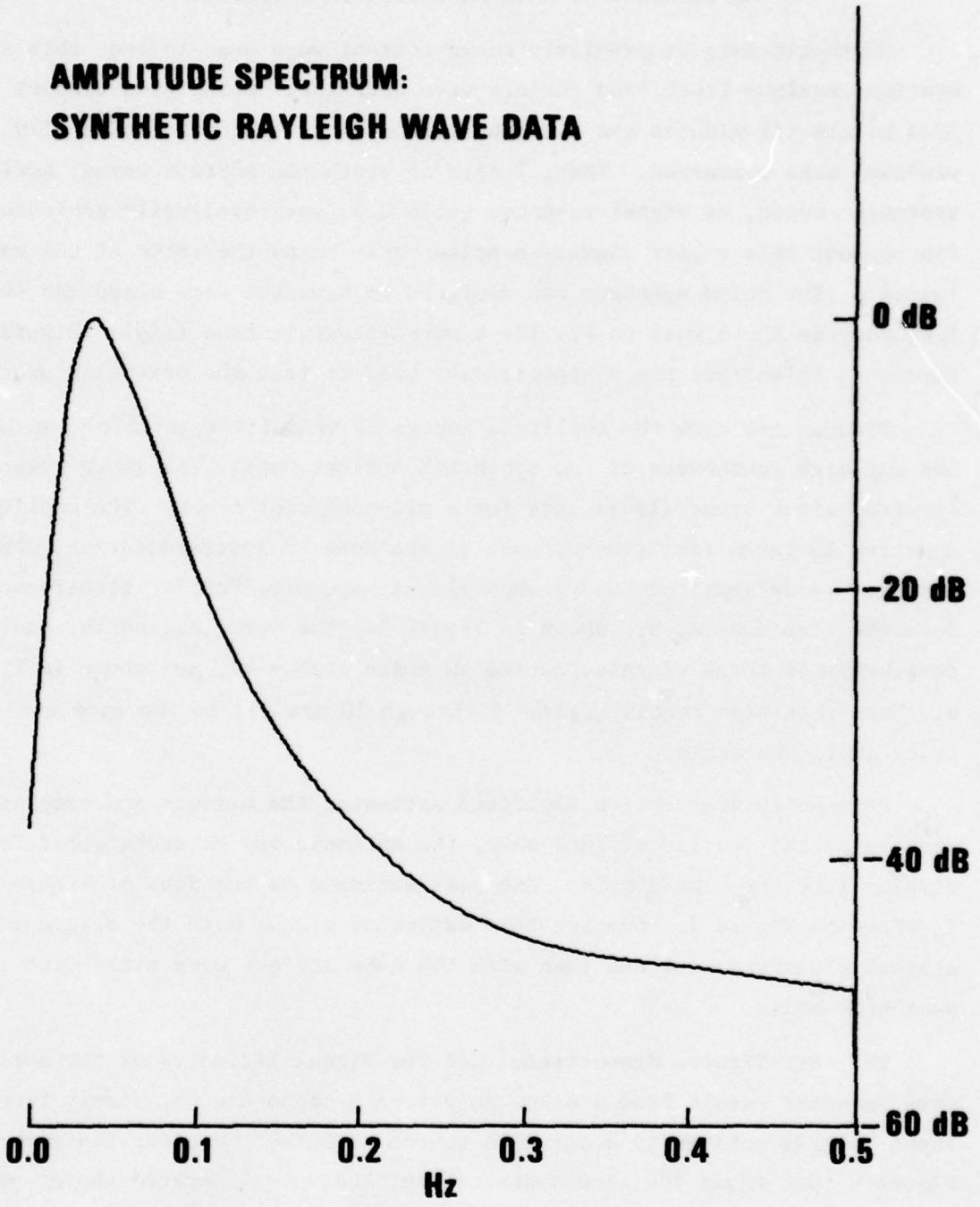


Figure 1 The amplitude spectrum of the synthetic Rayleigh wave data.

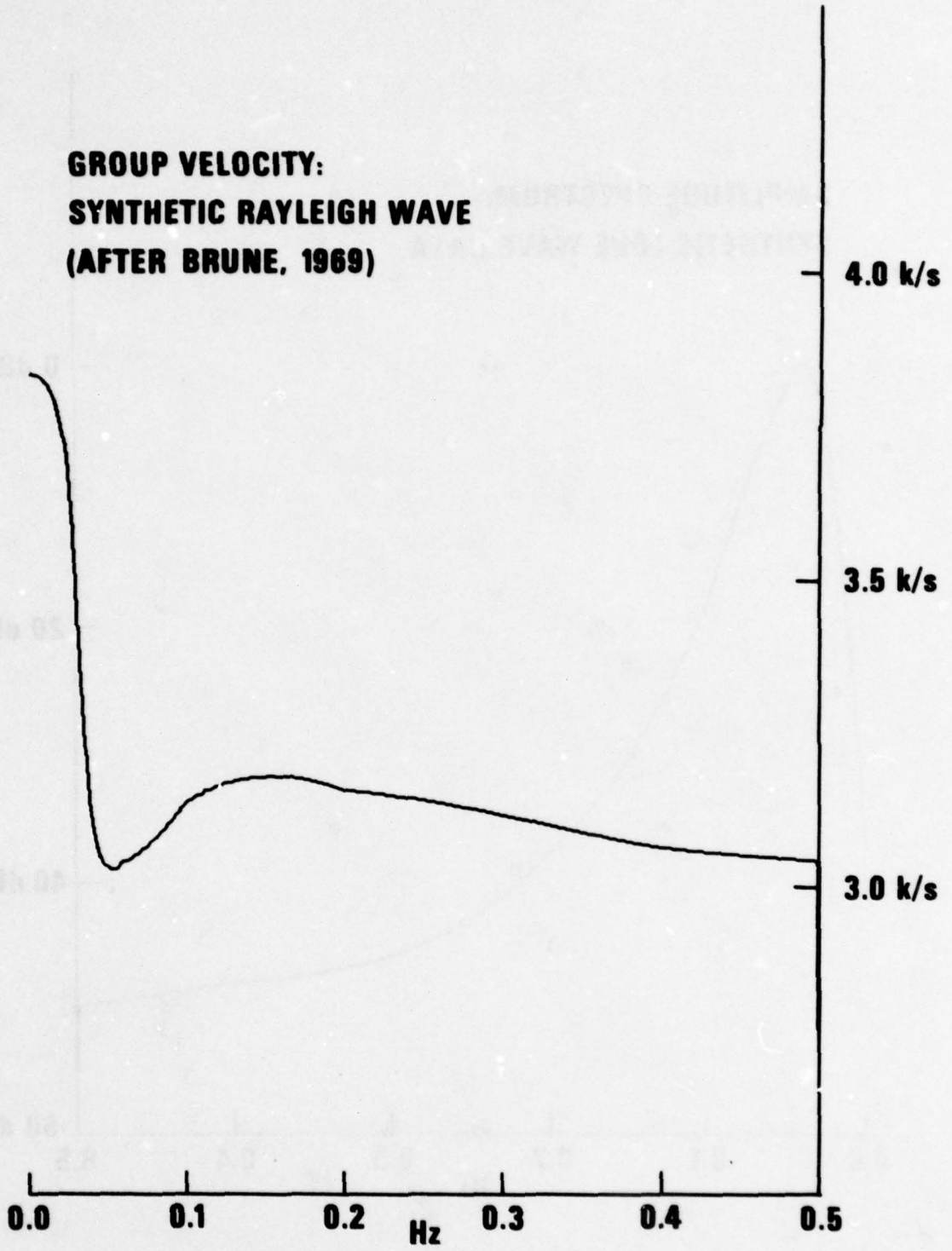


Figure 2 The group velocity spectrum of the synthetic Rayleigh wave.

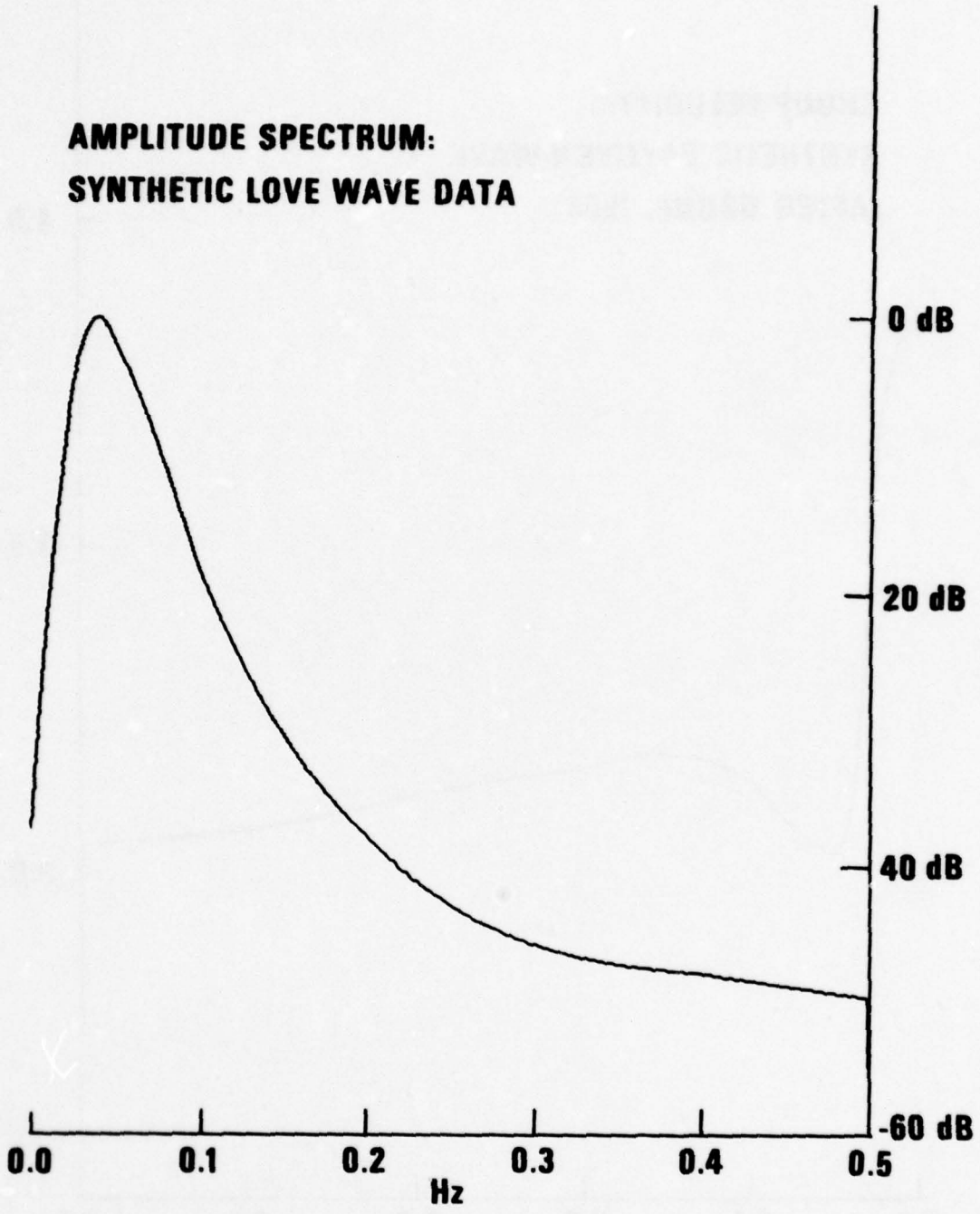


Figure 3 The amplitude spectrum of the synthetic Love wave data.

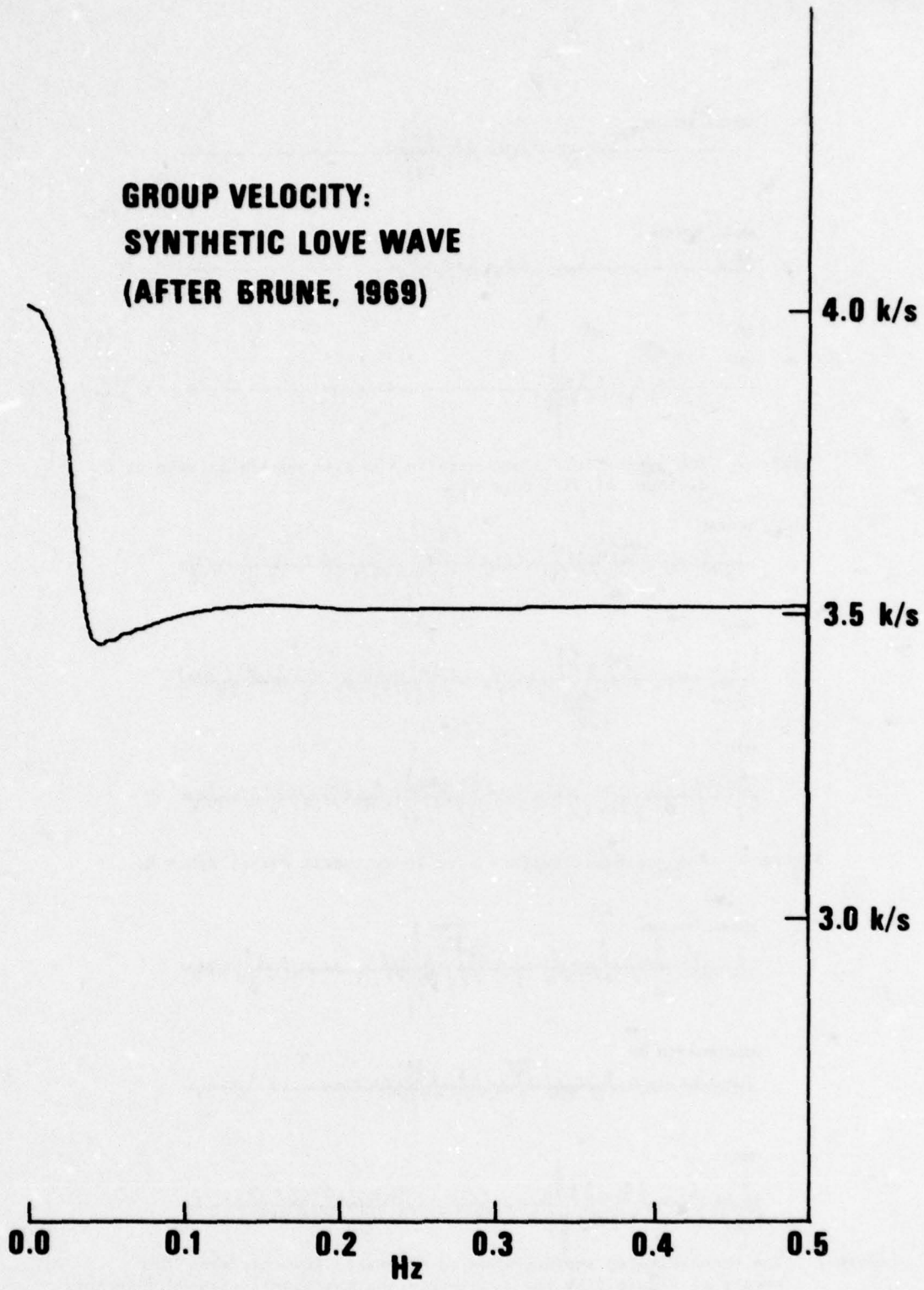


Figure 4 The group velocity spectrum of the synthetic Love wave.

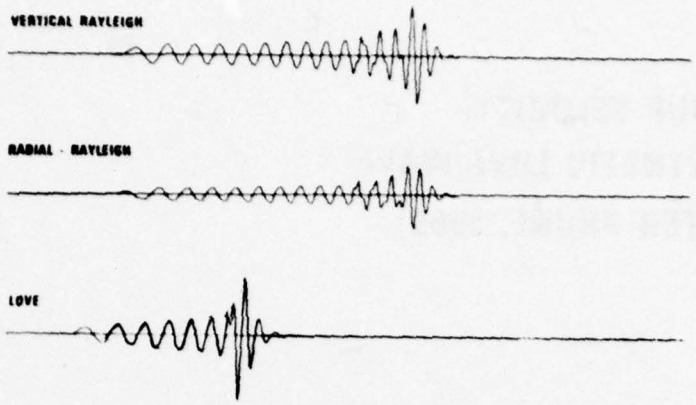


Figure 5 The synthetic surface wave in the time domain as seen at a distance of 58.5 degrees.

Figure 6 The synthetic surface wave in synthetic noise; S/N = ½.

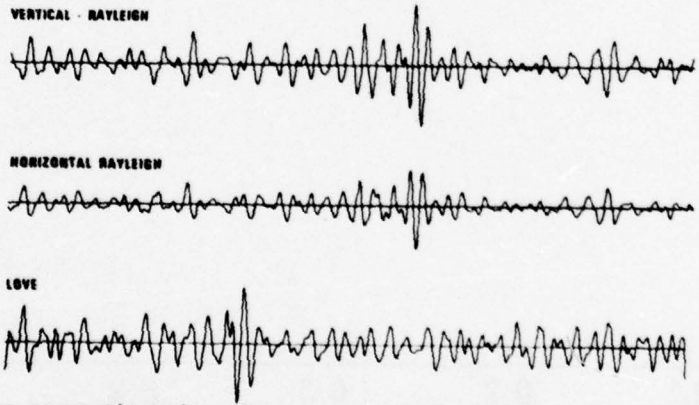


Figure 7 The reconstituted surface wave as estimated from the synthetic record of Figure 6 by the single-station maximum-likelihood detector.

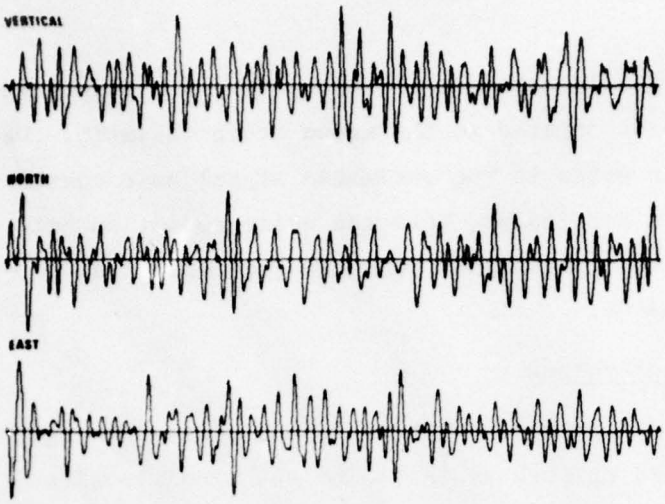


Figure 8 The synthetic surface wave in synthetic noise; S/N = 1/10.

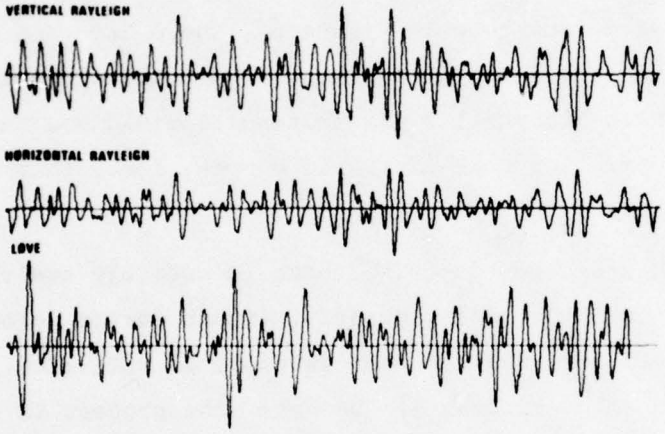


Figure 9 The surface wave as reconstructed from the synthetic record of Figure 8 by the single-station maximum-likelihood detector.

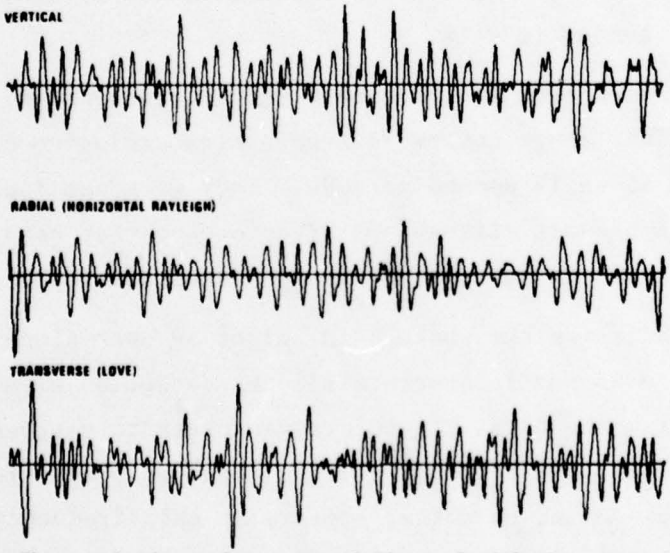


Figure 10 The synthetic surface-wave of Figure 8 with the components simply rotated to the known back azimuth.

The F-statistic (discussed in more detail below) is 1.8. Figure 10 displays the input records rotated to the known source azimuth. Observe the evident reduction in noise in the estimated signal as compared to the rotated records. In particular, note the noise reduction between the radial component of the input and the estimated horizontal component of the Rayleigh wave.

Measuring Detector Effectiveness

As noted above, 170 suites of synthetic data, corresponding to two days of records, were composed of pure synthetic noise. Additional suites of data were generated by burying the synthetic surface wave of Figure 5 in each of 170 windows of synthetic noise. In each window the signal-to-noise ratio was exactly 0.5, as in the case in Figure 6. These 340 time windows were then processed by the detection algorithm. (The printed result for one such window is shown in Figure 11.) These results yield one curve of the family of receiver operating characteristic curves, i.e., that at S/N = ½.

Note again that the processor described here is entirely analytic and that no computer search is involved (or needed), either during determination of the "best" back azimuth, or when the best waveform is estimated. Since the algorithm eliminates all searching of the data, the process is very rapid. On the IBM 360/44 it is 2 orders of magnitude faster than real time. Processing the entire 4 days of data lasted only 1 hour and 10 minutes, including generation of synthetic data.

The data were processed over the frequency range corresponding to 10-50 second periods. This range covers 72 frequencies exclusive of the 12 in the interval from 15 to 18 second periods. They were not included in the interests of realism because microseisms often occupy that band and it could not be counted upon to strengthen signal "detectability".

While the F-statistic, for the whole band, might be used alone as a detection criterion, it does not incorporate all the detection significance that can be extracted from the data. Therefore, in order to sharpen the sensitivity of the detector, the replicated signal back-azimuth information was used which, though distorted by noise, appears at each frequency in the signal band. Most of the energy from an event that arrives at a station

INPUT SIGNAL-TO-NOISE RATIO = 0.50
INPUT AZIMUTH = 0.0
NEXT NOISE GENERATOR NUMBER = 1175617537

ERROR OF ESTIMATED AZIMUTH IN DEGREES	F-STATISTIC x2	RANGE OF PERIODS IN SEC.	NO. OF FREQUENCIES
5.343	0.594E 01	48.8-26.9	18
5.513	0.536E 01	26.3-18.3	18
-2.011	0.286E 01	15.1-12.0	18
-65.885	0.326E 01	11.9- 9.9	18

RESULTS OF THE ANALYSIS OF 4 SEPARATE FREQUENCY BANDS OF THE RECORD

THE FOLLOWING LIST REPEATS, IN AZIMUTHAL ORDER, ALL THE ABOVE VECTORS AT THE LEFT HAVING F-STATISTICS GREATER THAN 2.8. FOR EACH VECTOR THE LAST COLUMN GIVES, IN DEGREES, THE DISTANCE FROM THE PRECEDING VECTOR.

AZIMUTH	FREQUENCY	F-STAT	PERIOD	DEGREES
5	0.021	6	26.95	7.4
6	0.038	5	18.29	0.2
294	0.084	3	9.94	288.6
358	0.066	3	12.05	63.9

THE MOST ANOMALOUS CONCENTRATION OF VECTORS IN THE ABOVE LIST OF 4 IS THAT OF THE GROUP OF 3 VECTORS BETWEEN 358 and 6 DEGREES. THE EQUIVALENT OF 3 OR MORE OUT OF 4 VECTORS WITHIN 7.5 DEGREES OR LESS WILL OCCUR AT RANDOM ON THE AVERAGE OF ONCE EVERY 0.76 DAYS.
THE LEAST SQUARES MEAN OF THOSE 3 UNIT VECTORS HAS AN AZIMUTH OF 3 DEGREES, WITH AN EMS ERROR OF 0.061.

ERROR OF ESTIMATED AZIMUTH IN DEGREES	F-STATISTIC x2	RANGE OF PERIODS IN SEC.	NO. OF FREQUENCIES
4.448	0.504E 01	48.0-18.0 15.0- 9.0	72

RESULTS OF THE ANALYSIS OF THE SYNTHETIC SIGNAL-IN-NOISE OVER ALL 72 FREQUENCIES

Figure 11 Sample Detector bulletin from evaluation tests.

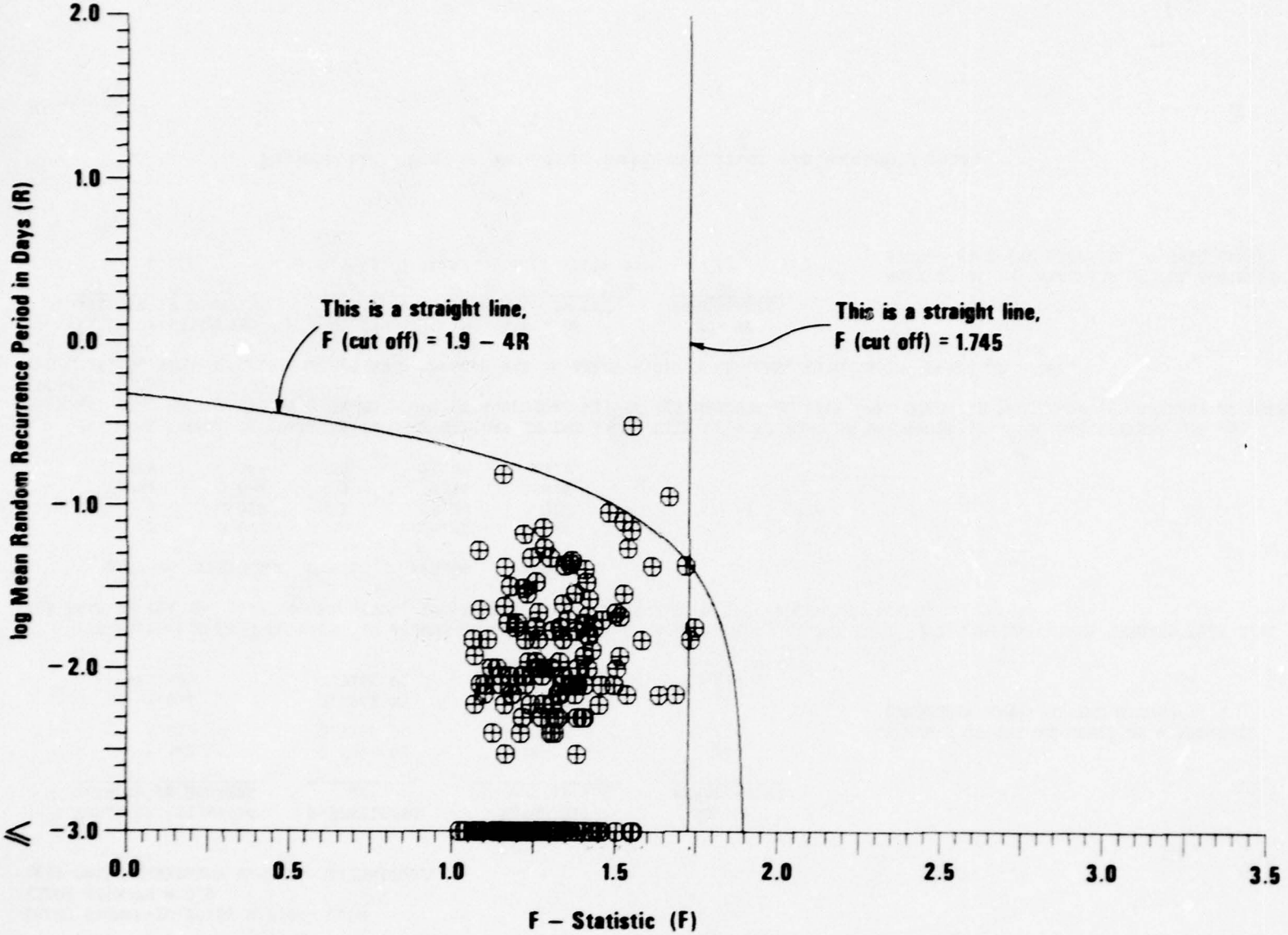


Figure 12 A plot of mean random recurrence period versus F-statistic for 170 times windows (2 days) of processed three-component synthetic noise data.

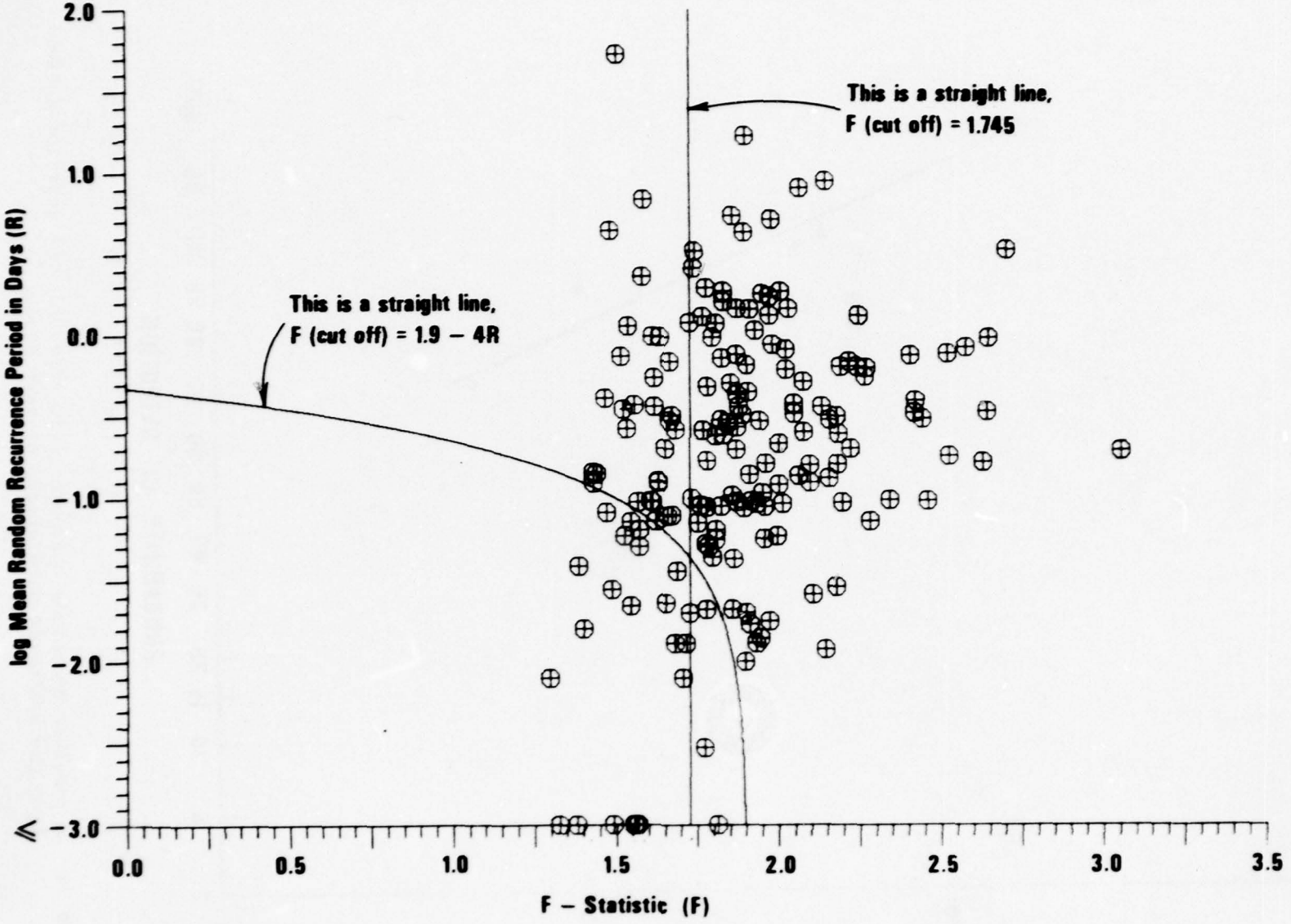


Figure 13 A plot of mean random recurrence period versus F-statistic for 170 time windows (2 days) of processed three-component synthetic surface-wave data in synthetic noise (S/N = ½).

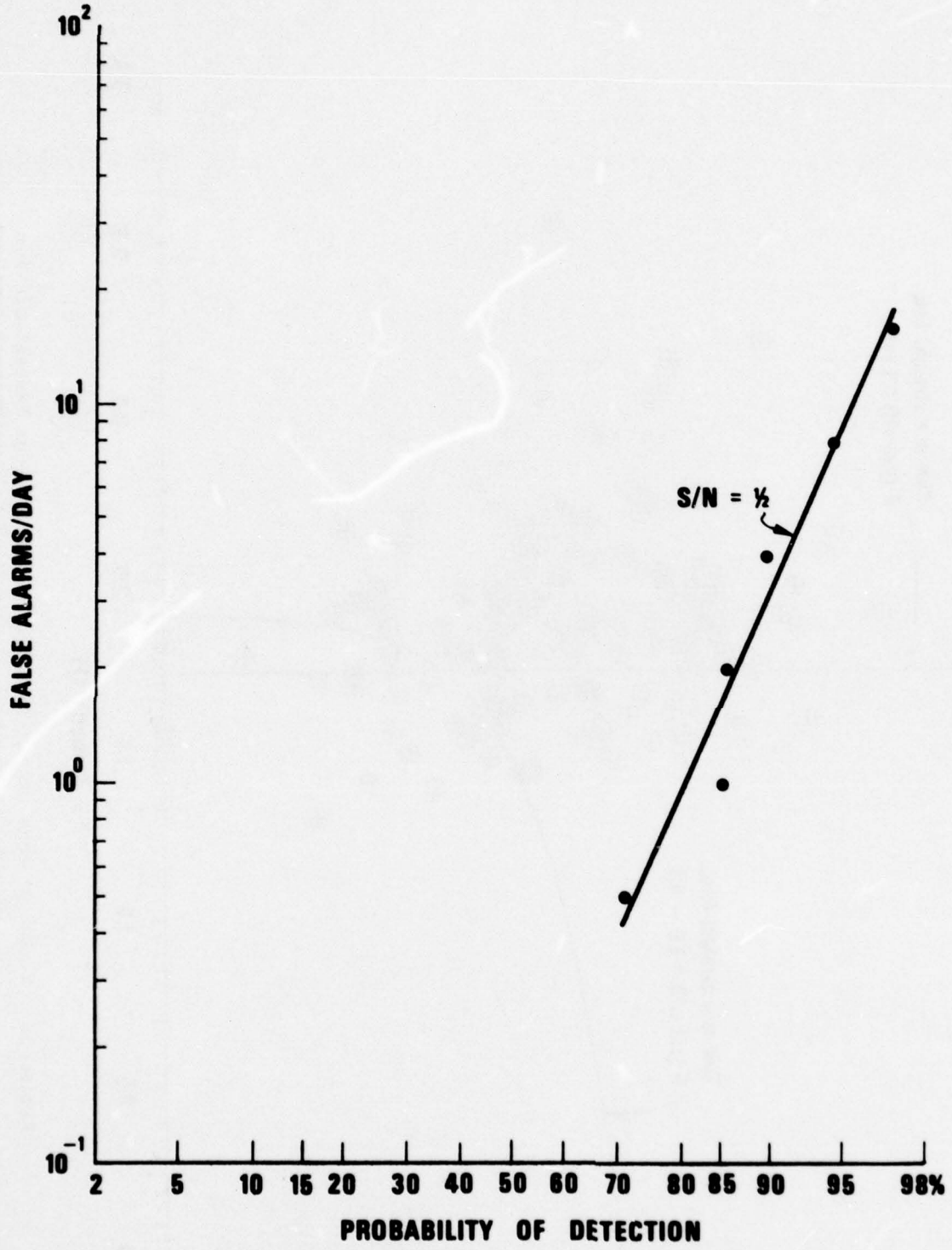


Figure 14 A receiver operating characteristic curve for the detector derived from the synthetic data of the present study.

will come out of the same 90 degree azimuthal sector. This persistence, over frequency, from one general direction is important in deciding if a signal has actually been observed. The problem is, of course, to interpret such a clustering of azimuths and translate that vague sense of significance into an exact quantity.

A Supplementary Detection Statistic

In an unpublished work (Smart, 1974) the author examined this problem, which may be stated as follows: Given an arbitrary number, M, of random azimuths, what is the probability that N or more of them will be clustered within θ degrees or less? Without presenting the derivation here (Smart 1974)

$$f(N,M,\theta) = (M-1)\left\{ M - M! \sum_{k=1}^{N-1} \frac{P^{N-k-1}(1-P)^{M+k-N}}{(N-k-1)!(M-N+k)!} \right\}$$

where $P = \theta/360$. The function f is the average number of times per set of M random azimuths that a cluster of N azimuths within θ degrees (or an equally improbable configuration) will occur.

Given a processor which produces sets of azimuthal measurements at a regular rate, the product of that rate and f, the mean random recurrence frequency of a specified cluster event from one of those sets, is the effective "false alarm" rate of all like events, i.e., those of equal or greater improbability. This expression for $f(N, M, \theta)$ can be used to sweep through a set of azimuths and identify and place a likelihood value on any anomalous clusterings. Of course, M may be any integer greater than 1, so the azimuthal sets can be large or quite small. The procedure is called azimuthal distribution analysis.

To make azimuthal estimates at several frequency bands in each time window, and thus generate a set of azimuths for distribution analysis, the 72 frequencies processed for the overall F-statistic were also processed as four separate bands of 18 frequencies apiece. This also yields

---

Smart, Eugene, 1974. Azimuthal distribution analysis, unpublished.

measurements of F, one for each band, and those F-statistics were employed at this point in a preliminary way. Each azimuth associated with an F-statistic less than 1.4 was dropped from distribution analysis. The distribution of the resulting set of four (or less) azimuths was then analyzed and a significance value placed upon the most anomalous cluster in the form of its mean random recurrence period expressed in days. The significance of clusters of less than three azimuths was arbitrarily set to zero.

Thus, two detection statistics are available at each time window. Below are the results of processing four days of synthetic data. Those results are also presented graphically in Figures 12 and 13.

During the two days of noise observations the F-statistic rose above 1.725 only twice. During the two days of signal-in-noise observations (S/N = ½) F fell to 1.725 or below 51 times (out of 170 observations). Thus, on the basis of the F-statistic alone, with a threshold of $F_T = 1.725$ an average of one false alarm a day, and a detection rate of 70% is expected in the population having S/N = ½. However, without increasing the false alarm rate, the number of missed signals can be reduced by one half.

If instead of

$$F_T = 1.725$$

a threshold of

$$F_T = 1.9 - 4R$$

is used, where R is the mean random recurrence period in days, there are still two false alarms in the two days of noise observations, but only 25 missed detections in the 170 observations of signal in noise. So by means of F and R combined, at a false alarm rate of one day, the processor detected 85% of the signals at S/N = ½. This datum constitutes one point on the receiver operating characteristic curve for that signal-to-noise ratio. The curve generated by the family of F thresholds with R slope -4 is displayed in Figure 14.

Synthetic data of known Rayleigh wave ellipticity ratio have been used thus far in evaluation of the detector. The ellipticity ratio, ϵ, is

the amplitude of the radial component of particle motion divided by that of the vertical. The ratio throughout these synthetic data is 2/3.

The F-statistic versus frequency measured by the processor may be used to further improve signal estimates. It provides an index of where the signal lies in the frequency domain, thus indicating the bandpass filter for the detector output. Such automatically adjusted band filtering can be incorporated into the program containing the detector, as a follow-on process.

For an example of the possibilities, see Figure 15, which shows the synthetic signal of Figure 5 mixed with noise at S/N = ½. Now examine Figure 16 which is the detector's estimate of the signal, automatically bandpass filtered in the processor itself at the band indicated by the F-statistic. This estimate is a good approximation of the input signal in Figure 5.

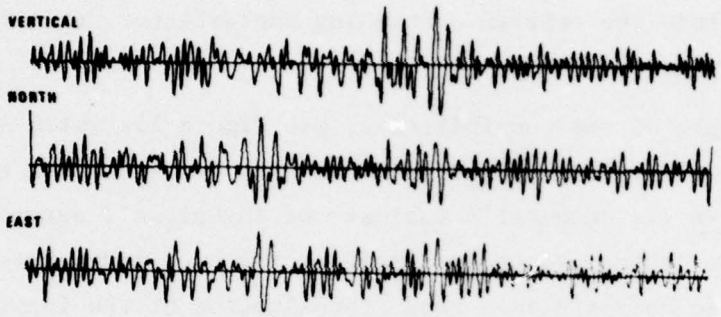


Figure 15 The synthetic signal once more at S/N = ½.

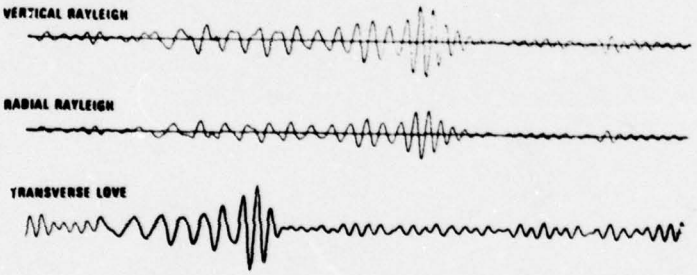


Figure 16 The detector's estimate of the signal buried in the record of Figure 15.

ANALYSIS OF REAL DATA

Setting Arrival Times and Surface-Wave Magnitudes Automatically

Von Seggern (1977) adapted the computer code generated in this study for application to real data, allowing the processor to determine surface wave arrival times and magnitudes automatically. Though this application is only one of several possibilities, his results demonstrate the usefulness of the process in real situations. Of the time windows automatically selected for surface-wave arrivals and M_s measurements he said

> These windows are good approximations to what a human analyst would set as the arrival window of 20-sec LR waves. This process, then, selects the best window, based on the F-statistic, for computation of Rayleigh-wave magnitudes... .

Figure 17 is a tabulation of von Seggern's results from application of the single-station surface wave processor to HGLP and SRO records of the 4 July 1976 Kazakh event. The records themselves appear in Figure 18.

Note that for some stations with evident LR signals no detection was declared by the automatic process. The lack of detection was due either to data irregularities in some cases or to the signal's poor S/N ratio in the .04 to .055 Hz band in others.

Automatic Back Azimuth Determination of Real L_g Signals

In a separate study (Smart, 1977) the author applied this surface-wave process to short period LRSM records of the L_g signals from SALMON and GNOME. The stations used were from 7 to 25 degrees distant from the events. The frequency band explored was about three octaves wide and centered around 1.0 Hz. For want of any measurements of $\varepsilon(f)$ an arbitrary ellipticity ratio of 1.0 was applied to all frequencies. Figures 19 and 21 present the resulting azimuthal estimates, the actual event locations, and the source locations as computed from the azimuthal estimates. Most of the estimated azimuths are in good agreement with the known station-event azimuths, especially considering that each estimate depends on measurements made at a single station. Note that the presence of prograde particle motion mixed with retrograde orbiting has introduced a 180-degree azimuthal ambiguity, which, however, does not interfere with computation of the site of origin,

				18-25 sec				28-51 sec			
		Δ^{o}	E-S azimuth	F-statistic	Start pt	δ(az)	nm·sec	F-statistic	Start pt	δ(az)	nm·sec
AMNO	6011	95	4	21.4	1281	-1^{o}	1610	26.6	833	-17^{o}	1000.
GUMO	6021	64	100	15.8	641	-23^{o}	2370.	28.4	449	-18^{o}	1890.
SNZO	6051	123	116	no detection				no detection			
MAIO	6031	20	233	73.5	193	1^{o}	7040.	44.0	65	-2^{o}	4240.
TATO	6061	41	111	no detection				20.8	257	16^{o}	1510.
CTA	6091	91	120	no detection				18.4	321	23^{o}	613.
CHG	6101	35	146	no detection				26.4	321	30^{o}	2390.
TLO	6111	57	294	no detection				15.3	385	0^{o}	1400.
ELA	6121	39	255	16.7	577	21^{o}	1560.	no detection			
KON	6141	39	312	50.3	385	24^{o}	1749.	33.4	257	4^{o}	2890.
KIP	6151	93	51	40.3	833	-1^{o}	595.	16.4	385	-9^{o}	648.
ALQ	6161	95	4	15.4	1089	12^{o}	581.	47.3	641	-22^{o}	928.
MAT	6151	44	84	no detection				no detection			
ZLP	6171	137	310	15.5	1345	-15^{o}	585.	no detection			

used rotated seismograms, so calculated azimuth should be near 0^{o}

$\Delta\theta < 30^{o}$

$F > 15$

max amp window provided above 2 conditions met

Figure 17 A tabulation of von Seggern's results from application of the single-station surface wave processor to HGLP and SRO records of the 4 July 1976 Kazakh event.

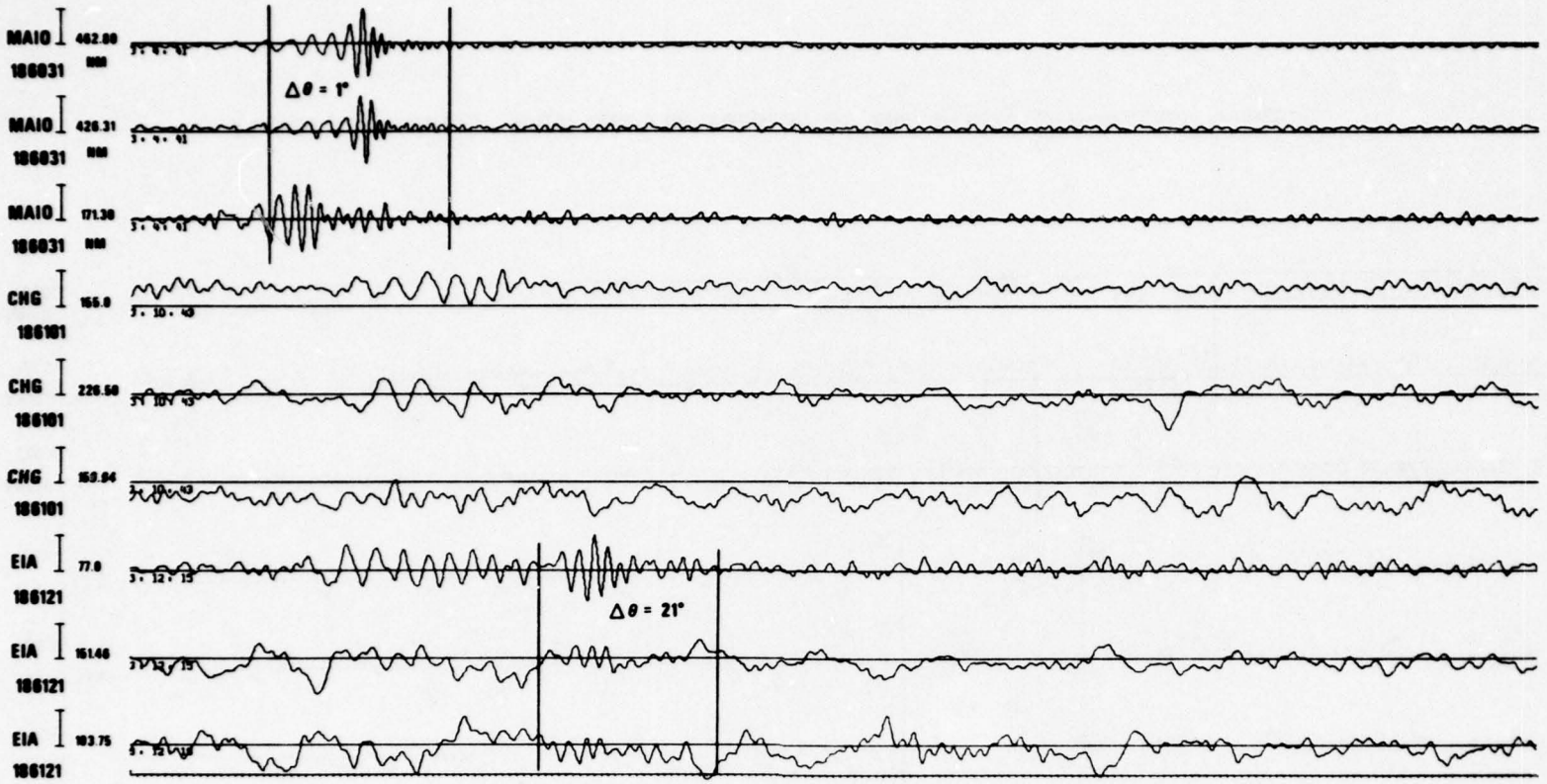


Figure 18 HGLP and SRO records of the 4 July 1976 Kazakh event.

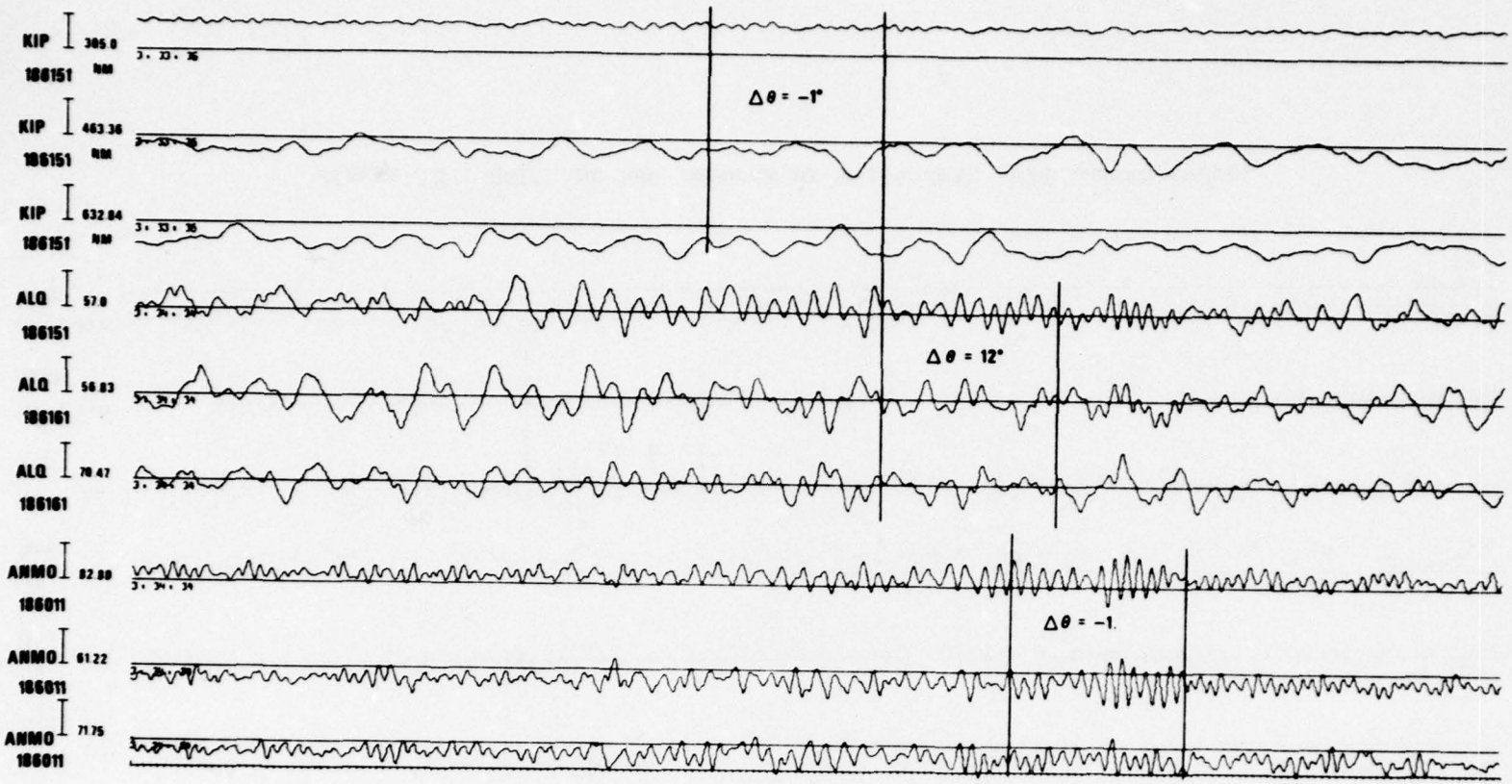


Figure 18 (Cont.) HGLP and SRO records of the 4 July 1976 Kazakh event.

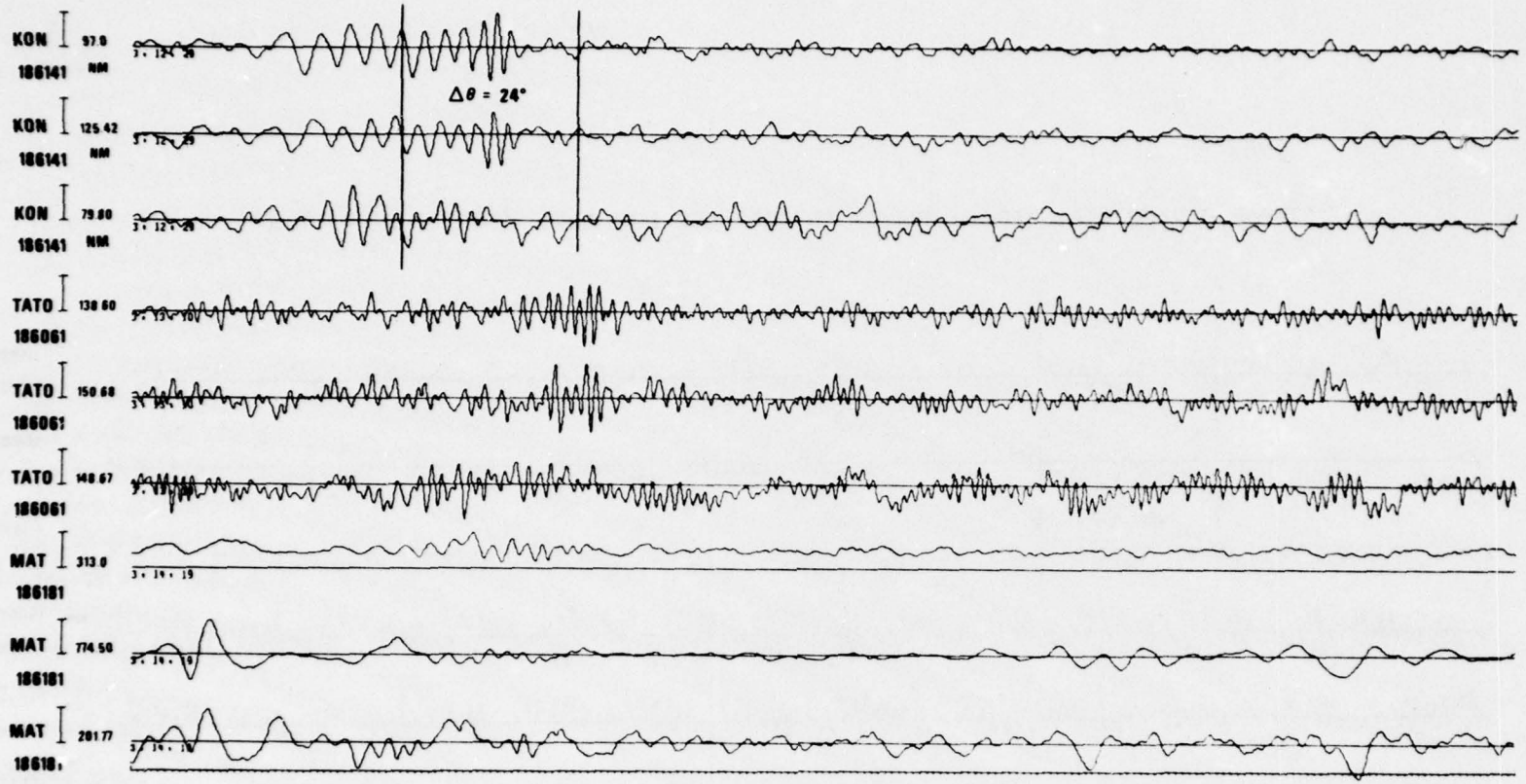


Figure 18 (Cont.) HGLP and SRO records of the 4 July 1976 Kazakh event.

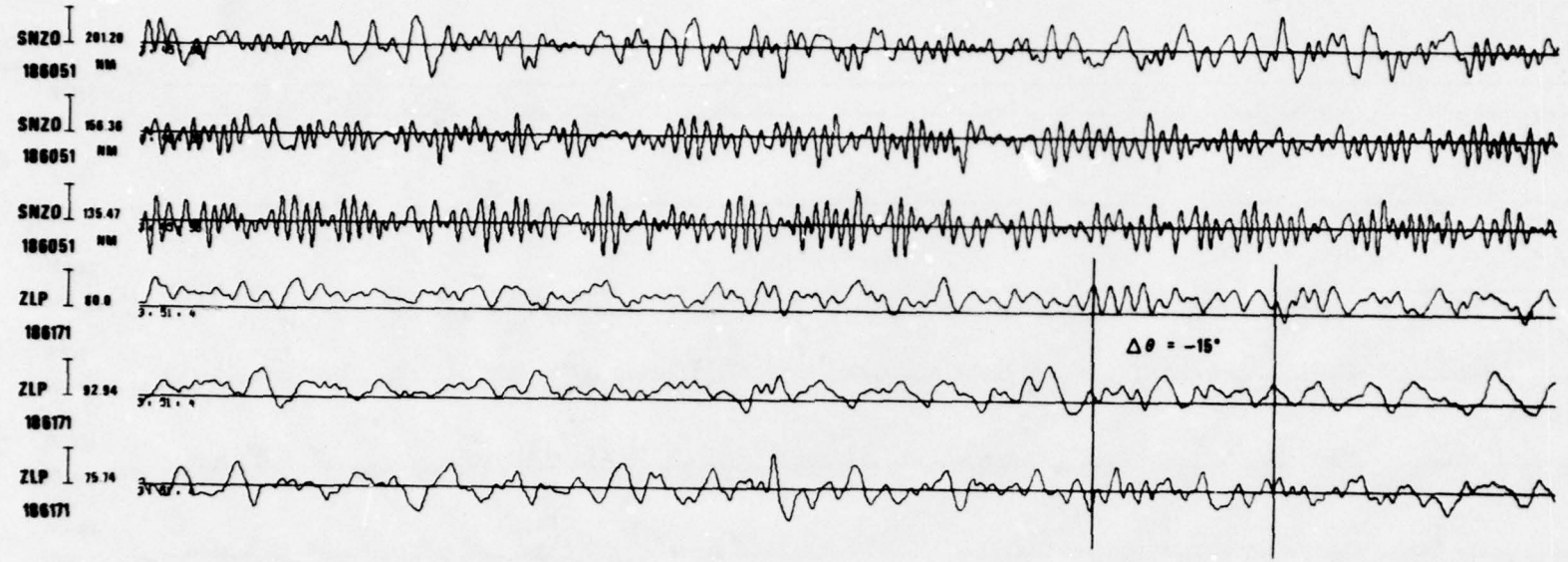


Figure 18 (Cont.) HGLP and SRO records of the 4 July 1976 Kazakh event.

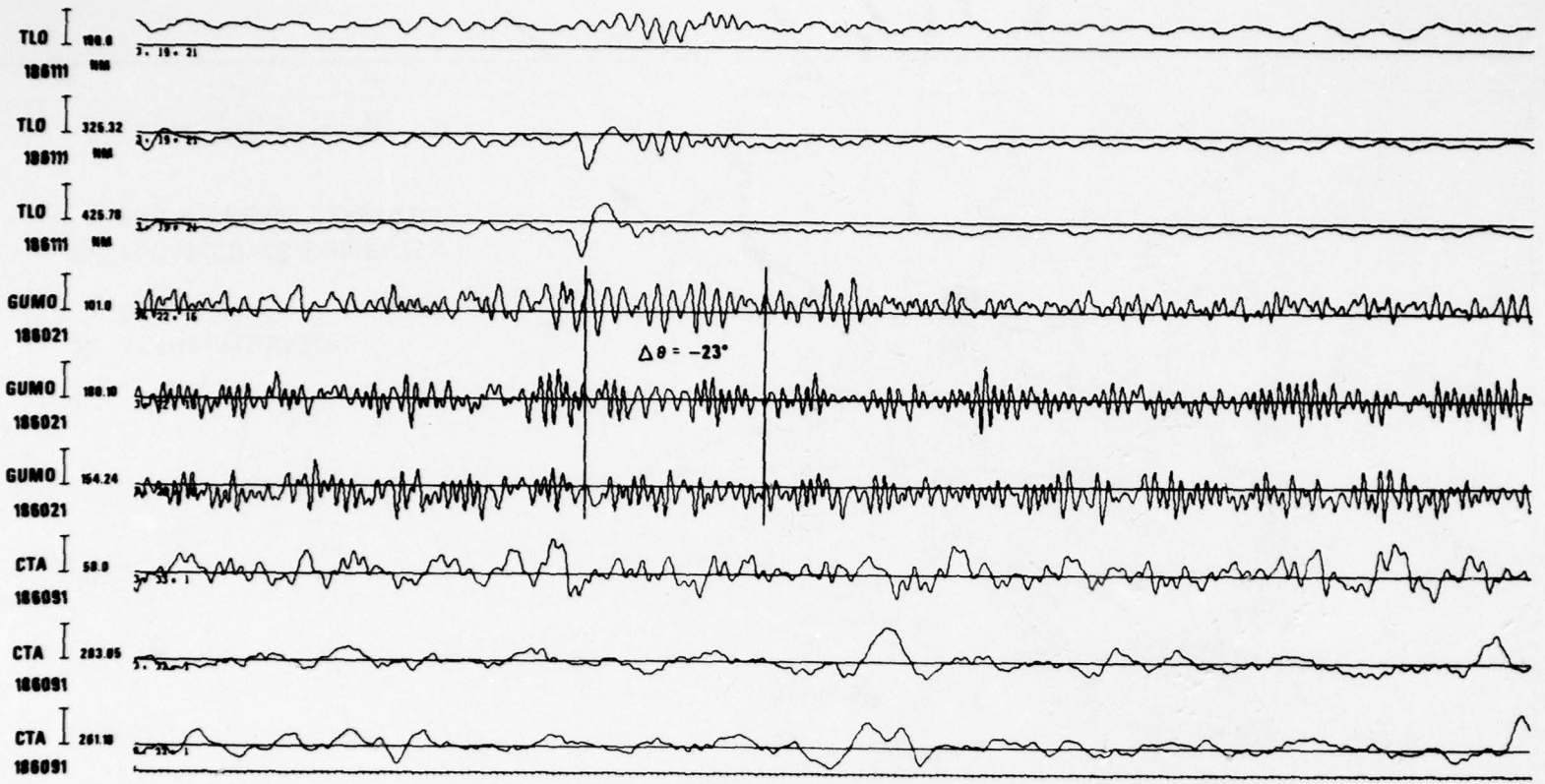


Figure 18 (Cont.) HGLP and SRO records of the 4 July 1976 Kazakh event.

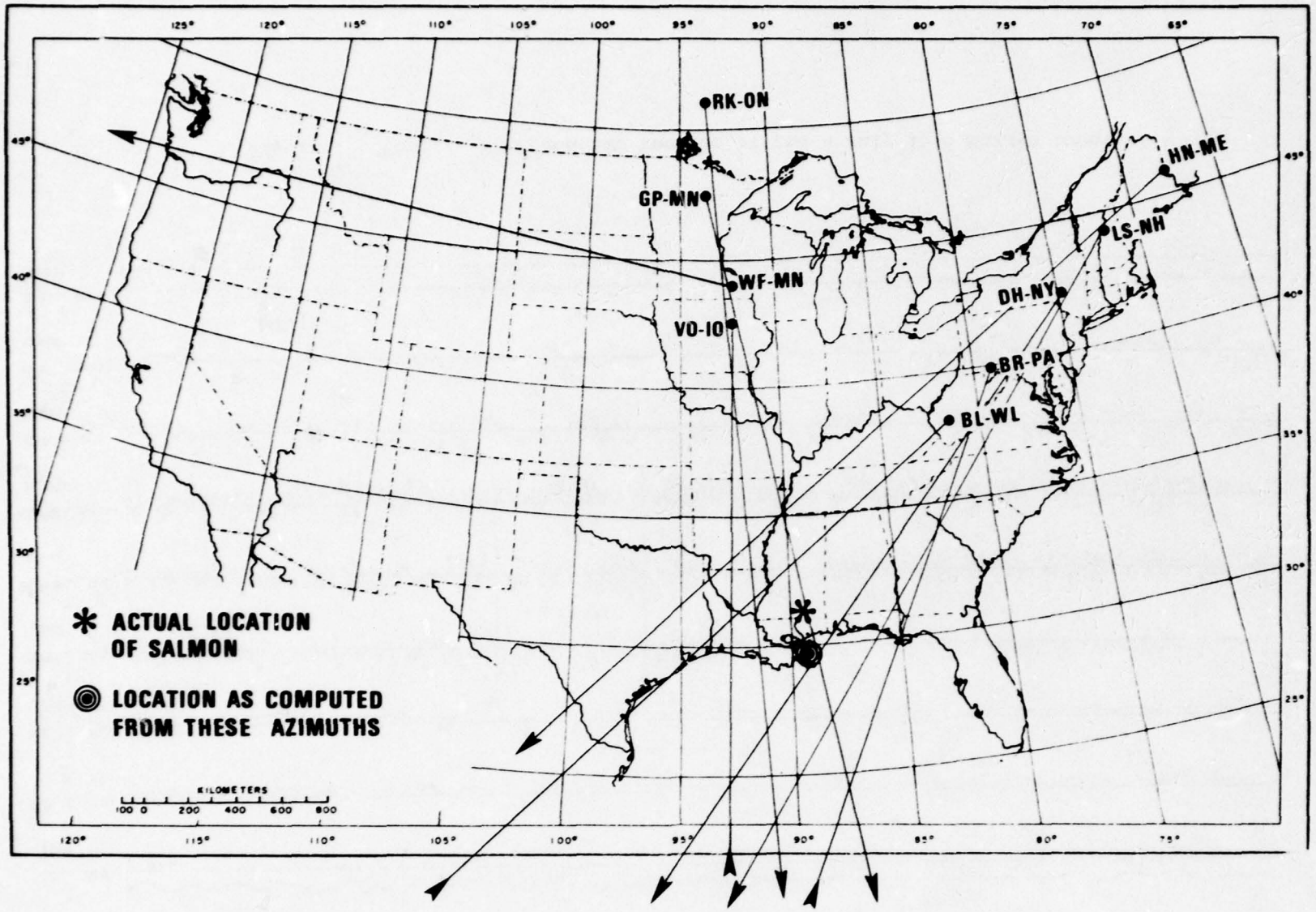


Figure 19 Automatic signal azimuth estimates of L_g surface waves from SALMON as seen at several LRSM stations.

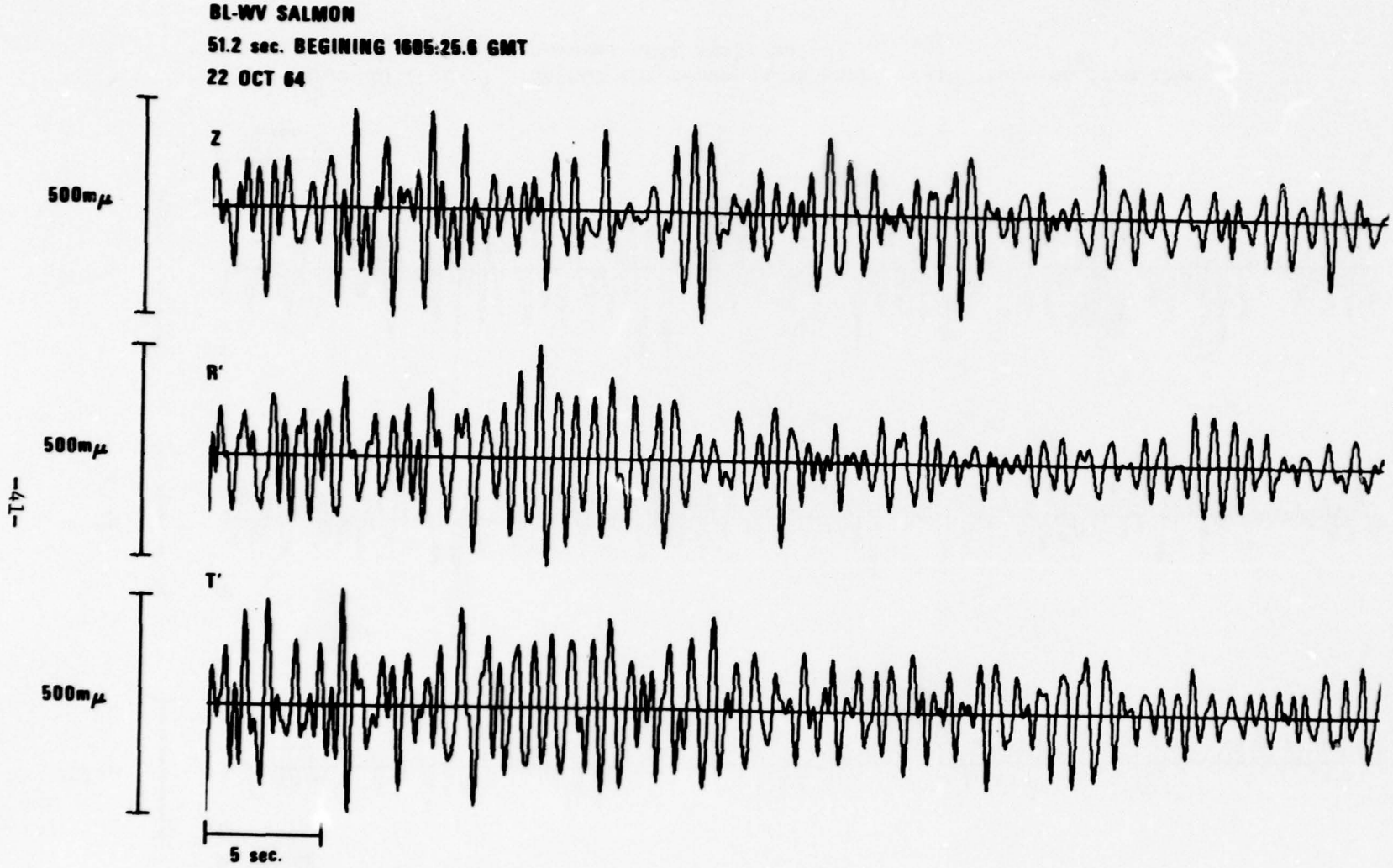


Figure 20 SALMON three-component short-period records from the several LRSM stations.

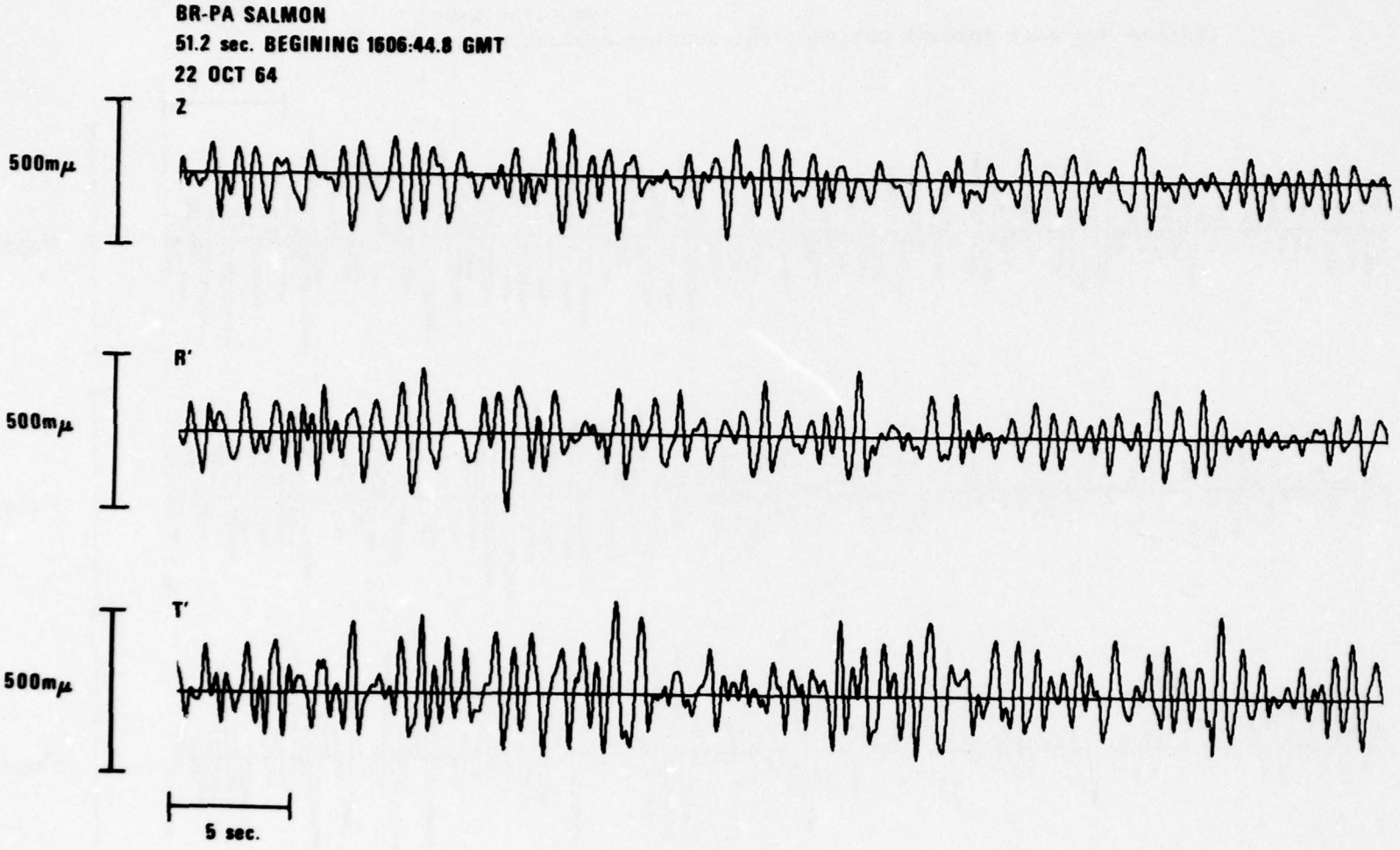


Figure 20 (Cont.) SALMON three-component short-period records from the several LRSM stations.

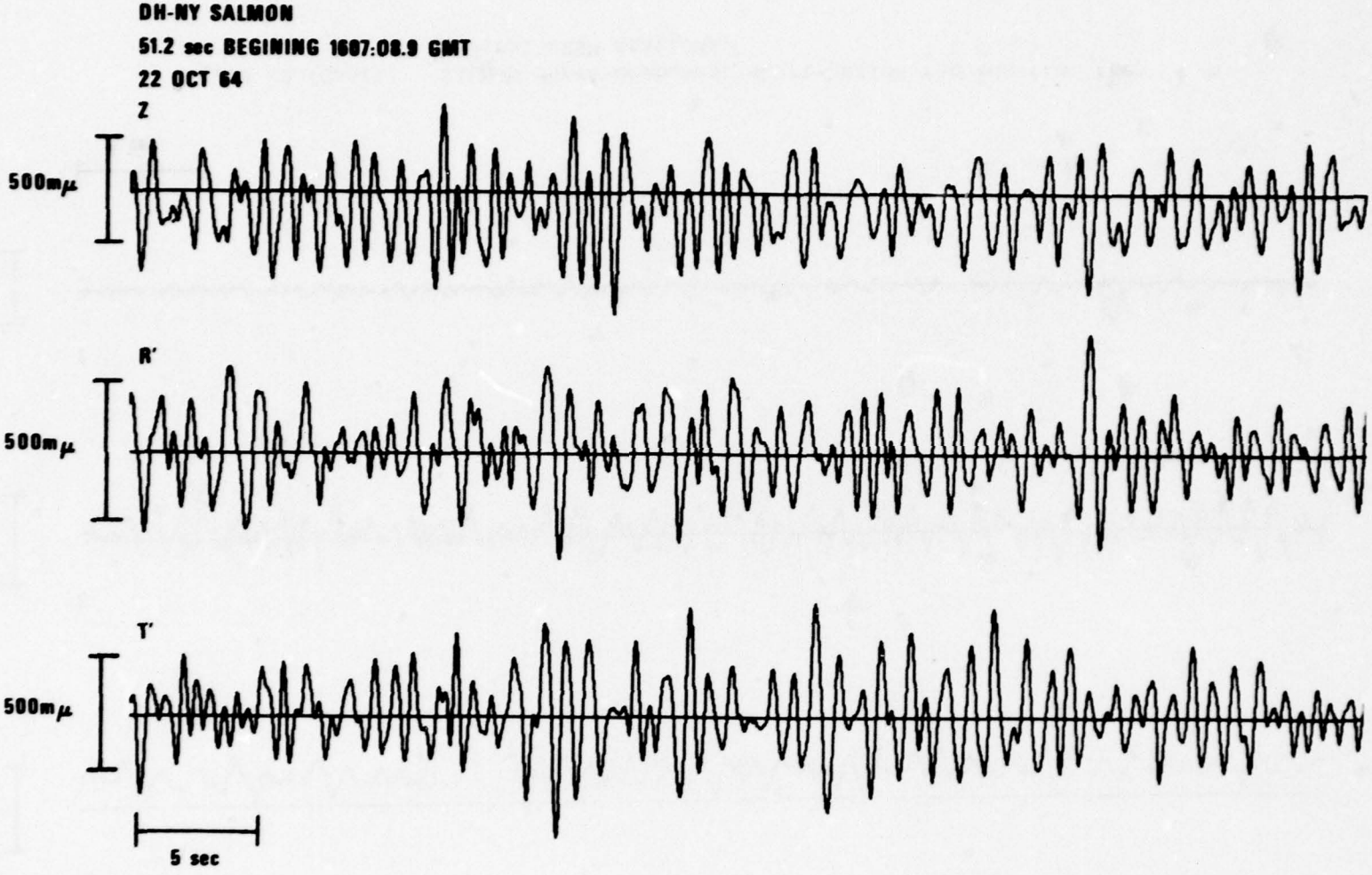


Figure 20 (Cont.) SALMON three-component short-period records from the several LRSM stations.

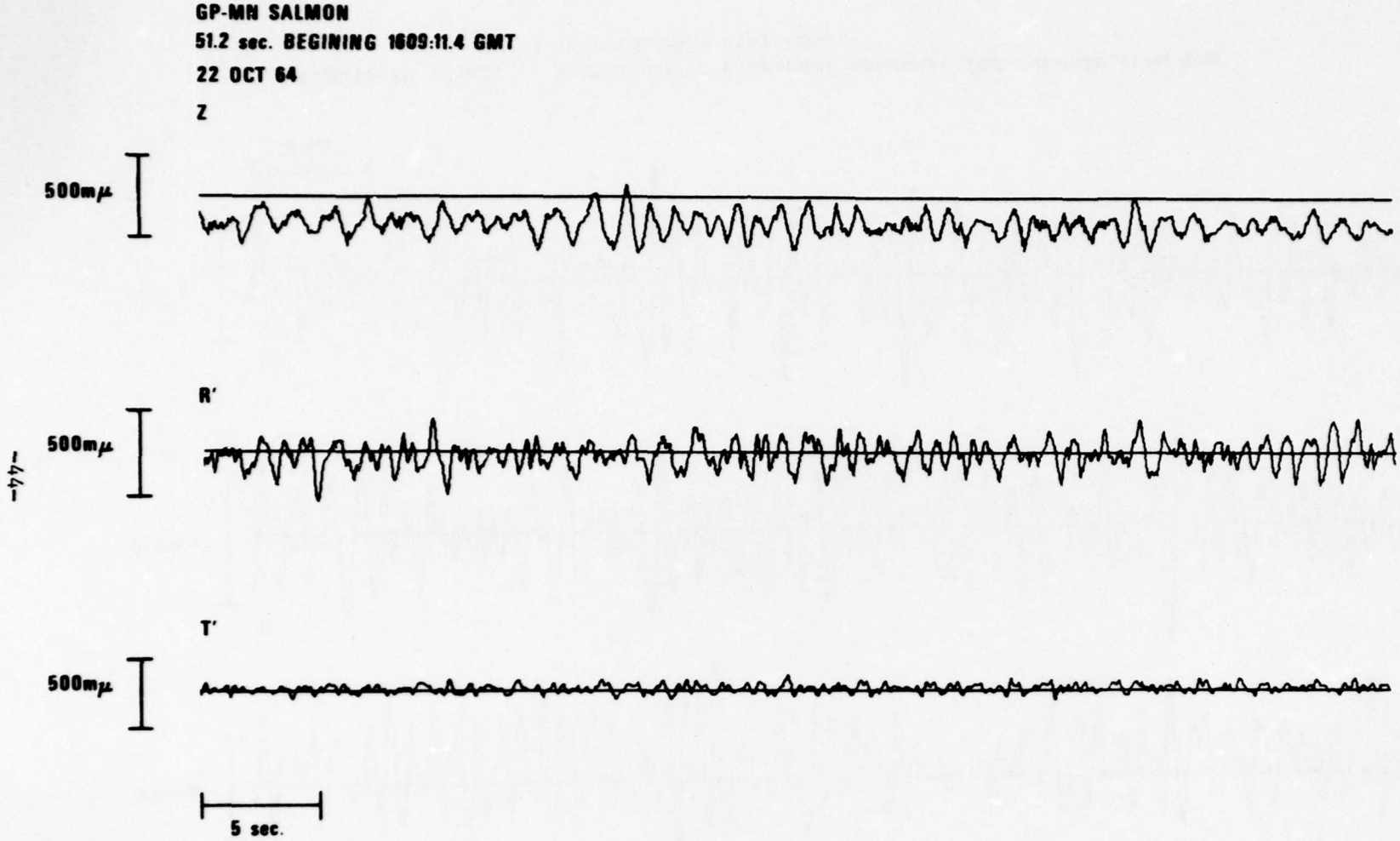


Figure 20 (Cont.) SALMON three-component short-period records from the several LRSM stations.

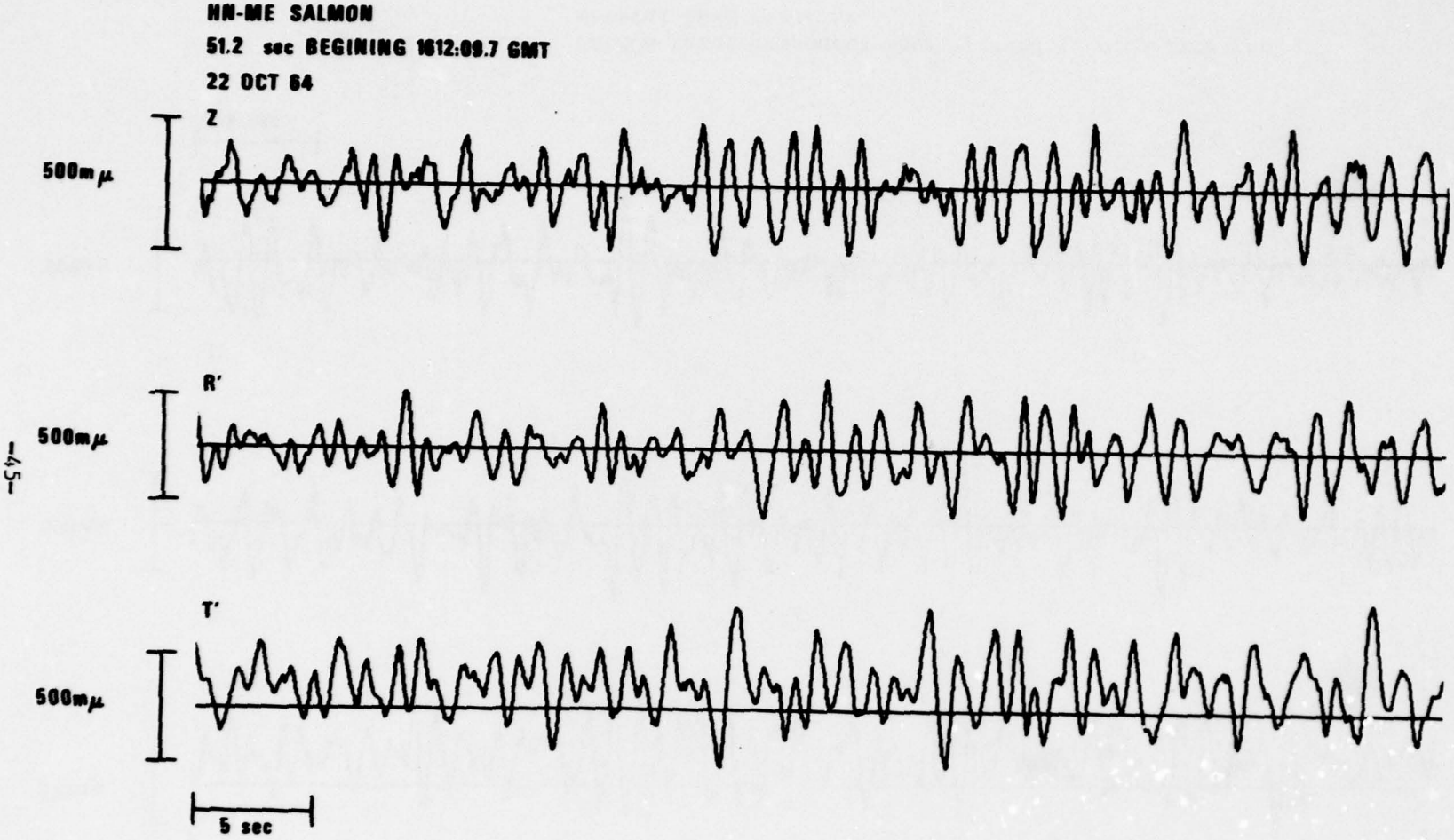


Figure 20 (Cont.) SALMON three-component short-period records from the several LRSM stations.

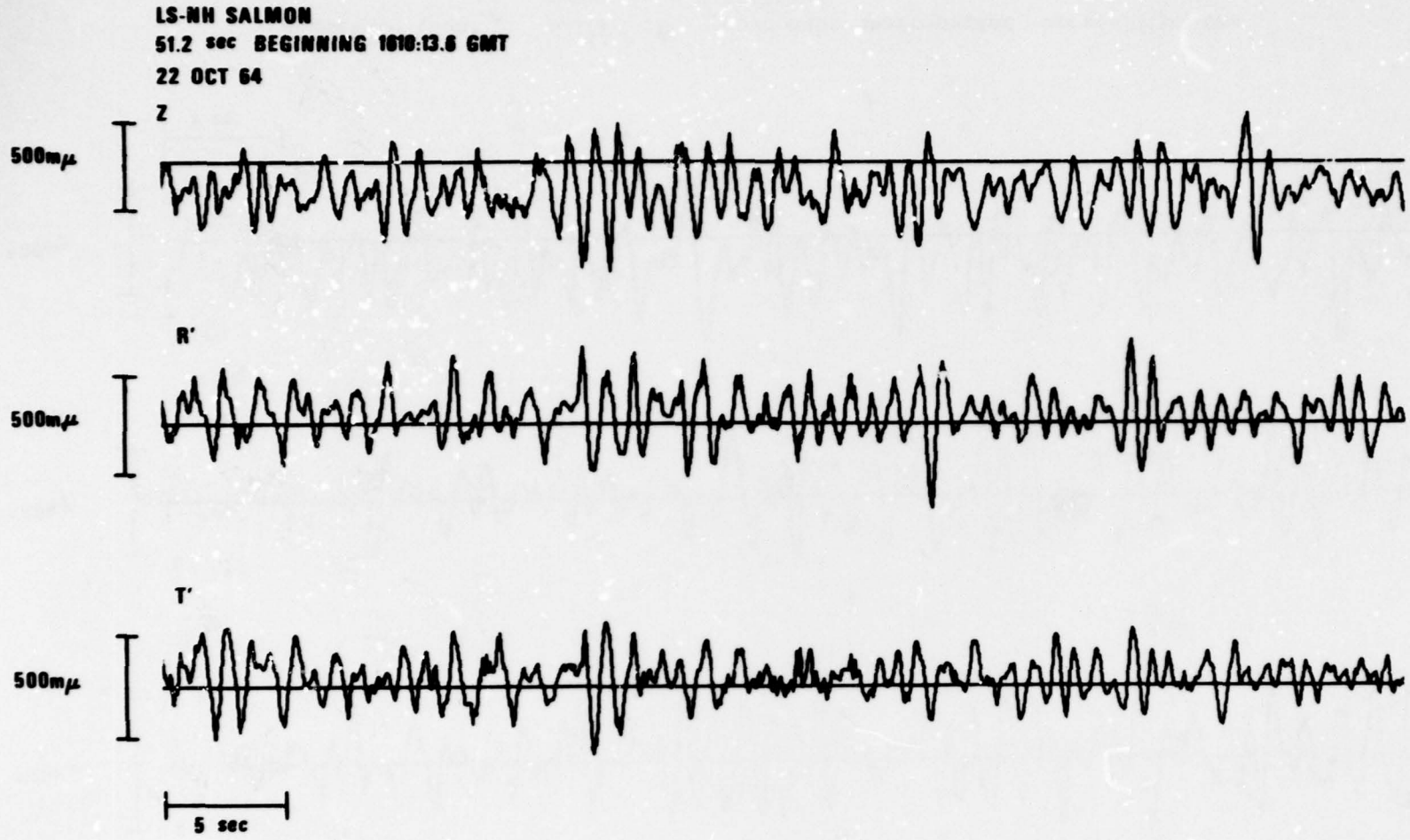


Figure 20 (Cont.) SALMON three-component short-period records from the several LRSM stations.

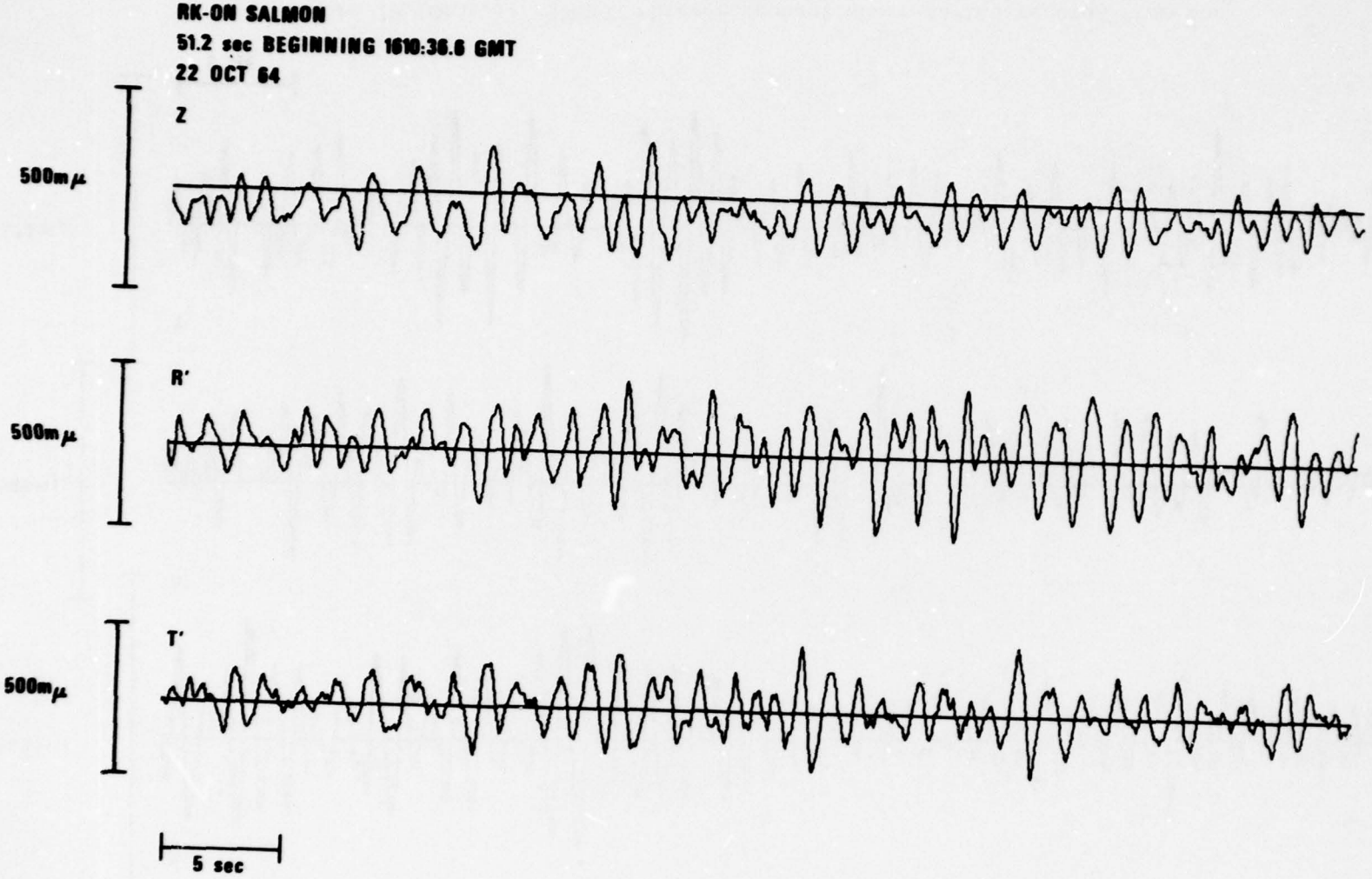


Figure 20 (Cont.) SALMON three-component short-period records from the several LRSM stations.

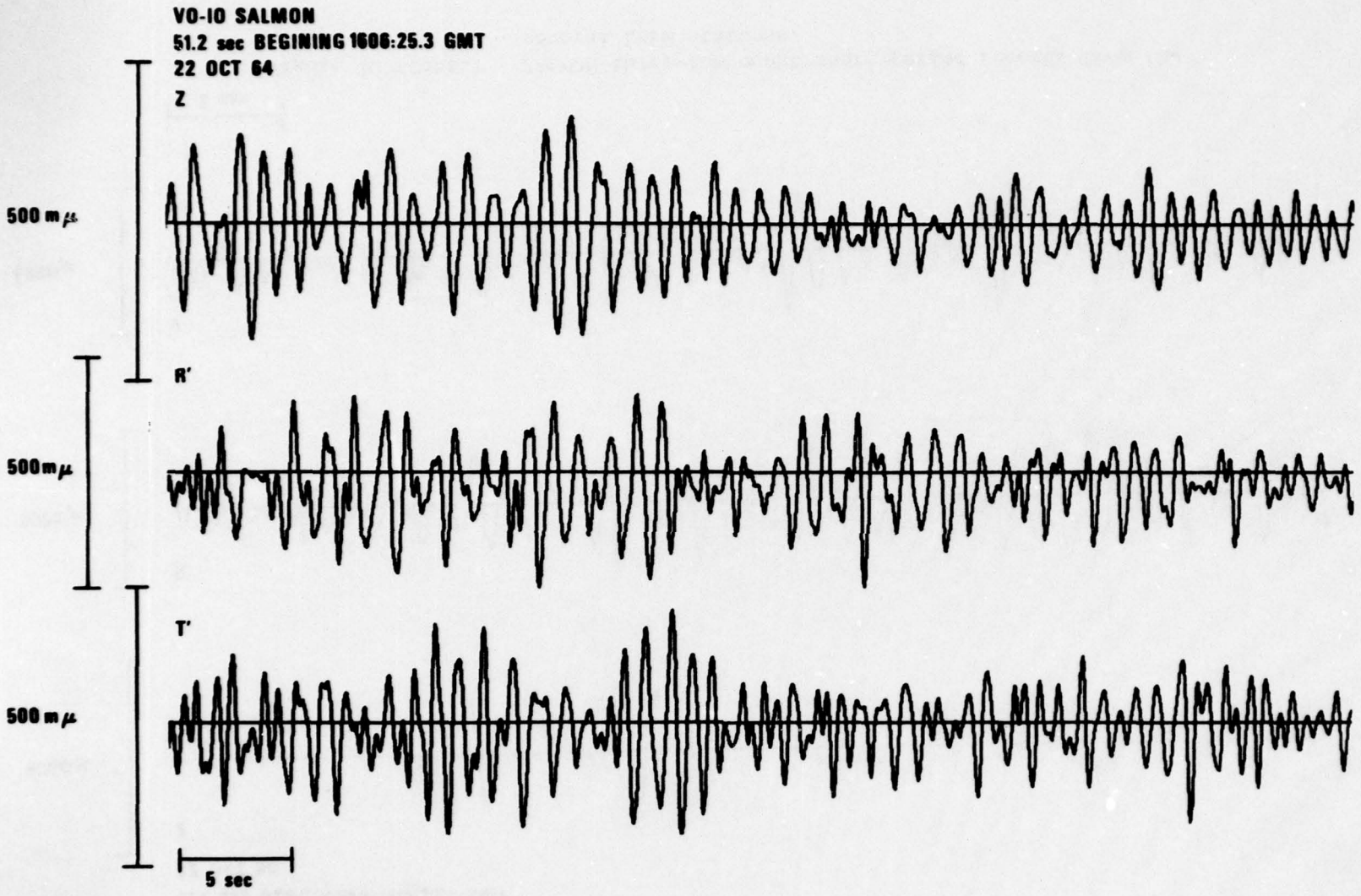


Figure 20 (Cont.) SALMON three-component short-period records from the several LRSM stations.

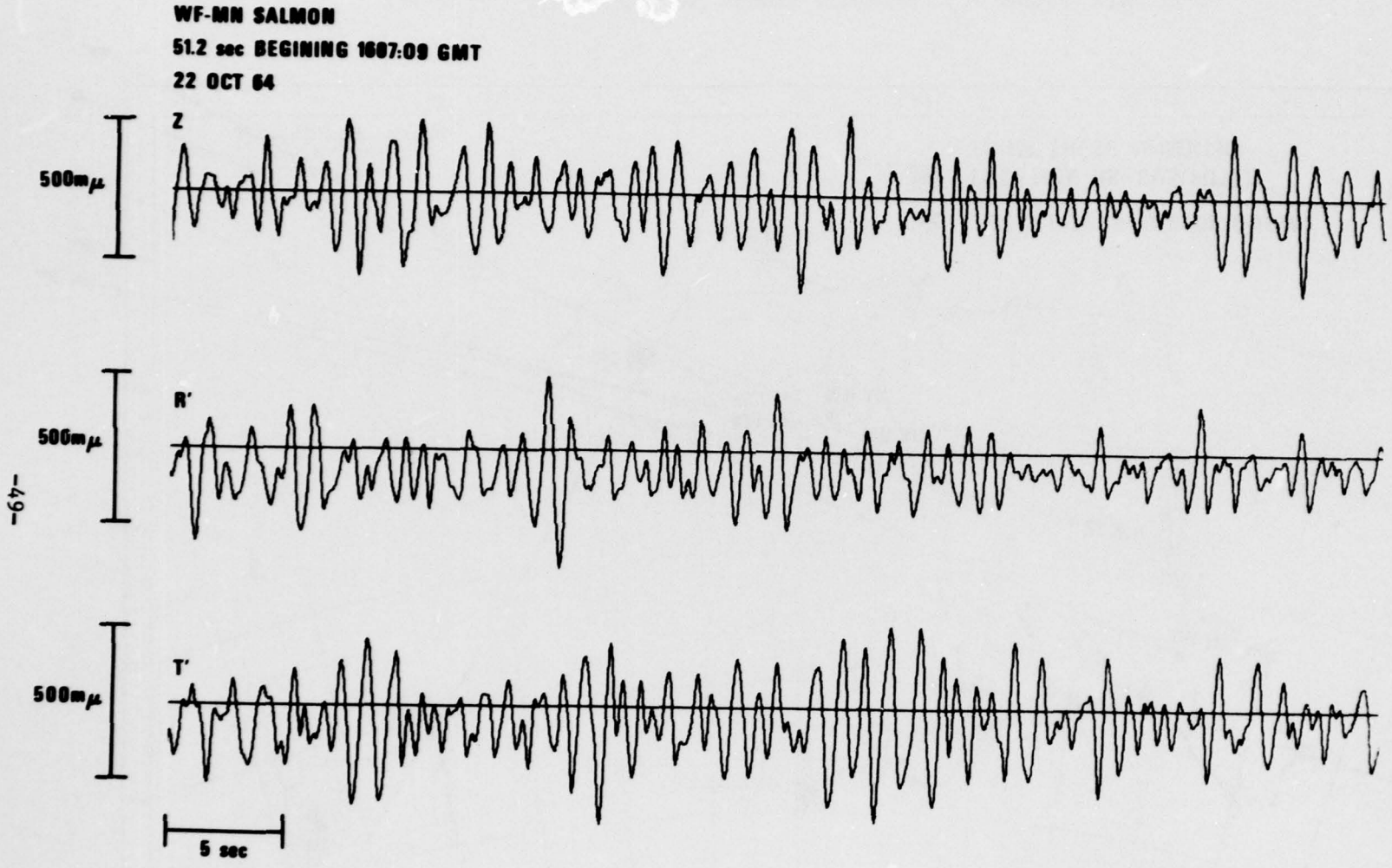


Figure 20 (Cont.) SALMON three-component short-period records from the several LRSM stations.

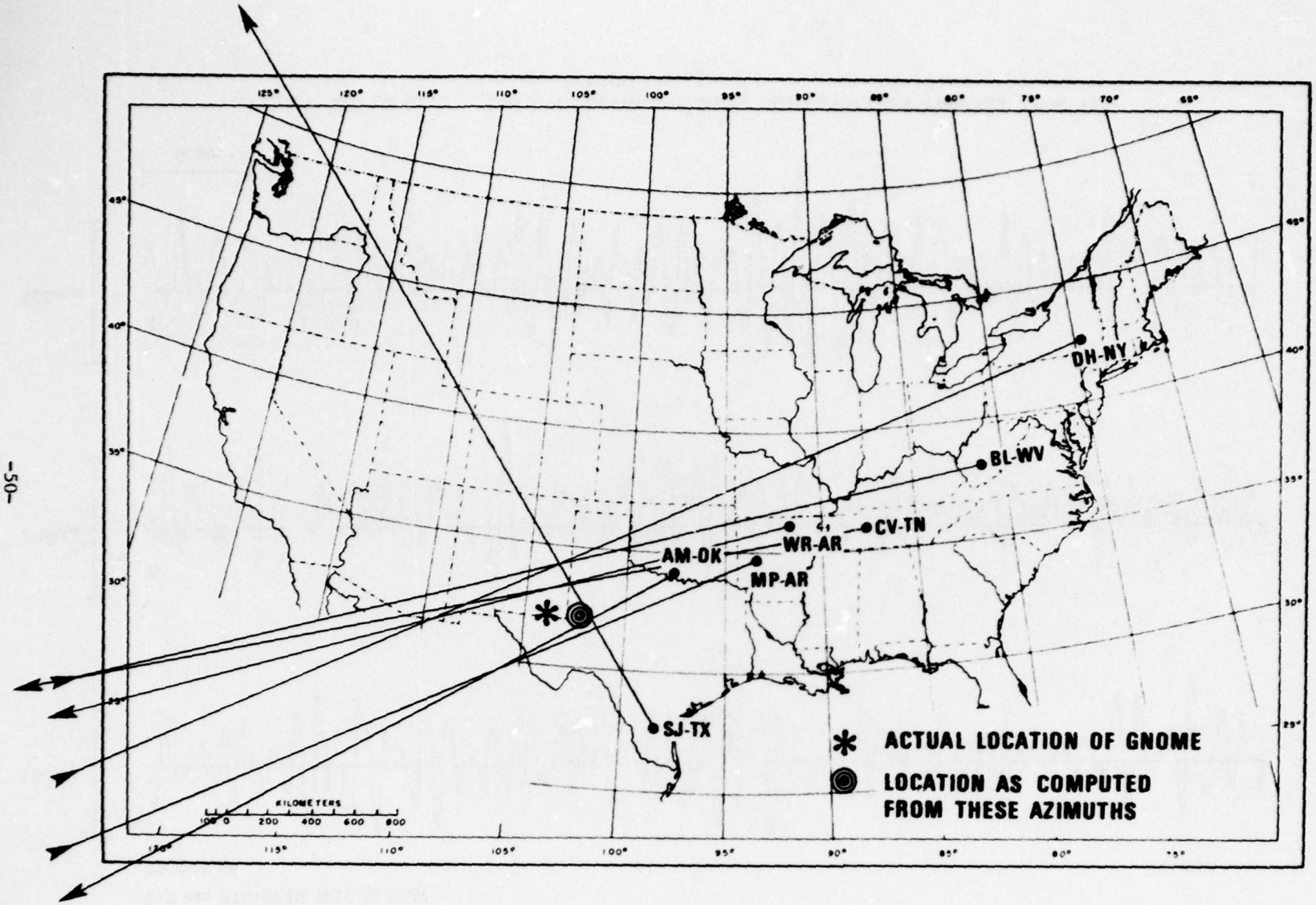


Figure 21 Automatic signal azimuth estimates of Lg surface waves from GNOME as seen at several LRSM stations.

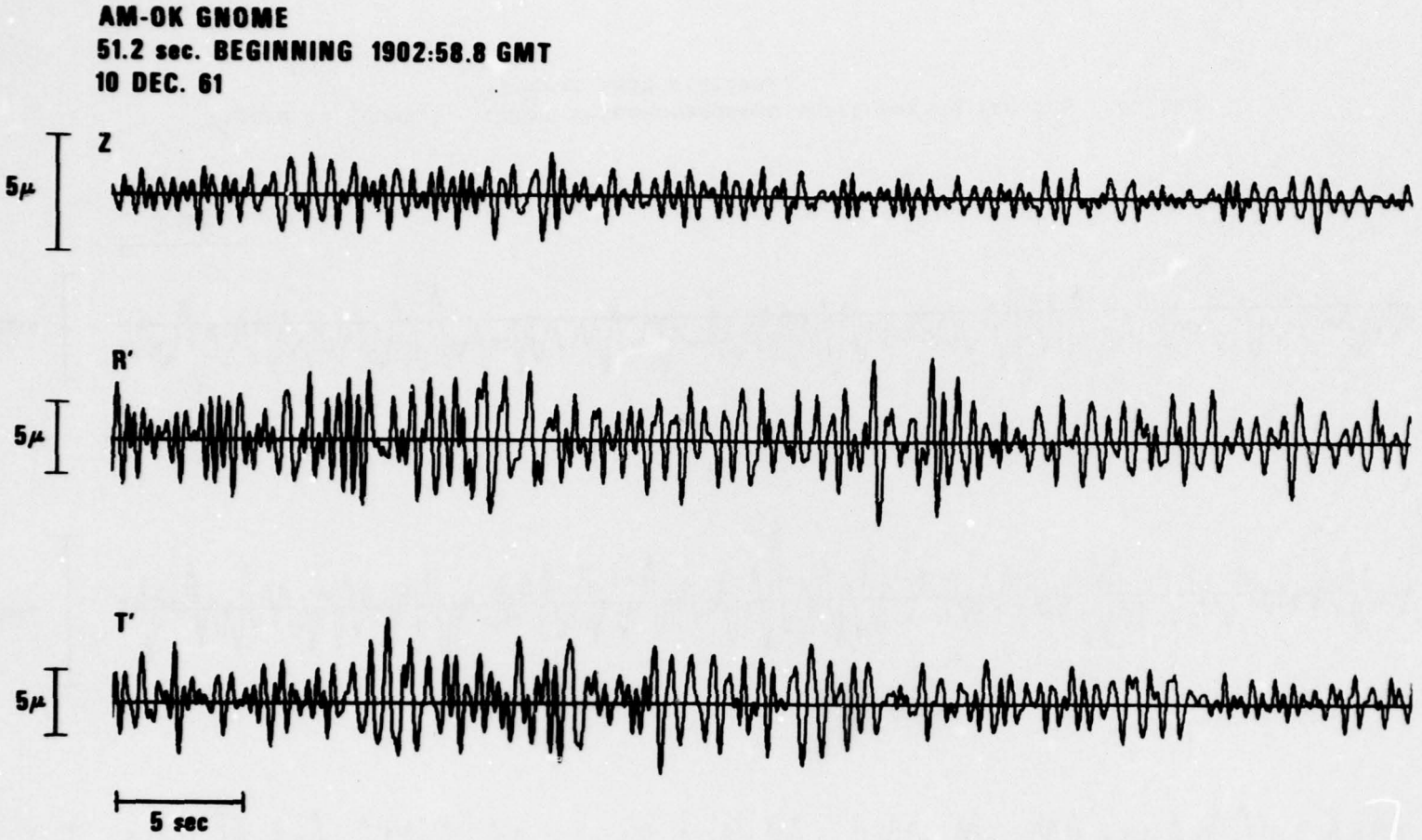


Figure 22 GNOME three-component short-period records from the several LRSM stations.

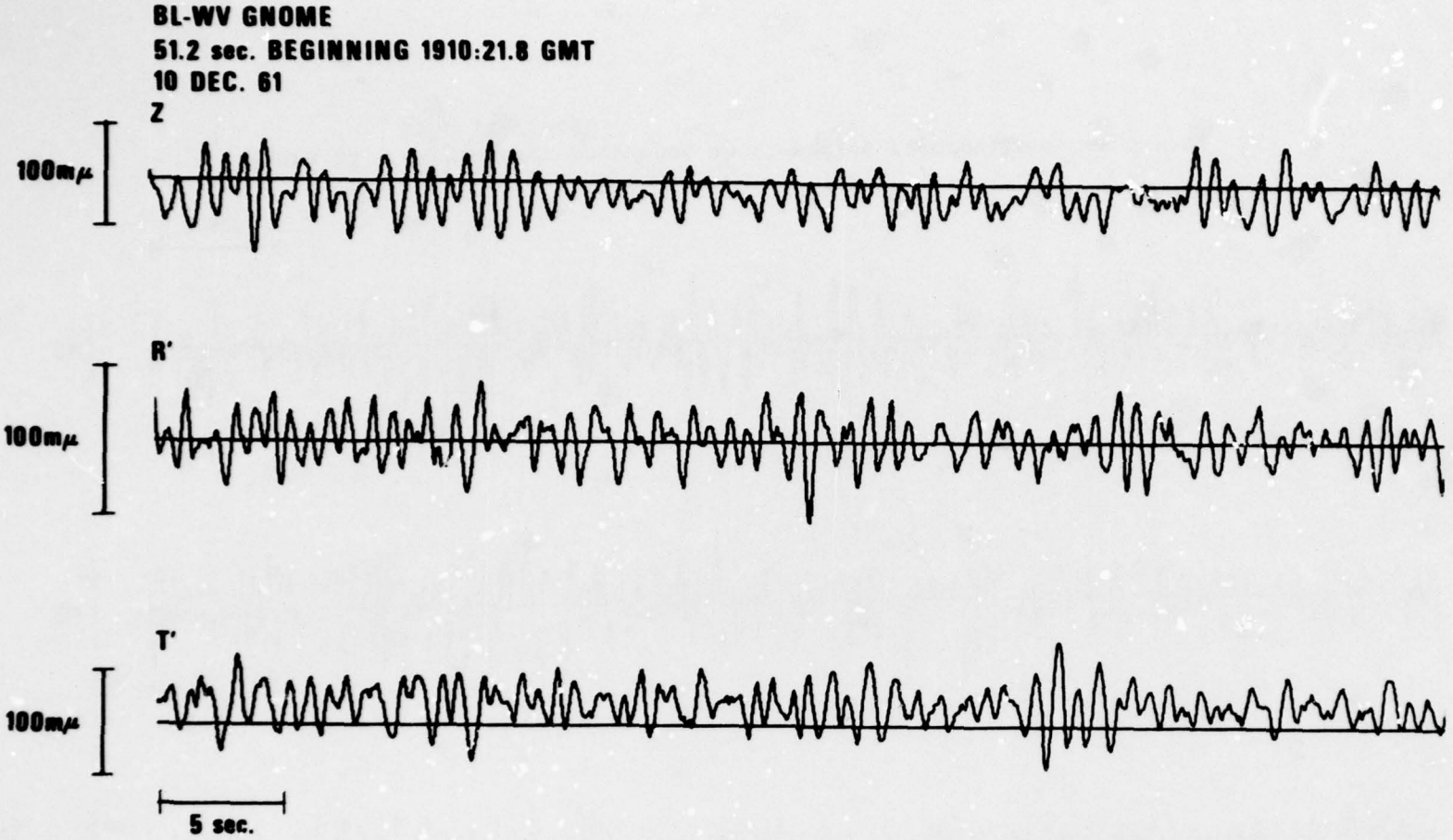


Figure 22 (Cont.) GNOME three-component short-period records from the several LRSM stations.

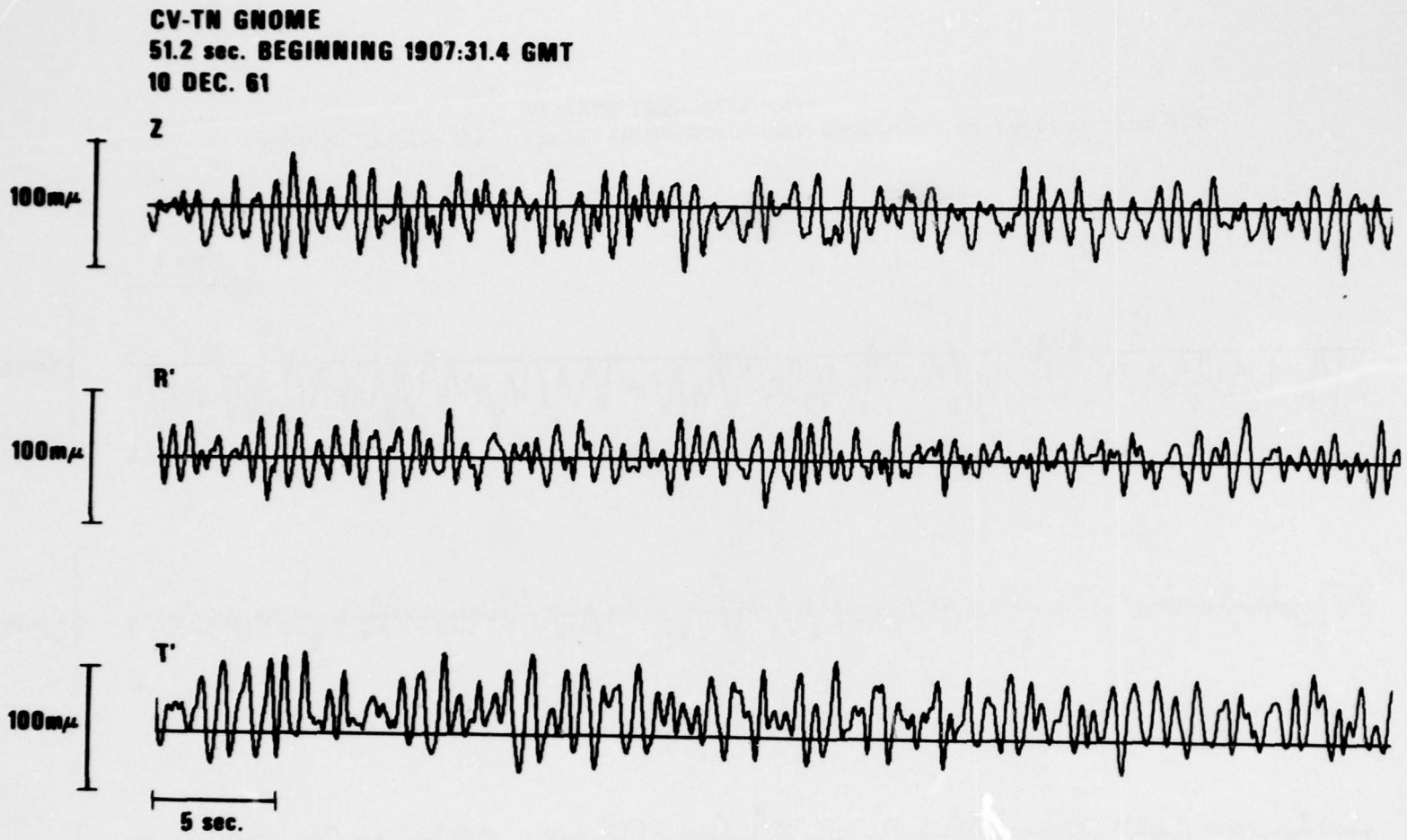


Figure 22 (Cont.) GNOME three-component short-period records from the several LRSM stations.

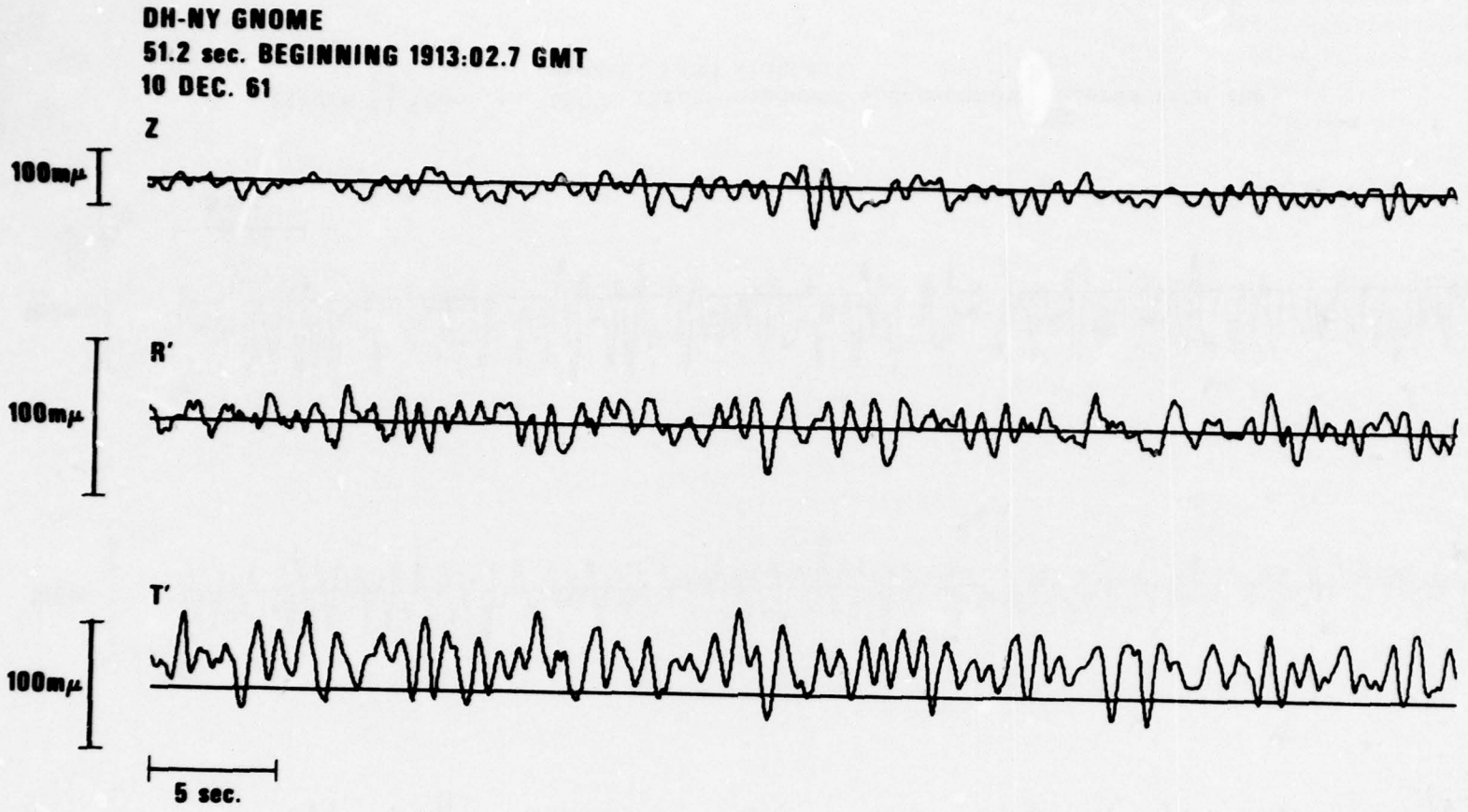


Figure 22 (Cont.) GNOME three-component short-period records from the several LRSM stations.

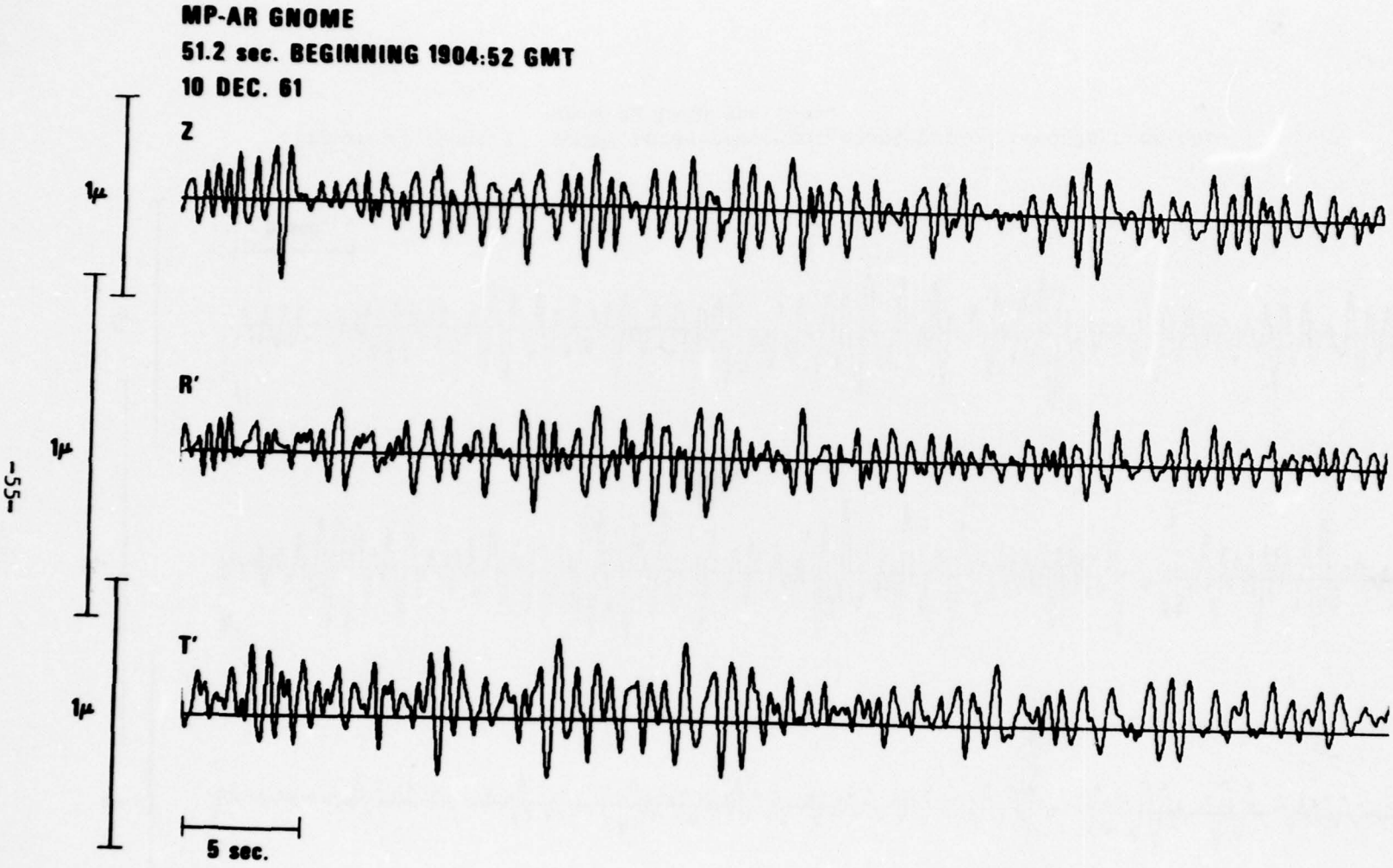


Figure 22 (Cont.) GNOME three-component short-period records from the several LRSM stations.

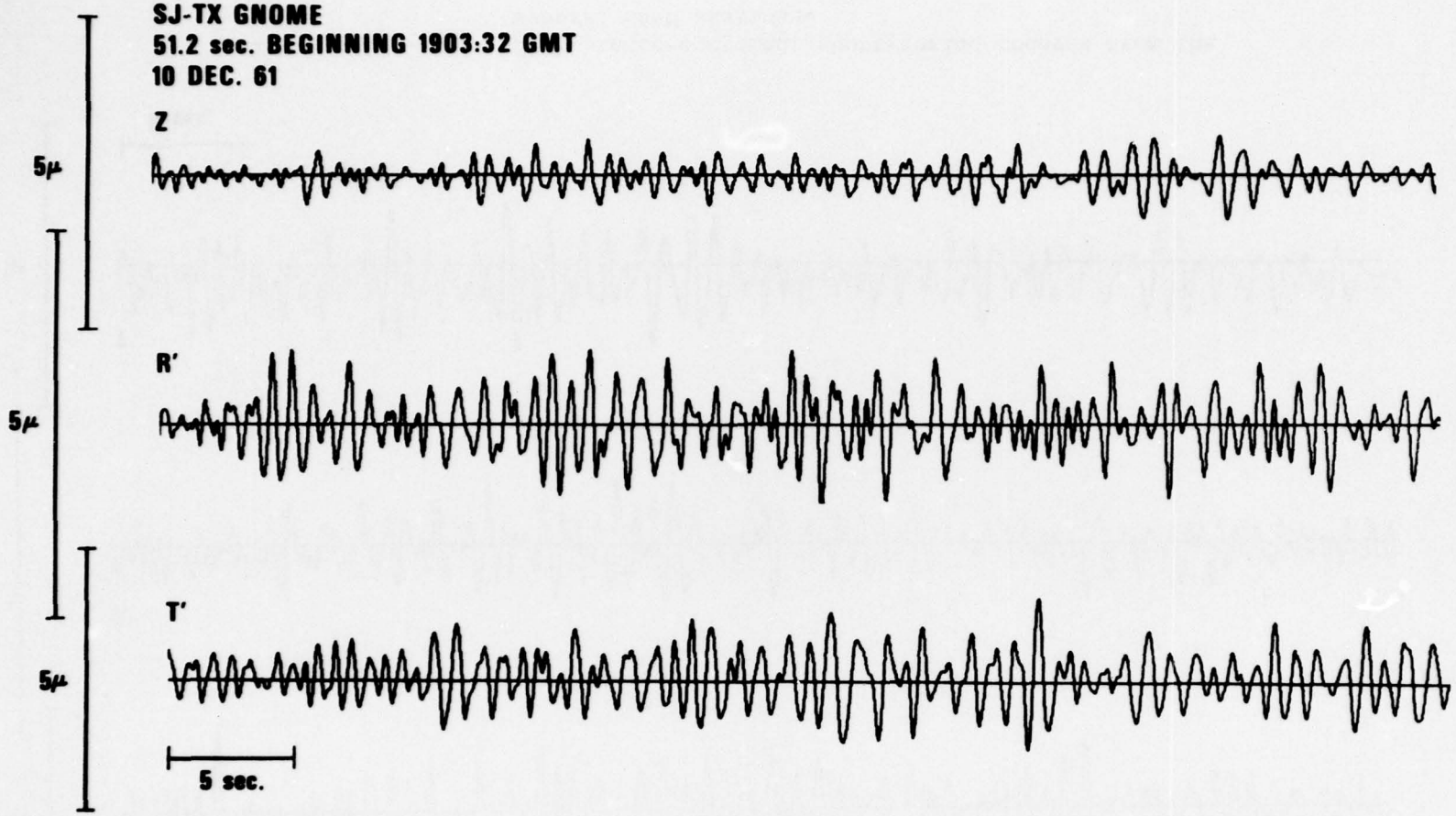


Figure 22 (Cont.) GNOME three-component short-period records from the several LRSM stations.

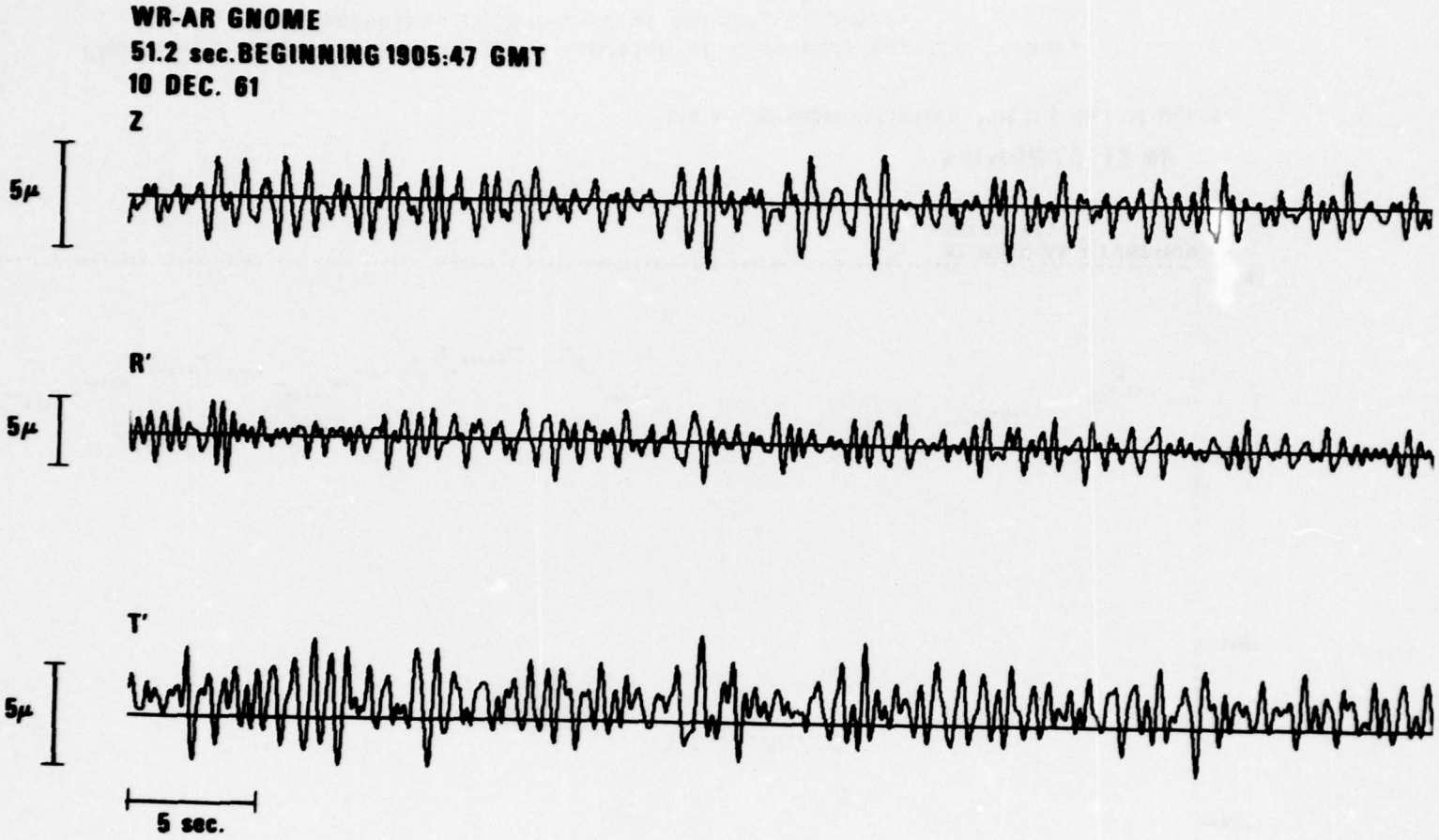


Figure 22 (Cont.) GNOME three-component short-period records from the several LRSM stations.

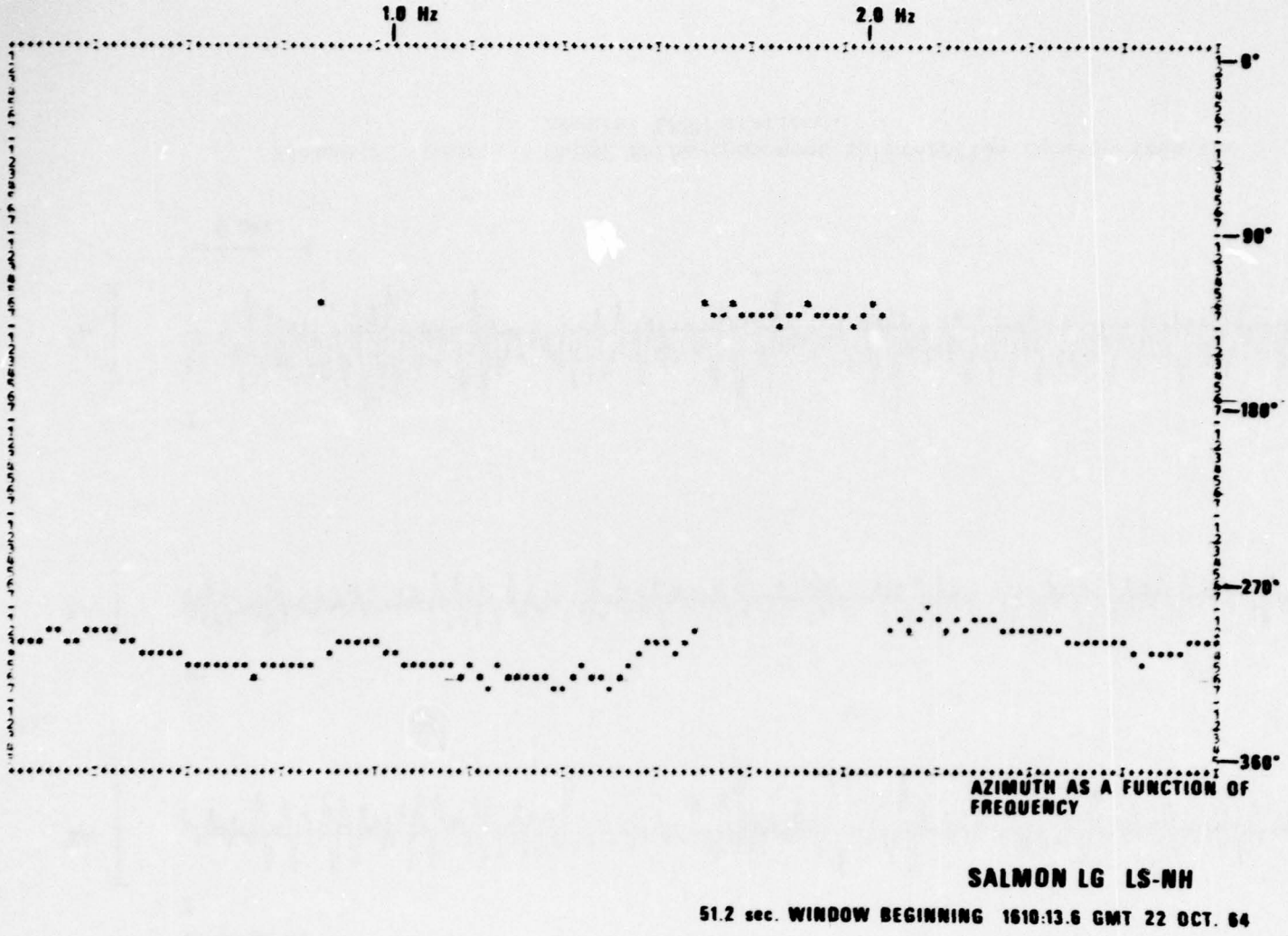


Figure 23. Apparent azimuth as a function of frequency for the Lisbon, New Hampshire, LRSM record of SALMON, L_g phase.

as may be seen in the figures. Figures 20 and 22 present the waveforms of those SALMON and GNOME records.

Estimation of apparent azimuth versus frequency for one such station (Figure 23) provides an interesting demonstration of the alternation from retrograde to prograde particle motion as frequency varies. Since retrograde motion is assumed in the model, the shift to prograde produces an 180-degree step in apparent azimuth. (The set of 3 sensors in this case was oriented not toward the north but rather away from the Nevada Test Site. To correct the azimuths in the figure add 94 degress.)

CONCLUSION

The detection processor described in this report has been shown to be a potentially powerful high-speed tool for automatically reviewing and/or editing three-component single-station data on a routine basis, with a performance competitive with that of an analyst.

Because it is sensitive to the ellipticity ratio of Rayleigh waves, the detector algorithm provides a new means for measuring that ratio, and thus it may offer the means for inferring the structure under a given recording station.

ACKNOWLEDGEMENTS

The author acknowledges valuable conversation with Dr. R. Blandford, Dr. Z. Der and Dr. R. Shumway.

REFERENCES

Barnard, T. E., and L. J. O'Brien, 1974. An evaluation of adaptive-beam-forming techniques applied to seismic data, ALEX (01)-TR-74-08, Texas Instruments, Inc., Dallas, Texas.

Brune, J. N., 1969. Surface Waves and Crustal Structure, The Earth's Crust and Upper Mantle, Geophysical Monograph 13, AGU, Washington, D. C.

Capon, J., R. J. Greenfield, and R. T. Lacoss, 1969. Long period signal processing results for the Large Aperture Seismic Array, Geophysics, 34, 305-329.

Choy, G., and K. McCamy, 1973. Enhancement of long-period signals by time varying adaptive filters, J. Geophys. Res., 78, 3505-3511.

Simons, R., 1968. PHILTRE-A surface wave particle motion discrimination process, Bull. Seism. Soc. Am., 58, 629-637.

Smart, E., 1974. Azimuthal distribution analysis, unpublished.

Smart, E., 1977. Automatic back-azimuth and spectral estimates of SALMON LG surface-waves recorded at lone 3-component LRSM installations, SDAC-TR-77- , Teledyne Geotech, Alexandria, Virginia.

Sobel, P. A., and D. H. von Seggern, 1975. Performance of the Philtre Processor at Low Signal-to-Noise Ratios, SDAC-TR-75-6, Teledyne Geotech, Alexandria, Virginia.

von Seggern, D. H., 1977. Methods of automating routine analysis tasks in preparing a global seismic bulletin, SDAC-TR-77-13, Teledyne Geotech, Alexandria, Virginia.