The Annals of Statistics
1988, Vol. 16, No. 1, 1–54

THE 1983 WALD MEMORIAL LECTURES

SOME STATISTICAL METHODS FOR RANDOM PROCESS DATA FROM SEISMOLOGY AND NEUROPHYSIOLOGY[1]

BY DAVID R. BRILLINGER

University of California, Berkeley

To Jeff Austin Brillinger, B.A.

Examples are presented of statistical techniques for the analysis of random process data and of their uses in the substantive fields of seismology and neurophysiology. The problems addressed include frequency estimation for decaying cosinusoids, signal estimation, association measurement, causal connection assessment, estimation of speed and direction and structural modeling. The techniques employed include complex demodulation, nonlinear regression, probit analysis, deconvolution, maximum likelihood, singular value decomposition, Fourier analysis and averaging.

Table of contents

Received December 1986; revised July 1987.

[1]Prepared with the partial support of National Science Foundation Grants CEE-79-01642, MCS-83-16634 and while the author was a Guggenheim Fellow.

AMS 1980 *subject classifications*. Primary 62M99, 62P99.

Key words and phrases. Array data, average evoked response, binary data, Fourier inference, complex demodulation, neurophysiology, point processes, probit analysis, spatial–temporal processes, seismology, system identification, time series.

I. Introduction

The purpose of statistics, ..., is to describe certain real phenomena.
A. Wald (1952)

The concern of these lectures is raw data distributed in time and/or space. The basic data are curves and surfaces. If n denotes the sample size and p denotes the dimension, then the concern is with the case of n much less than p. In the situations addressed, the phenomena have developed or are developing in time or space. They are complex, so that subject matter plays essential roles in the analyses made and in the interpretations and conclusions drawn. There need to be combinations of both physical and statistical reasoning. Indeed, a principal goal of the lectures is to bring out the key role that subject matter plays in the analysis of random process data. A further intention is to show that the fields of seismology and neurophysiology are rich in problems for statisticians, particularly those individuals with some interest in applied mathematics. The work presented involves a mixture of data analysis and structural modeling. The problems discussed are specific, but the techniques employed are broadly applicable. The data concerned is of high quality, so that detailed analyses are possible. The material presented consists of personal (collaborative) work and a few success stories of other particular methods that serve to tie the development together. An attempt is made to present problems from a unified point of view. Emphasis is on techniques, rather than novel substantive results.

The study of random process data provides a major interface of statistics with science and technology. Indeed, there has been an explosion in the collection of spatial–temporal measurements (corresponding in part to much of modern technology having become digital). Some particular issues and procedures become emphasized as a result of the interaction of statistics with technology. These include system identification, systems analysis, inverse problems, Fourier inference, bias versus variability (resolution versus precision), averaging functions, dynamics and micro-versus-macro study. These strains run through the examples presented. There is also a desire to display the broad range of data types with whose analysis statisticians must now be concerned.

Some provisos are necessary. There is no claim made that the analyses are definitive. What is presented is an overview, rather than specific details. Furthermore, there is little presentation of formalism. The reader is referred to the papers referenced for greater detail.

There are two lectures. The first concentrates on some statistical methods in seismology, the second on some corresponding methods in neurophysiology. It is interesting to see the same methods playing central roles in the analysis of data from two quite disparate fields. Indeed, one of the principal goals of the lectures was to bring out the universality of statistical techniques—by examples from these two fields.

II. Seismology

Jeffreys... attention to scientific method and statistical detail has been one of the main forces through which Seismology has attained its present level of precision.

Bullen and Bolt (1985)

1. The field and its goals. The term seismology refers to the scientific investigation of earthquakes and related phenomena. It has been defined as the "science based on data called seismograms, which are records of mechanical vibrations of the Earth" [Aki and Richards (1980)]. This latter definition allows the admission that seismologists also study vibrations caused by the sea, by volcanoes or by man. One further definition that has been given is: the science of strain–wave propagation in the Earth.

Whatever the definition, the broad goals of seismology are to learn the Earth's and a planet's interior composition and to predict the time, size, location and strength of ground motion in future earthquakes. Workers in the field seek to provide valid explanations of earthquake-related phenomena and to understand these phenomena so that life may be made safer.

Specific problems addressed include the detection, location and quantification of earthquakes, the distinguishing of earthquakes from nuclear explosions and the determination of wave velocity in the Earth's interior as a function of depth.

The accumulation of knowledge in seismology has displayed a steady back-and-forth between new insight concerning the waves and new insight concerning the media through which the waves propagate. Among major "discoveries" one can list are the inner core, the liquid central core, the Mohorovic discontinuity, the movement of tectonic plates causing earthquakes themselves and the locating of numerous gas and oil fields.

The field is largely observational, with the basic instruments the seismogram and clock. There are important experiments too, where tailored impulses are input to the Earth and the resulting vibrations studied. The field experienced the "digital revolution" in the 1950s and now poses problems exceeding the capabilities of even today's supercomputers.

Statistical methods have played an important role in seismology for many years—in large part due to the efforts of Harold Jeffreys [see Jeffreys (1977), for example]. Vere-Jones and Smith (1981) provide a review of many contemporary instances. Statistics enters for a variety of reasons. The data sets are massive. There is substantial inherent variability and measurement error. Models need to be refined, fitted and revised. Inverse problems need to be addressed. Experiments need to be designed. Sometimes the researcher must fall back on simulations. The basic quantity of concern is often a (risk) probability. In particular, it may be pointed out, that in the construction of the Jeffreys and Bullen (1940) travel time tables, one has an early, perhaps greatest success, of the use of robust/resistant methods. [B. A. Bolt's (1976) presidential address "Abnormal seismology" is well worth reading in this connection.]

Seismologists deal with data of a variety of types. The important forms are digital waveforms from spatial arrays of seismometers of various dimensions

(where the instruments have been arranged in such a fashion that an earthquake signal may be seen as a moving, changing entity) and catalogs (containing lists of an event's times, locations, sizes and other characteristics) for geographic regions of interest.

Seismology is not without its controversies. There are fundamental ones, such as whether or not plate tectonics is a validated theory. There are practical ones, such as does the size of the motion of an earthquake increase steadily as one approaches the fault or does it level off? As is so often the case, the existing data and analysis methods prove inadequate to resolve these disputes conclusively.

A general reference that provides much of the pertinent seismological background is Bullen and Bolt (1985). We turn to a presentation of some specific problems and techniques.

2. Free oscillations of the Earth. This subject is one of the principal developments in seismology over the last 25 years. Whenever there is a great earthquake, the Earth vibrates for days afterwards. The seismogram then consists, approximately, of a sum of an infinite number of exponentially decaying cosinusoids plus noise; see expression (2). The frequencies of the cosinusoids and the corresponding rates of decay relate to the Earth's composition. Measured values may be used to make inferences about that composition. The techniques of complex demodulation, nonlinear regression and regularization may be employed in this connection. Some details on these techniques will follow.

As is the case with many natural systems, the vibratory motion of the Earth may be described by a system of equations of the form

$$\frac{d\mathbf{Y}(t)}{dt} = \mathbf{A}\mathbf{Y}(t) + \mathbf{X}(t), \tag{1}$$

with $\mathbf{X}(\cdot)$ a (vector-valued) input. In the case that the input is $\mathbf{b}\delta(t)$, with $\delta(\cdot)$ the Dirac delta function (corresponding to the earthquake shock) and initial conditions $\mathbf{Y}(0-) = 0$, the general solution of (1) may be written as

$$\begin{aligned}\mathbf{Y}(t) &= \exp\{\mathbf{A}t\}\mathbf{b} \\ &= \sum_j \zeta_j \exp\{\mu_j t\}\mathbf{u}_j, \qquad t > 0,\end{aligned}$$

where $\mu_j, \mathbf{u}_j$ are the (assumed distinct) latents of the matrix $\mathbf{A}$. The spectrum occurring is discrete because of the finiteness of the Earth as a body. Focusing on one of the coordinates of $\mathbf{Y}(t)$ and assuming the presence of noise, one has

$$Y(t) = \sum_k \alpha_k \exp\{-\beta_k t\}\cos(\gamma_k t + \delta_k) + \varepsilon(t), \tag{2}$$

with $-\beta_k$ and γ_k the real and imaginary parts of the μ_j and $\varepsilon(\cdot)$ the noise. This model may be checked by complex demodulation of the series $Y(t)$ in the neighborhood of frequencies γ_k, as estimated from the periodogram. Provided the bandwidth of the demodulation is not too great, a single cosinusoid should be

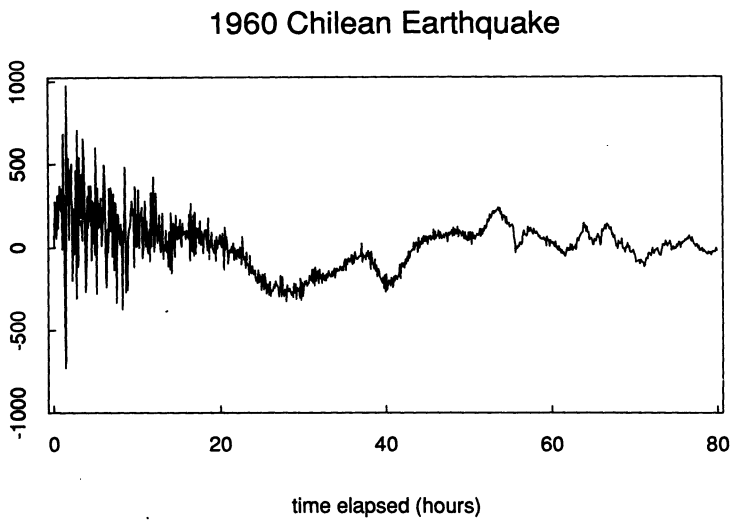

FIG. 1. *Record of the Chilean great earthquake of May* 22, 1960, *as recorded by the tiltmeter in the Grotta Gigante at Trieste. The tides have been partially removed.*

included, the log amplitude should fall off linearly with time and the phase angle should be approximately constant. Details are given later, specifically at (5).

Figure 1 is a plot of the seismogram recorded at Trieste of the 1960 Chilean great earthquake after partially removing the tides. Details re the data and the tidal removal procedure may be found in Bolt and Marussi (1962). Figure 2, a plot of the lower-frequency portion of the periodogram of this data, suggests the

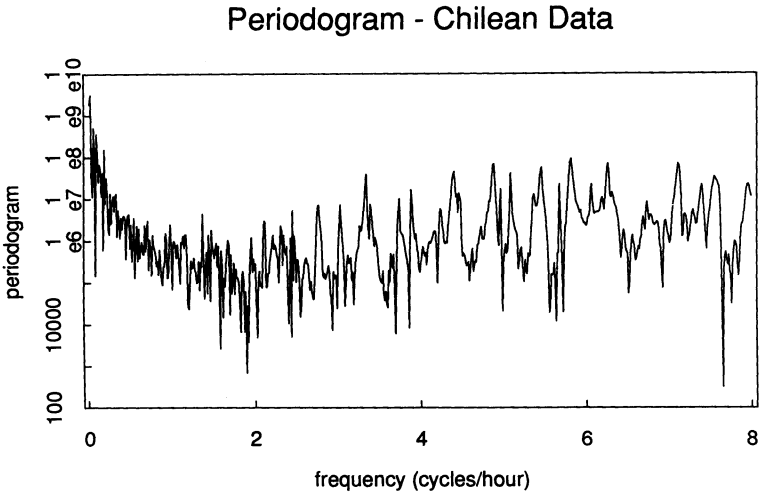

FIG. 2. *The periodogram of the data of Figure* 1 *based on* 2548 *data values. Only ordinates corresponding to frequencies less than* 8 *cycles/h have been graphed. The y-axis is logarithmic.*

presence of a variety of periodic components. The periodogram of a stretch of time-series values $Y(t)$, $t = 0, \ldots, T-1$, is defined as follows. Set

$$d_Y^T(\lambda) = \sum_{t=0}^{T-1} Y(t)\exp\{-i\lambda t\}, \qquad -\infty < \lambda < \infty. \tag{3}$$

Then the periodogram at frequency λ is defined as

$$I_{YY}^T(\lambda) = (2\pi T)^{-1}|d_Y^T(\lambda)|^2. \tag{4}$$

For data from the model (2), $I_{YY}^T(\lambda)$ may be expected to show peaks for λ near the γ_k.

The basic ideas of complex demodulation are frequency isolation by narrow-band filtering to focus on a single term in expression (2), followed by frequency translation to slow the oscillations down. The specific steps are: (i) $Y(t) \to Y(t)\exp\{i\lambda t\}$ (modulation), followed by (ii) local smoothing in t of $Y(t)\exp\{i\lambda t\}$ to obtain $Y(t, \lambda)$, the complex demodulate at frequency λ. In the case that $Y(t) = \alpha\exp\{-\beta t\}\cos(\gamma t + \delta)$, one has

$$\begin{aligned} Y(t, \lambda) &\approx \tfrac{1}{2}\alpha e^{i\delta} e^{-\beta t} e^{i(\lambda-\gamma)t}, \quad \text{for } \lambda \text{ near } \gamma \\ &\approx 0, \quad \text{otherwise.} \end{aligned} \tag{5}$$

Hence $\log|Y(t, \lambda)| \approx \log(\alpha/2) - \beta t$ and $\arg\{Y(t, \lambda)\} \approx \delta + (\lambda - \gamma)t$. Plots of these quantities versus t provide checks on model adequacy and provide preliminary estimates of parameters. Figures 3 and 4 present such plots for the Chilean data at two frequencies, 3.885 and 5.6775 cycles/h. These frequencies were determined by noticing the locations of peaks in the periodogram, setting λ equal to them, demodulating and then in some cases employing a nearby λ to get a more nearly horizontal phase plot. The fluctuations in the amplitude plot can be due to noise, to leakage from other frequency components or to split peaks among other things. The rate of decay β is found to generally vary with frequency in the present seismological situation. Results for the Chilean data for a variety of frequencies may be found in Bolt and Brillinger (1979).

The parameters could be estimated from the complex demodulate pictures, for example, by fitting regression lines. It is generally more effective to proceed via nonlinear regression. This has the further advantage of providing estimated standard errors. Suppose one has a model

$$Y(t) = S(t; \theta) + \varepsilon(t),$$

with $S(\cdot)$ known up to the finite-dimensional parameter θ and $\varepsilon(\cdot)$ a noise series. In the present case, $S(t) = \alpha\exp\{-\beta t\}\cos(\gamma t + \delta)$ and $\theta = \{\alpha, \beta, \gamma, \delta\}$. For the next step, it is convenient to take $\lambda_j = 2\pi j/T$ and to write $Y_j = d_Y^T(\lambda_j)$, $E_j = d_\varepsilon^T(\lambda_j)$ and $S_j(\theta) = d_S^T(\lambda_j)$. One will estimate θ by minimizing

$$\sum_{j \text{ in } J} |Y_j - S_j(\theta)|^2, \tag{6}$$

for J a range of subscripts with λ_j near γ. The logic of this is as follows. There are a variety of central limit theorems for empirical Fourier transforms [see, for example, Brillinger (1983)]. Suppose that the noise series $\varepsilon(\cdot)$ is stationary and

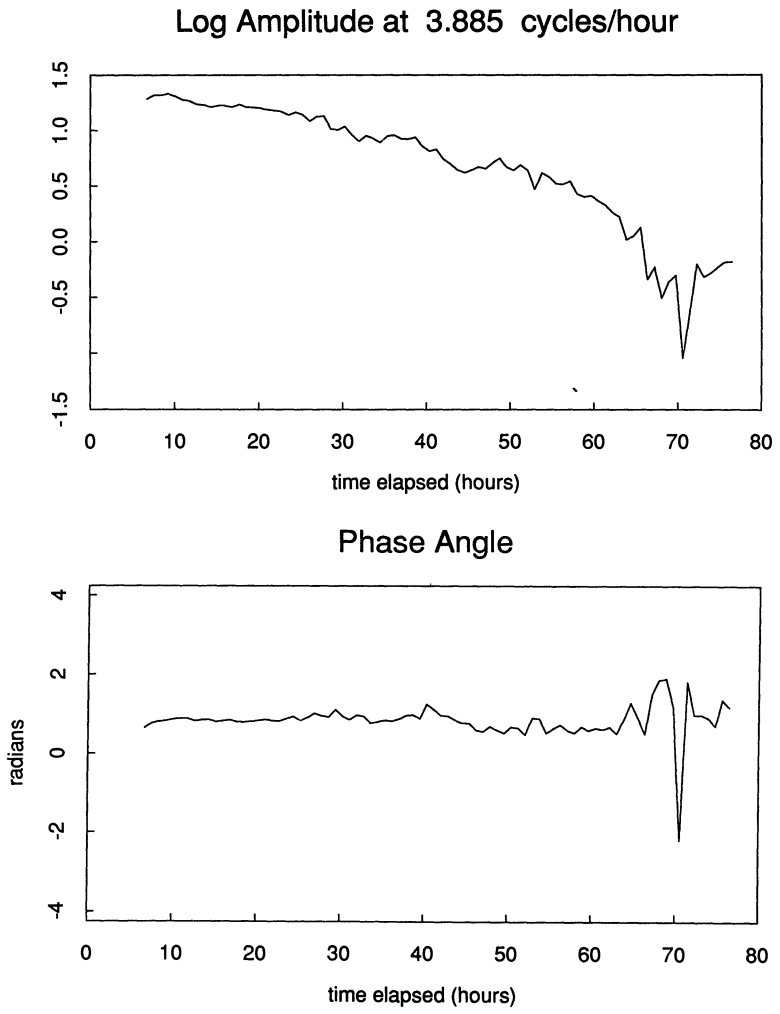

FIG. 3. *The result of complex demodulating the data of Figure* 1 *at a frequency of* 3.885 *cycles/h. The upper graph gives the logarithm of the running amplitude. The lower graph gives the running phase. The bandwidth of the filter employed is* 0.594 *cycles/h.*

mixing with power spectrum $f_{\varepsilon\varepsilon}(\lambda)$. Then for large T, E_j is approximately complex normal with mean 0 and variance $2\pi T f_{\varepsilon\varepsilon}(\lambda_j)$. Further the variates E_j, E_k are approximately independent. It follows that the determination of an estimate of θ to minimize expression (6) is approximately the maximum likelihood procedure. The statistical properties of such estimates were indicated in Bolt and Brillinger (1979) and developed in detail in Hasan (1982). For example, one finds the asymptotic variance of $\hat{\gamma}$ to be proportional to

$$\frac{4\pi f_{ee}(\gamma)}{T^3\alpha^2},$$

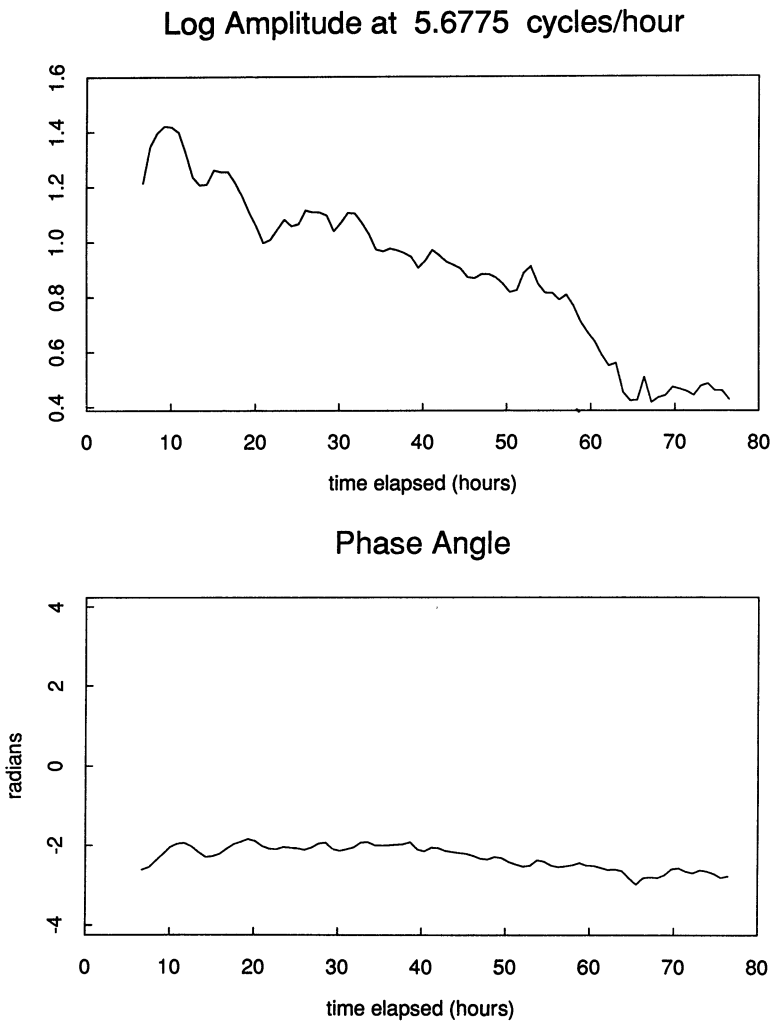

FIG. 4. *Complex demodulation as for Figure 3, but at the frequency of 5.6775 cycles/h.*

having considered a limiting process with $\beta = \phi/T$ as $T \to \infty$. The inverse cubic dependence on sample size is on first glance surprising. It comes from the narrowness of the peaks when they are present.

Complex demodulation is an exploratory technique. Hence one has to be conscious of the possibility of employing it at frequencies of "false" peaks. In practice, it is found that the nearness of the phase plot to a straight line is a highly sensitive indicator of the presence of a periodic component.

Earlier in the paper, it was noted that progress in seismology shows a to-and-fro between new knowledge of waves and new knowledge of the structure of the Earth. This occurs in the case of free oscillations. Suppose one has an initial model for the Earth in terms of some physical parameters, e.g., expres-

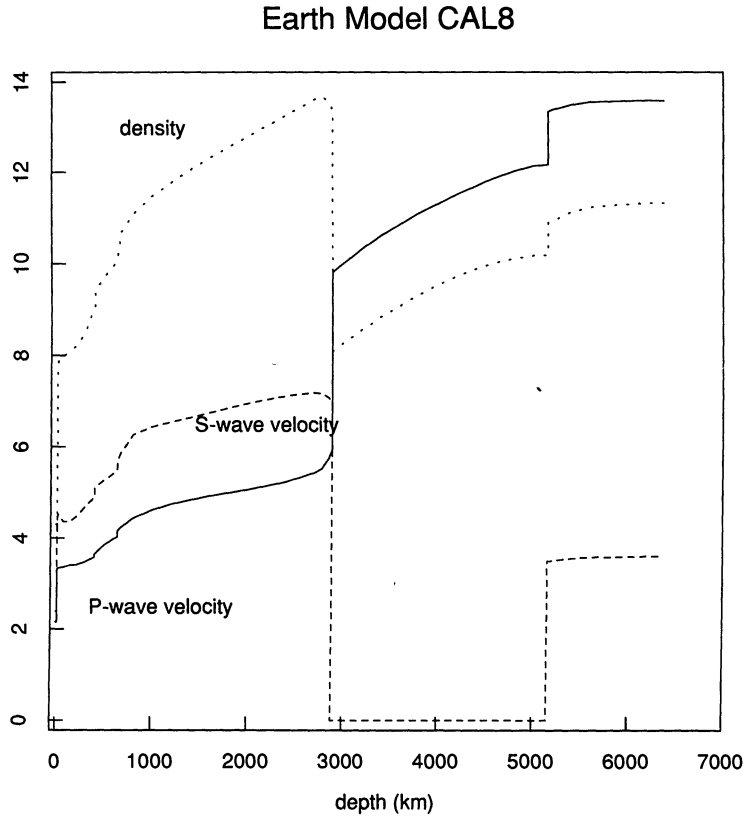

FIG. 5. *The* CAL 8 *Earth model. The curves give the assumed density (grams per cubic centimeter), P-wave velocity (kilometers per second) and S-wave velocity (kilometers per second) as a function of depth assuming a spherical Earth. Given such a model, one can compute implied periods of free oscillation. Of interest is the inverse problem, given periods what is a corresponding Earth model?*

sions for density, shear wave velocity and compression wave velocity as functions of depth, say $\rho(r)$, $c_S(r)$ and $c_P(r)$, respectively, r denoting depth. Figure 5, based on the data in Tables 3 and 4 of Bolt (1982), shows what is meant by an Earth model. Given such a model, one can compute the implied frequencies of free oscillation γ_k. How to do this is described in Chapter 6 of Lapwood and Usami (1981), for example. The relationship involved is nonlinear, but perturbations may be expressed linearly via kernels. Specifically, suppose one perturbs the parameters by amounts $\Delta\rho$, Δc_S and Δc_P, respectively, then the perturbation of the frequency of the kth free oscillation is given by

$$\Delta\gamma_k \approx \int_0^R A_k(r)\,\Delta\rho(r)\,dr + \int_0^R B_k(r)\,\Delta c_S(r)\,dr + \int_0^R C_k(r)\,\Delta c_P(r)\,dr,$$

for kernels A_k, B_k and C_k. This expression is said to lay out the "direct problem": Given $\Delta\rho$, Δc_S and Δc_P find $\Delta\gamma_k$. Now suppose a great earthquake

occurs. Then new estimates of the frequencies γ_k are available. One has the "inverse problem": Given the observed $\Delta\gamma_k$, find $\Delta\rho$, Δc_S and Δc_P. Because the new frequencies are just estimates, one seeks a model only approximately achieving them. It seems worth setting out the type of problem involved here in a specific notation. Let Y and Θ denote normed spaces. Let $\mathbf{X}$ denote a map from Θ to Y, $\mathbf{Y} = \mathbf{X}\theta$. The values $\mathbf{Y}$ and $\mathbf{X}$ are given, a value for θ is desired. Let α denote a scalar. Some, basically similar, methods for selecting a θ currently being employed include: (a) regularization, choose θ to minimize $\|\mathbf{Y} - \mathbf{X}\theta\|^2 + \alpha\|\theta\|^2$; (b) sieve, choose θ subject to $\|\theta\| \leq \alpha$ to minimize $\|\mathbf{Y} - \mathbf{X}\theta\|$; (c) residual, choose θ subject to $\|\mathbf{Y} - \mathbf{X}\theta\| \leq \alpha$ to minimize $\|\theta\|$. A characteristic of the solutions obtained is that one has to be content with the estimation of some form of average of the unknown θ. Chapter 12 of Aki and Richards (1980) contains a discussion of inverse problems in geophysics. A characteristic that distinguishes the present Earth model problem, from the usual inverse problems, is that there are discontinuities present in the model—corresponding to the Earth's layers. The above perturbation approach of a nonlinear problem to a linear one has been employed by geophysicists for many years; see Jeffreys and Bullen (1940), for example.

Several other references to the study of free oscillations may be noted. Hansen (1982) extends the procedure of Bolt and Brillinger (1979) to handle the case of several eigenfrequencies present in the nonlinear regression fit. Dahlen (1982) sets down the asymptotic results for the case of tapered data, that is, when convergence factors have been introduced into the Fourier transform computations. Zadro and Caputo (1968) look for nonlinearities via bispectral analysis.

3. Estimation of fault-plane parameters. That there exists a see-saw between the study of the Earth's structure and the study of earthquake sources was pointed out earlier. In this section it will be indicated how a (nonlinear) probit analysis may be employed to estimate basic characteristics of the source of an earthquake.

An important quantity read off the seismic trace of an earthquake at a particular observatory is the sign of the increment at the arrival of the first energy from the event. This sign corresponds to whether the initial motion is a compression or a dilation. In many cases, following the observation of an earthquake at a number of stations, if the observed signs of first motion are plotted on a map centered at the epicenter of the event a (radiation) pattern results. Figure 6, taken from Brillinger, Udias and Bolt (1980), provides such a plot for one of the aftershocks (event 4) of the Good Friday 1964 Alaskan event. (Unfortunately, due to the locations of the particular stations recording the event, this figure does not provide a particularly good example of the ideal radiation pattern, but the data were of special interest. Were the stations well scattered, in an ideal circumstance one would see mainly solid circles in two opposite quadrants and mainly open circles in the other two quadrants. In this case only two of the four quadrants have been covered. The implication will be that one of the planes will be poorly determined.) Following Byerly (1926), plots such as this have been employed to learn about the source. Before describing

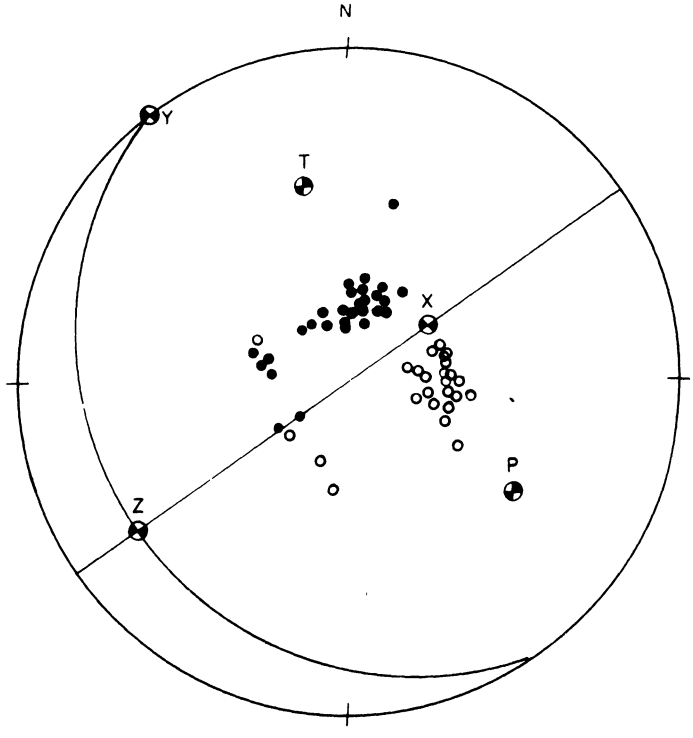

FIG. 6. *The P-wave first-motion data for the earthquake of the Alaskan sequence that took place March* 30, 1964 *at* 0200. *The solid circles refer to compressions, i.e., first motion upward, the open circles to dilations, i.e., first motion downward.* [*Reproduced with permission from Brillinger, Udias and Bolt* (1980).]

what may be learned, some details of the earthquake process will be set down. The usual assumption (the elastic rebound theory) is that earthquakes are due to faulting. A crack initiates at a point and (in the case of pure slip) spreads out to form a fault plane. As the crack passes a given point, slip takes place (on the fault plane) resulting in a stress drop and the radiation of seismic waves. The radiated (P-) waves may be shown to have a quadrantal pattern with one of the axes parallel and the other perpendicular to the fault plane of the event. It follows, and this is what Byerly (1926) contributed, that the data may be used to estimate the fault-plane orientation. Having an estimate of the fault plane and the direction of motion on that plane is important to geology and geophysics. Researchers seek to tie together surface and subsurface features, to consider regional stress directions and to use the results to confirm and extend the theory of plate tectonics. The results can be crucial to seismic risk computations.

Byerly proceeded graphically and this has continued to generally be the working approach. However, the results so obtained are subjective, have no attached measure of uncertainty and may not be easily combined with estimates derived from other events at the same site.

The problem may be approached in formal statistical fashion as follows. The data available consist of hypocenter of earthquake, locations of observatories, directions of observed first motions (compressions or dilations) at the observatories and a store of knowledge concerning the Earth's structure [velocity models as given in Jeffreys and Bullen (1940), for example]. It may further be argued that seismographic noise is approximately Gaussian [see Haubrich (1965)]. Let a fault plane be described by three angles $(\theta_T, \phi_T, \theta_P)$. Let $A_{ij}(\theta_T, \phi_T, \theta_P)$ denote the theoretical expression for the wave amplitude on the focal sphere for event i at station j. This expression may be found in Brillinger, Udias and Bolt (1980). (The focal sphere is a "little" sphere of unit radius around the hypocenter. In carrying out the amplitude computation, one has to trace the ray from the hypocenter to the observatory through the focal sphere.) Let Y_{ij} denote the realized amplitude of the seismogram at the onset of the event. Then one can write $Y_{ij} = \alpha_{ij}A_{ij} + \varepsilon_{ij}$, with α_{ij} a scale factor and ε_{ij} normal mean 0 and variance σ_{ij}^2 variate. Here α_{ij} reflects the attenuation the signal experiences in traveling from the source to the observing station, whereas ε_{ij} represents noise caused by disturbances unrelated to the earthquake of concern. Let $y_{ij} = 1$ if $Y_{ij} > 0$ and $= 0$ otherwise. It follows that

$$\text{Prob}\{y_{ij} = 1\} = \text{Prob}\{Y_{ij} > 0\} = \Phi(\rho_{ij}A_{ij}),$$

writing $\rho_{ij} = \alpha_{ij}/\sigma_{ij}$, for this signal-to-noise ratio. The model may be further expanded by including a term γ_{ij} to allow for reader and recorder errors, now writing

$$\text{Prob}\{y_{ij} = 1\} = \gamma_{ij} + (1 - 2\gamma_{ij})\Phi(\rho_{ij}A_{ij}). \tag{7}$$

Precise data correspond to γ and σ small (hence ρ large) and imprecise to γ near 0.5 or ρ near 0.

The model is seen to take the form of a nonlinear probit (with a term γ corresponding to "natural mortality"). An example of a corresponding likelihood is provided by

$$\prod_{ij} \Phi(\rho_i A_{ij})^{y_{ij}}\big(1 - \Phi(\rho_i A_{ij})\big)^{1-y_{ij}}, \tag{8}$$

assuming ρ to depend on event alone and $\gamma = 0$. One can now proceed to estimate the unknown parameters $\theta_T, \phi_T, \theta_P, \rho_i$ by maximum likelihood.

Figure 6 includes the fitted planes for the case of event 4 of the Alaska sequence. These particular estimates were computed restricting the likelihood (8) to the observations of event $i = 4$ and including a γ term as in (7).

It is critical to assess the fit of any model. In Brillinger, Udias and Bolt (1980), this was done by comparing the theoretical and estimated probability functions. Figure 7 is based on a pooled analysis of some 16 of the Alaskan events (labeled by i previously) that seemed to go together. It has been assumed that the ρ_i are all equal in the fit studied. The figure provides the empirical probability that the observed first motion agrees with the theoretical as a function of amplitude. The fitted values $z = \hat{\rho}\hat{A}_{ij}$ have been grouped into cells of width 0.1 in the analysis.

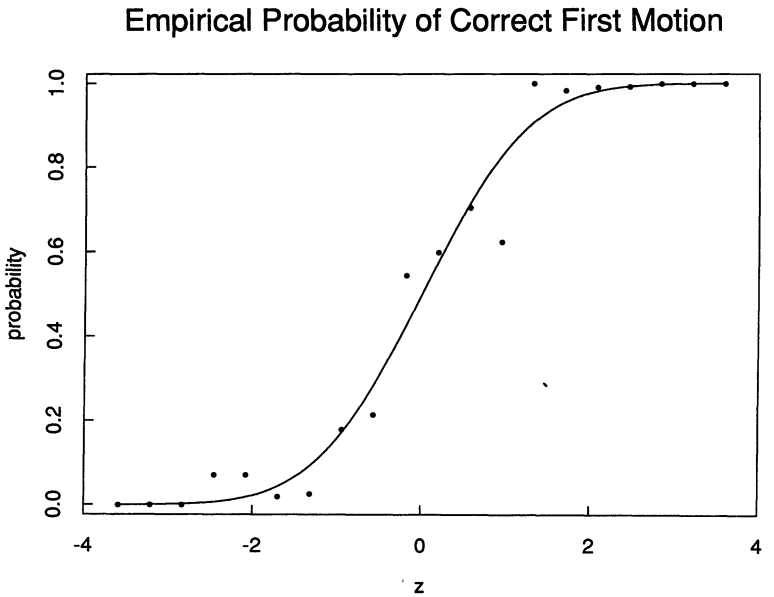

FIG. 7. *A plot of the statistic* (9) *of Section* 3 *and* $\Phi(z)$ *for the data of* 16 *events of the Alaskan sequence of* 1964. *The plot is meant to assess the validity of the model* (7). *Here z refers to the values* $\hat{\rho}\hat{A}_{ij}$.

What is plotted are $\Phi(z)$ and

$$(9)\qquad \#\{(i,j)|\operatorname{sgn} Y_{ij} = \operatorname{sgn} \hat{A}_{ij},\ z-h<\hat{\rho}\hat{A}_{ij}<z+h\}/\#\{(i,j)|z-h<\hat{\rho}\hat{A}_{ij}<z+h\},$$

for $h = 0.5$. Here $\hat{A}$ refers to $A(\hat{\theta}_T, \hat{\phi}_T, \hat{\theta}_P)$ and # refers to the count of the number of elements in the set. The fit seems adequate.

The results of further computations of this type may be found in Brillinger, Udias and Bolt (1980) and Buforn (1982). The maximization program VA09A of the Harwell subroutine library, see Hopper (1980), proved effective in determining the maximum likelihood values. The estimates were, however, nonunique and poorly determined in some cases of small data sets.

An important by-product of such analyses is to form clusters of like fault-plane solutions for events in the same region, in order to get at motions occurring on the same fault plane; see Udias, Munoz and Buforn (1985), for example. The maximum likelihood standard errors are useful in this connection. The practical implication of the work just reported is that first motions for large collections of events may be handled routinely and that geophysical conjectures may be checked formally. The final fault-plane solution may be plotted in traditional fashion allowing examination of the data for difficulties. What remains is for more realistic seismic source models than the one treated in the papers listed to

be fitted statistically. An elementary reference to the subject matter of concern here is Boore (1977).

4. Quantification of earthquakes. One of the important and difficult questions of seismology is how to measure the "size" of an earthquake. Size is an essential feature that a seismologist makes use of in attempts to deal with earthquake hazards and to understand the basic phenomena of concern. Specifically, the seismologist is not only interested in estimating the direction of movement at the source, he is further interested in the overall deformation that took place and the amount of energy that was released. Among the physical quantities of interest for a given earthquake are the seismic moment (a measure of the seismic energy released from the entire fault) and the stress drop (difference between the initial and final stress.)

For a variety of seismic source models, seismologists have related the seismic moment and stress drop to characteristics of the amplitude spectrum $|S(\lambda)|$, the modulus of the Fourier transform of the signal. Suppose that the seismogram is written as

$$Y(t) = s(t; \theta) + \varepsilon(t),$$

where $s(\cdot)$ is the signal, θ is an unknown parameter and $\varepsilon(\cdot)$ is a noise disturbance. If $S(\lambda; \theta)$ denotes the Fourier transform of $s(t; \theta)$, then what is given, from the source model, is the functional form of $|S(\lambda, \theta)|$. A reason for working in the Fourier domain here is that distracting phase information is eliminated. Common forms (for displacement measurements) include

$$|S(\lambda; \theta)| = \alpha/\sqrt{1 + (\lambda/\lambda_0)^{\beta}} \text{ and } \alpha/\left\{1 + (\lambda/\lambda_0)^2\right\},$$

with $\theta = \{\alpha, \beta, \lambda_0\}$. The seminal paper on the determination of such functional forms and on the relationship of their parameters to the "size" of the earthquake is Brune (1970/1971). Estimates of the seismic moment and stress drop may be determined once estimates of α and λ_0 are available. That the parameters relate to size and duration will be seen for a particular functional form in the discussion that follows. The empirical practice has been to estimate the unknowns graphically from a plot of the modulus of the amplitude of the empirical Fourier transform $|d_Y^T(\lambda)|$. The following formal procedure was suggested in Brillinger and Ihaka (1982).

The asymptotic distribution of $|d_Y^T(\lambda)|$ may be evaluated in the case of stationary ε using a central limit theorem of the type mentioned in Section 2. The asymptotic distribution is found to depend on $|S(\lambda; \theta)|$ and $f_{\varepsilon\varepsilon}(\lambda)$ alone. Hence one needs an expression only for the modulus of S, and as stated above, this is what the seismologist generally provides. Next, with the model $Y(t) = s(t; \theta) + \varepsilon(t)$ and small noise,

$$|d_Y^T(\lambda)| = |S(\lambda; \theta)| + \left(d_\varepsilon^T(\lambda) + d_\varepsilon^T(-\lambda)\right)/2 + \cdots,$$

showing variation around $|S|$ not depending on $|S|$. However, when deviations of $|d_Y^T|$ from a final fitted form are plotted versus the fitted values, dependence of the error on $|S|$ is apparent. An example is provided in Figure 8. This is the

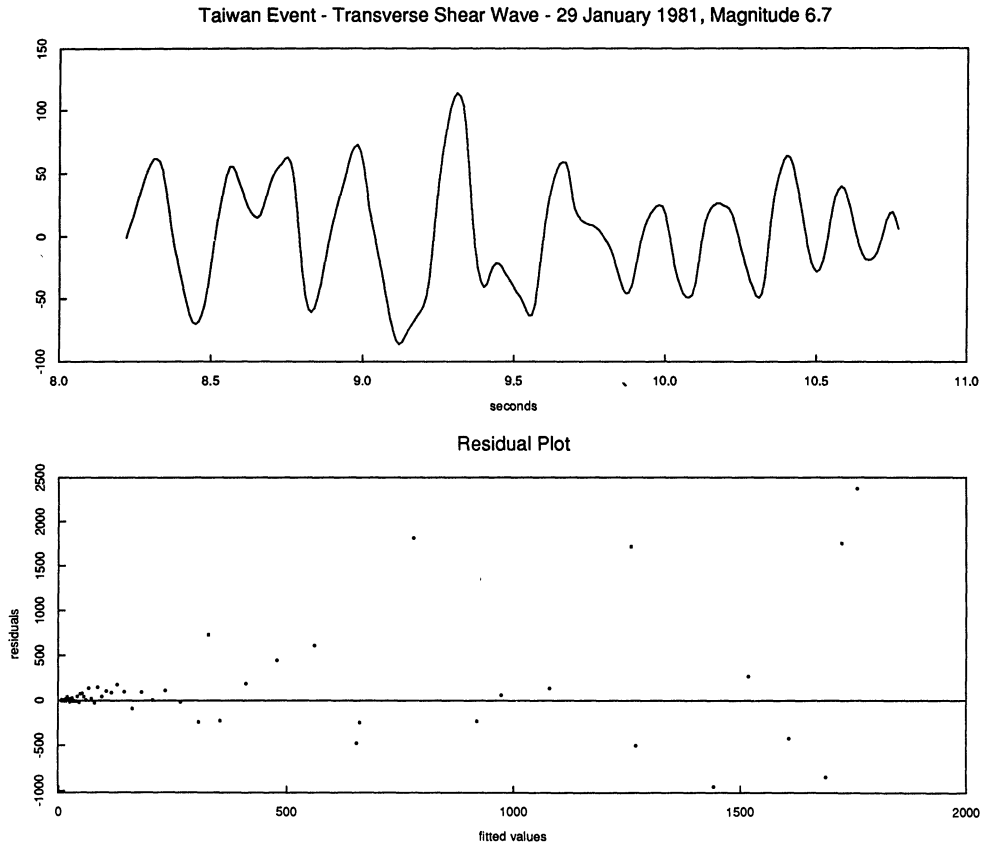

FIG. 8. *The upper graph gives the transverse S-wave component of the vibrations of the January* 29, 1981, *magnitude* 6.7, *Taiwan earthquake as recorded by the central accelerometers of the Smart* 1 *array. The array is approximately* 30 *km northwest of the epicenter of the event. The lower graph plots the differences between the amplitudes of the Fourier transform of the data and corresponding (final) fitted values. The data stretch consisted of* 256 *points.*

result of computations for an earthquake of magnitude 6.7 that occurred in Taiwan on January 29, 1981. The data were recorded by one of the instruments of the Smart 1 array; see Bolt, Tsai, Yeh and Hsu (1982). The upper graph of the figure provides the transverse S-wave portion of the recorded accelerations. The lower graph provides the deviations plot just referred to. This plot suggests that the noise is in part "signal generated" in this case. There are various physical phenomena that can lead to signal-generated noise. These include multipath transmission, reflection and scattering. The following is an example of a model that includes signal-generated noise:

$$Y(t) = s(t) + \sum_k \left(\gamma_k s(t - \tau_k) + \delta_k s^H(t - \tau_k)\right) + \varepsilon(t), \tag{10}$$

with the τ_k time delays, with s^H the Hilbert transform of s and with γ_k, δ_k

reflecting the vagaries of the transmission process. [The inclusion of the Hilbert transform allows the presence of phase shifts. The Hilbert transform is discussed, for example, in Brillinger (1975a), page 32]. With the γ_k, δ_k, τ_k random and after evaluating the large sample variance, one is led to approximate the distribution of $Y_j = d_Y^T(\lambda_j)$ by a complex normal with mean $S(\lambda_j; \theta)$ and variance $\Gamma_j = 2\pi T(\rho^2|S(\lambda_j; \theta)|^2 + \sigma^2)$, where now ε has been assumed to be white noise (of variance σ^2), and also it is assumed that $E\gamma_k$, $E\delta_k = 0$ and that the process τ_k is Poisson. The ratio ρ^2/σ^2 measures the relative importance of signal-generated noise. This variance is seen to depend on the "signal" through $|S|$ and leads to

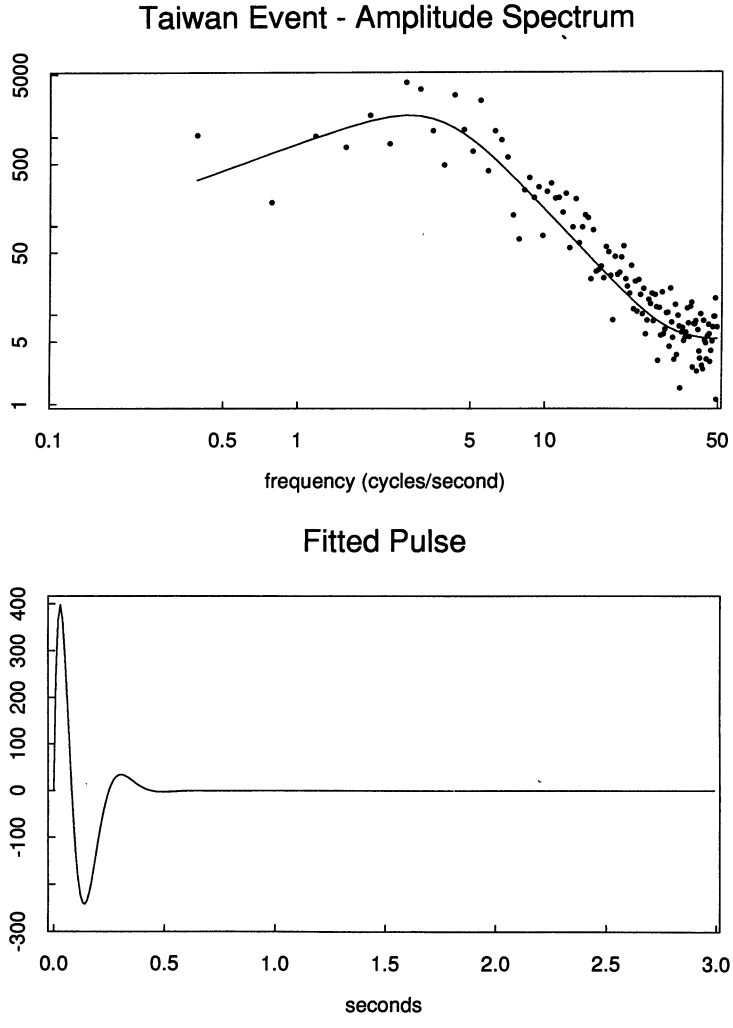

FIG. 9. *The upper graph provides the amplitudes of the Fourier transforms of the Taiwan data of Figure 8 and the corresponding fitted expected values as computed for the model of Section 4. Both scales of the plot are logarithmic. The lower graph provides the fitted pulse* $\hat{s}(t)$.

wedging of the type present in Figure 8. One can proceed to estimate θ by deriving the marginal likelihood based on the $|Y_j|$. This likelihood may be evaluated and found to be

$$\prod_j \left(\exp\left(-\frac{|Y_j|^2 + |S_j|^2}{\Gamma_j} \right) I_0\left(\frac{2|Y_j|\,|S_j|}{\Gamma_j} \right) \frac{1}{\Gamma_j} \right),$$

where I_0 denotes a modified Bessel function. The upper graph of Figure 9 shows a fit of the model $|S(\lambda)| = \alpha|\lambda|/(1 + (\lambda/\lambda_0)^4)$ to the data of Figure 8. This functional form was settled on after the degree of fit of two more elementary forms was examined.

Details may be found in Ihaka (1985). We remark that this model fit corresponds to a time domain pulse $s(t) = \alpha\lambda_0^2 p(\lambda_0 t)$, where

$$p(t) = \left[\sin\frac{t}{\sqrt{2}} - t\sin\left(\frac{t}{\sqrt{2}} + \frac{\pi}{4} \right) \right] e^{-t/\sqrt{2}},$$

for $t > 0$ and $p(t) = 0$ otherwise. The expression $s(t) = \alpha\lambda_0^2 p(\lambda_0 t)$ indicates how λ_0 corresponds (inversely) to the duration of the event and how α corresponds to size. The lower graph of Figure 9 provides a plot of the fitted pulse. Once estimates of α, λ_0 are at hand, these may be converted to estimates of the seismic moment and stress drop via theoretical relationships developed by geophysicists.

The maximum likelihood fit of the model was carried out by a computer program written by Ihaka. This program also generates standard error estimates and standardized residuals. These later may be used to assess the goodness of fit of the model. Figure 10 provides a plot of the standardized residuals against the

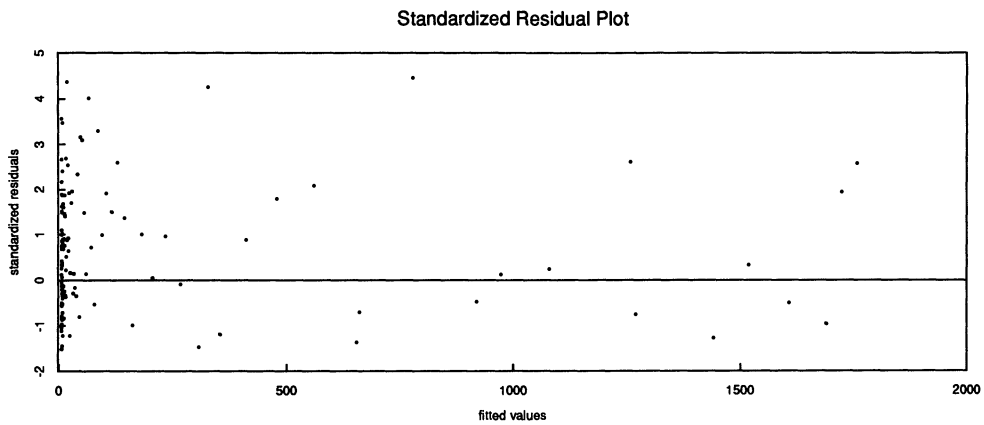

FIG. 10. *A standardized residual plot, based on the model* (10), *corresponding to the lower graph of Figure* 8. *The differences between the amplitudes of the Fourier transform values and their fitted expected values have been divided by their fitted standard deviations to obtain standardized residuals.*

fitted values of the same format as the residual plot of Figure 8. The wedging corresponding to signal-generated noise in the later plot is no longer present; however, there is a definite suggestion that the fit might be improved in the region where the signal has low amplitude. Luckily, this is the region of least importance. It awaits future analysis. It might be handled by allowing the series $\varepsilon(t)$ to have a nonconstant spectrum.

5. Array data. Today it would be a strange thing indeed for an earthquake to be recorded on just one seismometer. In fact, from the very earliest days, readings of the same event at geographically scattered observatories have been made use of. Since the 1960s, seismometers have been deliberately arranged in geometric designs over distances of the order of miles to hundreds of miles in order to allow extraction of traditional information and sometimes ellicitation of new information.

An important use has been the estimation of the direction from which a seismic signal is arriving and the velocity with which it is moving. One manner in which this is done is by the computation of estimates of frequency–wavenumber spectra. The procedure may be described as follows. Suppose one has array data; $Y(x_j, y_j, t)$, $j = 0, \ldots, J$ and $t = 0, \ldots, T-1$. Here (x_j, y_j) denotes the coordinates of the location of the jth sensor. The frequency–wavenumber periodogram of this data is given by

$$(11) \qquad \left| \sum_j \sum_t Y(x_j, y_j, t) \exp\{-i(\mu x_j + \nu y_j + \lambda t)\} \right|^2, \qquad -\infty < \mu, \nu, \lambda < \infty.$$

A motivation for this definition is the following. Suppose one has a plane wave $Y(x, y, t) = \rho \cos(\alpha x + \beta y + \gamma t + \delta)$ of temporal frequency γ and wavenumber $\kappa = (\alpha, \beta)$. Then the periodogram will have a peak near (α, β, γ). (Incidentally, this wave is moving with apparent velocity $\gamma / \sqrt{\alpha^2 + \beta^2}$ from azimuth given by $\tan \phi = \beta / \alpha$.) An example of array data is given by Figure 11. What is plotted are the locations of nine of the seismometers of the Smart 1 array located in Taiwan. Also plotted are the portions of the traces used in the computations. These traces correspond to the vertical P-wave part, of the January 29, 1981 earthquake. (The initial near-flat part is the noise, saved in a buffer, just before the onset of the wave.) The estimated epicenter of this earthquake was 30 km southeast of the array. Figure 12 gives a central portion of the frequency–wavenumber periodogram, for this data, as computed via formula (11), at frequency λ corresponding to 1.944 cycles/s. (The temporal frequency 1.944 was picked on the basis of a times-series analysis of the individual seismograms.) There is seen to be a large peak in the southeast quadrant, at an azimuth that turns out to correspond to that of the epicenter of the event. The radial distance corresponds to the velocity of P-waves.

Seismologists working with this type of data have often preferred to employ, what they call, the "high-resolution" or "Capon" statistic [see Capon (1969)] instead of the periodogram (11). The high-resolution statistic typically shows more dramatic peaks than the periodogram. Before defining it, we introduce

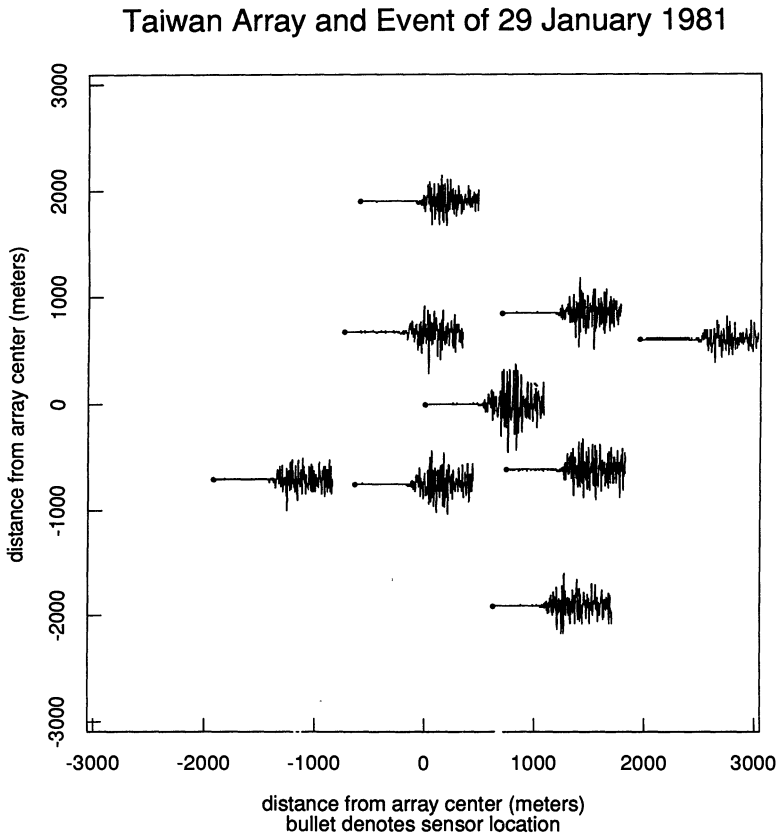

FIG. 11. *The vertical P-wave portion of the January 29 Taiwan earthquake as recorded at nine of the sensors of the Smart 1 array. The "bullets" are plotted at the physical locations of the sensors. Noise immediately preceding the arrival of energy from the event had been saved in a buffer.*

some notation. Let $\mathbf{Y}(t)$ denote the j-vector $[Y(x_j, y_j, t)]$. Set

$$\mathbf{Y}_k = T^{-1} \sum_{t=0}^{T-1} \mathbf{Y}(t) \exp\left\{-i\frac{2\pi kt}{T}\right\},$$

for $k = 0, 2, \ldots$. Further let $\mathbf{B} = [\exp\{-i(\mu x_j + \nu y_j)\}]$. If $\lambda = 2\pi l/T$, l an integer, then the periodogram (11) is proportional to $|\bar{\mathbf{B}}^T \mathbf{Y}_l|^2$. Next define

$$\mathbf{M} = \sum \mathbf{Y}_k \bar{\mathbf{Y}}_k^T,$$

with the sum over k with $2\pi k/T$ near λ. Now the high-resolution statistic at frequency λ may be defined as $1/\bar{\mathbf{B}}^T \mathbf{M}^{-1} \mathbf{B}$. If $Y(x, y, t) = \rho \cos(\alpha x + \beta y + \gamma t + \delta) +$ noise, this statistic may be expected to show a peak for (μ, ν) near (α, β) and λ near γ. This statistic has been introduced, in part, in order to be able to present the next example.

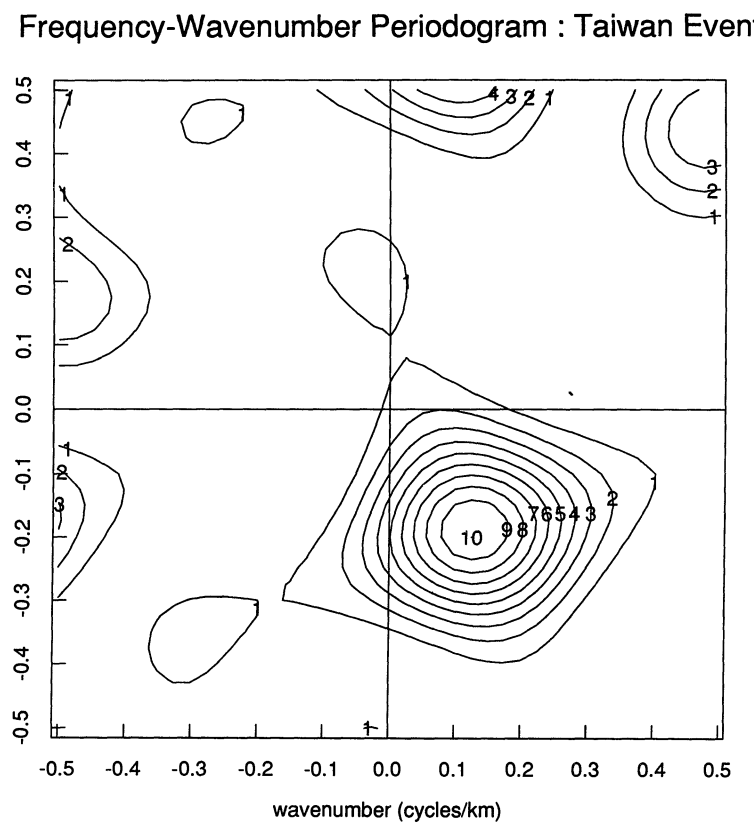

FIG. 12. *The frequency–wavenumber periodogram of the data of Figure* 11. *The time series stretches contain* 720 *points. The temporal frequency employed is* 1.94 *cycles/s.*

Figure 13 is reproduced from Scheimer and Landers (1974). It shows the high-resolution statistic computed for two portions of data recorded by the Large Aperture Seismic Array (LASA) in Montana following a strip-mining blast. These computations confirmed the validity of the high-resolution approach. The statistic for one portion shows a single large peak in the direction of the blast. The statistic for the following portion shows energy arriving from various directions. This analysis provided empirical proof of the existence of scattering of seismic waves. That this phenomenon existed had been theorized for years. A frequency–wavenumber data analysis has provided the confirmation.

Spectral analyses are (too) often thought of as being appropriate only for stationary data. As the preceding example shows, the technique may be highly useful in nonstationary cases as well. As a second example we mention the results of Bolt, Tsai, Yeh and Hsu (1982). If, in fact, an earthquake is caused by faulting, then the direction of the source of seismic energy will be changing as the fault is ripping, that is, as the fault tip is advancing. In the paper cited, Bolt, Tsai, Yeh and Hsu present high-resolution spectra for succeeding time stretches

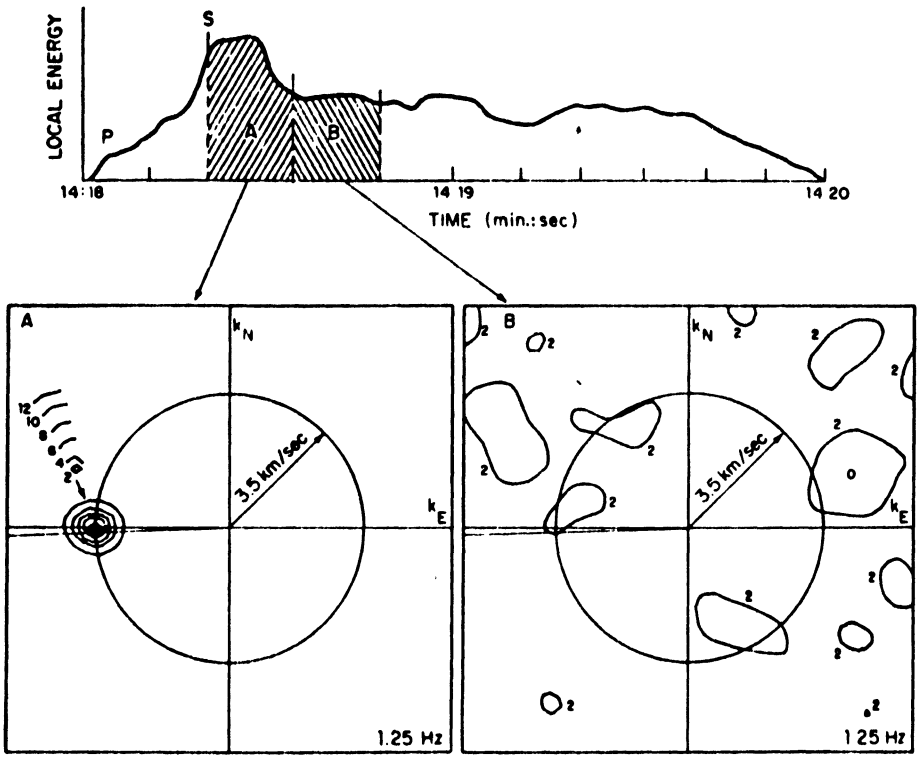

FIG. 13. *The high-resolution (or Capon) spectrum computed for S-wave data recorded, following a strip-mining blast, at the Large Aperture Seismic Array located in Montana. The temporal frequency was* 1.25 *cycles/s (corresponding to the principal frequency of the S-wave).*

of the January 29, 1981 Smart 1 event. There is an apparent shift in direction with time. Their work may have been the first experimental measurement of a seismic dislocation moving along a rupturing fault.

In each of the preceding two examples, frequency–wavenumber analysis has allowed researchers to confirm the presence of suspected scientific phenomena.

6. Exploration seismology (reflection seismology). The problem of learning the Earth's crustal structure can be approached as one of system identification. The approach to be described takes advantage of the fact that the Earth happens to be made up of layered strata. Signals, such as powerful impacts or explosions, can be deliberately input to the Earth and the consequent vibrations recorded by an array of seismometers or geophones. Such experiments may be carried out in a search for gas and oil, or in a scientific study of the general geological makeup of a region of interest. The results of these experiments may be viewed as one of the grand success stories for statistical techniques generally, and of least squares particularly. An unusual aspect of the inferences made is that in many cases one gets to examine their validity, by the later

drilling of a well. No worked example is presented in this section of the paper, in large part because data sets are hard to come by. The material is presented, however, because it provides a case where a rather complete solution (design through confirmation) can be presented and because the basic experimental technique is also employed in the neurophysiological case, where so much less is known.

In its simplest form, the energy of an initiated seismic disturbance propagates through the Earth with a spreading wavefront. When it meets an interface between geological strata, part of the energy may be reflected back and part continue forward due to the difference in acoustic impedance at the interface. The sensors record the returning reflected energy echoes. Knowledge of subsurface velocities allows estimation of the depths and angles of inclination of the various reflectors, whereas knowledge of the locations of reflectors allows estimation of velocities. (One notes again a see-saw in the collection of knowledge.) In practice, the initiating impacts will be repeated a number of times at the same location and at points of a grid. The power of averaging is again used.

If the input signal is taken to be $X(t)$ and if $Y(t)$ denotes the corresponding output, then the two may be modeled as related, assuming linearity and time invariance, by

$$Y(t) = \int a(t-s)X(s)\,ds. \tag{12}$$

The function $a(\cdot)$ is called the impulse response, since if the Dirac delta function $\delta(t)$ is taken as input, then the resulting output is $Y(t) = a(t)$. The function $a(\cdot)$ evidences the reflectors and velocities in the earth beneath the source and receiver. The model and its interpretation may be motivated as follows. Suppose a pulse is applied at time τ. Suppose in consequence a wave is generated, travels at velocity v_1 to a reflector at distance d_1 and a proportion α_1 is reflected back. With $X(t) = \delta(t-\tau)$, then $Y(t) = \alpha_1\delta(t-\tau-2d_1/v_1)$. (This is actually the naive model for radar or sonar.) Suppose further that the transmitted portion continues downward at velocity v_2 to a reflector at distance d_2 and a portion of its energy is reflected back, some of which is transmitted by the first reflector to reach the receiver. Now the response has the form $Y(t) = \alpha_1\delta(t-\tau-2d_1/v_1) + \alpha_2\delta(t-\tau-2d_1/v_1-2d_2/v_2)$. This last is seen to correspond to the system of expression (12) with impulse response $a(t) = \alpha_1\delta(t-2d_1/v_1) + \alpha_2\delta(t-2d_1/v_1-2d_2/v_2)$. One can clearly extend this model to situations with many layers, many velocities and many corresponding transmission and reflection coefficients. Peaks in the function $a(t)$ may be seen as corresponding to reflectors. (It must be noted that unfortunately such an elementary interpretation is likely to be complicated in practice by interfering phenomena such as ghost reflections. Some techniques have been developed to handle these.) The basics of exploration seismology are discussed in Wood and Treitel (1975), Waters (1978) and Robinson (1983).

The problem has now been formulated as one of system identification; given stretches of corresponding input X and output Y, determine an estimate of the impulse response $a(\cdot)$. In the case that a pulse close to a Dirac δ may be

generated and that the function $a(\cdot)$ drops off to 0 reasonably quickly, a convenient procedure results from taking

$$X(t) = \sum_{m=1}^{M} \delta(t - m\Delta),$$

as input, and the "average evoked response"

$$\hat{a}(s) = \frac{1}{M} \sum_{m=1}^{M} Y(s + m\Delta),$$

as an estimate of $a(s)$. This input corresponds to applying pulses periodically. The estimate corresponds to stacking and averaging.

Suppose one sets $m_{YX}(t) = Y^*X(t)$ for some convolution operation "$*$." Then from (12) one has

$$m_{YX}(t) = \int a(t - s)m_{XX}(s)\, ds$$

and one has a deconvolution (or inverse) problem to solve. Suppose one decides to seek an $X(\cdot)$ such that

$$\int a(t - s)m_{XX}(s)\, ds \approx a(t),$$

to allow elementary processing. In terms of Fourier transforms, the left-hand side here may be expressed as $\int \exp\{i\lambda t\}A(\lambda)M_{XX}(\lambda)\, d\lambda$, with M_{XX} a Fourier transform of m_{XX}. Then what is wanted is an X such that $M_{XX}(\lambda) \approx 1$ on the support of $A(\cdot)$. If $A(\lambda)$ is known to be near 0 for $0 < \lambda < \lambda_0$ and for $\lambda > \lambda_1$, then a possible function is the "chirp" signal

$$X(t) = \cos\left(\left[\lambda_0 + (\lambda_1 - \lambda_0)\frac{t}{\tau}\right]t\right), \quad \text{for } 0 \le t \le \tau.$$

In the seismic case, the values of λ_0, λ_1 have been determined in various experiments. The chirp probe originated in radar work during World War II [see Cook and Berenfield (1967)]. It may be seen to attach near equal power to the frequencies between λ_0 and λ_1. In the seismic case special devices have been developed to input the chirp signal to the earth. The signal is input repeatedly and the results averaged. The response is then convolved with the chirp function, that is, m_{YX} is formed to estimate $a(\cdot)$. Structure can appear dramatically during the cross-correlation processing described here.

In practice, subtle further processing is employed to handle wavefront curvature, ghost reflections and other natural phenomena that may be present.

7. Other topics. There are other problems arising in seismology to which statistical methodology can be applied fruitfully. These include analysis of the coda (i.e., of the irregular trailing part of the disturbance), analysis of scattering, risk analysis, nonlinear phenomena, point process studies, polarization, cepstral analysis, discrimination of earthquakes from explosions [see, e.g., Tjøstheim

(1981)], seismicity study, travel time table construction, attenuation laws, earthquake location and azimuthal dependence of characteristics. Vere-Jones and Smith (1981) discuss several of these problems. In some cases work has begun.

8. Discussion. Seismologists have long been serious users of statistical methods. One finds Harold Jeffreys making the following statement in the entry, "Seismology, statistical methods," in the *International Dictionary of Geophysics*: "The uncertainty is as important a part of the result as the estimate itself. ... An estimate without a standard error is practically meaningless." Hudson (1981) remarks: "The success of the Jeffreys–Bullen travel time tables was due in large part to Jeffreys' consistent use of sound statistical methods." When I asked my colleague B. A. Bolt what he saw as the role of statistics in seismology, he replied: "Seismology is largely an inferential science. ... The role of statistics in seismology is to provide a rigorous procedure for turning observations on seismic waves, etc., into probabilistic statements about properties of the (real) Earth."

One may note that work in seismology is characterized by massive data sets, inherent variability and measurement error, defining/fitting/refining models, design of experiments, simulation, probabilistic description, needs for robust/resistant procedures, predictive situations, inverse problems and combination of observations. Statistics has much to offer in all these connections.

9. Update. Since the lectures were presented in 1983, work has progressed on various of the topics covered. Abrahamson (1985) has employed Smart 1 data to better see the movement of the fault rupture tip. Chiu (1986) studies the problem of estimating the parameters of a moving energy source. Lindberg (1986) develops "optimal" tapers to employ in the estimation of the frequencies of free oscillations. The approach of Kitagawa and Gersch (1985) to nonstationary data seems likely to prove of broad practical applicability. The book by Udias, Munoz and Buforn (1985) goes into substantial detail over the formal estimation of fault-plane parameters. Copas (1983) sets down an expression for the variance of a statistic like that of (9). Brillinger (1985) develops a maximum likelihood statistic for detection and estimation of a plane wave given array data. Donoho, Chambers and Larner (1986) develop a robust/resistant procedure for better aligning the seismic traces of a section. Mendel (1983, 1986) presents maximum likelihood state space-based methods for handling the data of reflection seismology. Shumway and Der (1985) indicate how the EM method may be employed to deconvolve pulses hidden in seismic traces. The nongaussianity of seismograms obtained in reflection seismology is being taken specific advantage of; see Giannakis and Mendel (1986). The techniques of Donoho (1981) and Lii and Rosenblatt (1982) seem bound to prove useful in the seismological case. The thesis, Ihaka (1985), has been completed. Ogata [e.g., Ogata (1983) and Ogata and Katsura (1986)] has carried out a variety of likelihood-based analyses of earthquake times as a point process. Many statisticians have begun working on statistical aspects of inverse problems. We specifically mention O'Sullivan (1986).

One can speculate on where the field of statistical seismology will go in the coming years. It seems clear that there will be much concern with non-Gaussian noise and signals, that vector-valued spatial–temporal data and analysis will become the norm, that large-scale conceptual models will be set down and that there will be a variety of techniques developed for borrowing strength in situations with scanty data, e.g., risk estimation.

III. Neurophysiology

> ..., *modern biometry is the interdisciplinary endeavor to build structural stochastic models of biological phenomena.*
>
> J. Neyman (1974)

10. The field and its goals. Neurophysiology is the branch of science concerned with how the elements of the nervous system function and work together. The functioning is seen to involve chemical mechanisms, electrical mechanisms and physical arrangement. The studies extend from the movements of individual ions, through to the mass behavior of the components of the brain.

The goals of neurophysiologists range to the heroic: how to explain things like memory, emotion, learning, sleep, expectation, behavior. At a less ambitious level, neurophysiologists are concerned with how a single nerve cell responds to stimuli, transmits information and changes with alterations of the environment.

The neuron is both the functional and structural unit of the nervous system. The brain is a multiprocessor of dramatic complexity. The elements of the nervous system may be said to differ from those in the seismic case, in that they apparently have purposes.

The field is largely experimental with researchers collecting varied and extensive data sets. The data include photographs made via electron microscopes, fluctuating voltages and current levels within single nerve cells and finally electroencephalograms (the brain's electrical potential at points near the skull.) The studies are sometimes simply observational, but often complex experimental designs are employed.

Important techniques that are made use of include staining to identify individual neurons, insertion of microelectrodes to make measurements within individual cells and the averaging of whole suites of responses to a stimulus of interest in order to reduce what can be the dominant effects of noise. Many experiments are computer controlled and computer processed.

Discoveries made by neuroscientists include the following. Nerve cells communicate with each other in both a chemical and electrical fashion, the voltage pulse that travels along a neuron's output fiber is of near constant shape and there are a broad variety of nonlinear phenomena that occur. A number of verifiable physical laws and effective deterministic models (such as the Hodgkins–Huxley equations) have been set down. Much insight has been gained, especially at the level of small groups of neurons. At the level of the brain itself, knowledge is mainly phenomenological. Here the brain is viewed as a black box and studied by system identification techniques. Whatever the approach, discoveries have been made leading to lifesaving and life-improving clinical diagnoses.

Statistical methods entered with the quantification of the field. No single individual scientist seems to have had a dominating effect, rather there have been many contributing workers—researchers concerned with electroencephalograms (EEGs) and researchers concerned with small collections of neurons. Statistical methods entered both because of high noise levels and because a variety of phenomena seemed to be inherently stochastic. Evidence for this last is presented in Burns (1968) and Holden (1976). Pertinent books on neurophysiology include Freeman (1975), Aidley (1978) and Segundo (1984). General reviews of statistical models and methods in neurophysiology are given in Moore, Perkel and Segundo (1966) for the cases of single neurons and of small groups of neurons and by Glaser and Ruchkin (1976) for EEGs. Statistical methods for classification and pattern recognition, for handling artifacts and for data summarization are in common use.

Neurobiology is one of the most active branches of science. The physiological phenomena with which it is concerned are fundamental and in most cases barely understood.

11. Neuronal signaling. One of the important means by which nerve cells communicate is via spike trains. The inlays at the tops of the three graphs of Figure 14 give examples of spike times representative of three different sorts of neuronal behavior; pacemaker (near-periodic), bursting (activity occurs in bursts) and bursting with acceleration (of firing within bursts).

Suppose that a neuron fires at times τ_n, $n = 0, \pm 1, \pm 2, \ldots$. A convenient formal representation of its temporal behavior is provided by writing

$$Y(t) = \sum_n \delta(t - \tau_n),$$

with $\delta(\cdot)$ the Dirac delta function. This representation leads to results analogous to ordinary time-series results in many cases. In the case that the τ_n are random, one has a stochastic point process $\{\tau_n\}$. A principal descriptor of a point process is provided by its rate function. This is given by

$$\lim_h \text{Prob}\{\text{point in } (t, t + h]\}/h,$$

as h tends to 0. In the stationary case, where the stochastic properties of the process do not depend on the time origin, the rate function is constant and so only crudely useful then.

The autointensity function is an important parameter in the stationary case. It is defined as

$$\lim_h \text{Prob}\{\text{point in } (t, t + h] | \text{point at } 0\}/h,$$

as h tends to 0. It is a point process analog of the autocovariance function of time-series analysis in a general sense. This parameter may be used, for example, to describe the behavior of spontaneously firing neurons. Figure 14 presents examples for three cases. In the first case, the neuron is firing approximately periodically. The (estimate of) the autointensity is seen to oscillate (with period

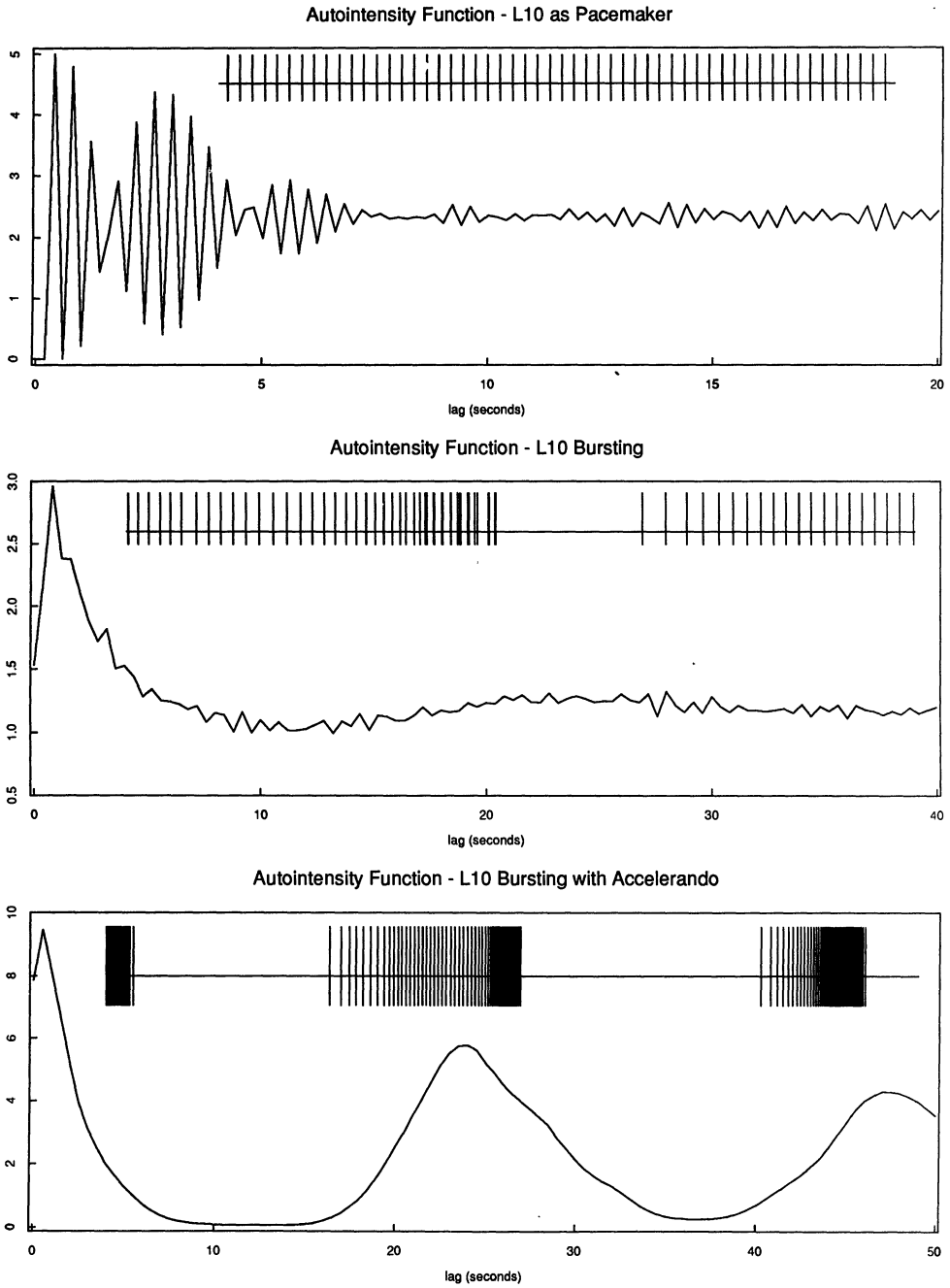

FIG. 14. *Point process data (spike train) from the nerve cell* L10 *of Aplysia californica. The cell is behaving in three different fashions. The inlays at the tops of the three graphs give brief stretches of the data (but not on the same time scales as the autointensities). The functions plotted are estimates of the autointensity functions based on* 1538, 1019, 1631 *spikes, respectively.*

equal to the interval between the points). In the second case, the neuron is evidencing activity in bursts. The probability that the neuron fires again soon after it has fired is high. In the third case, the neuron is also firing in bursts; however, now there is structure within the bursts, the rate of firing is seen to increase therein. The bursts here are at regular intervals.

The autointensity functions have been estimated, for this figure, by the statistic

$$\#\{|\tau_n - \tau_m - t| < h/2\}/Nh,$$

with N the total number of points, with h a small binwidth and with t lag. (Here # refers to the count of the number of points in the set.) The data analyzed are for the cell L10 of *Aplysia californica*, the sea hare. They were collected and previously analyzed by Bryant, Marcos and Segundo (1973). The experimental procedures and details of the data preparation may be found in that reference.

A question that arises in the study of small networks of neurons is which neurons are interacting with which? In other words, which spike trains are associated with which others? A useful parameter to employ in the study of such questions is provided by the cross-intensity function. Supposing one has spike trains named M and N, then the cross-intensity function of N given M at lag t is defined as

$$\lim_h \text{Prob}\{N \text{ point in } (t, t+h] | M \text{ point at } 0\}/h,$$

as h tends to 0. If the M spike train consists of points σ_m and the N train of points τ_n, then this cross-intensity may be estimated by

$$\#\{|\tau_n - \sigma_m - t| < h/2\}/Mh,$$

with M denoting the number of M points in the data set, with h a small binwidth and with t lag. Figure 15 presents three examples of estimated cross-intensity functions. The first graph refers to data from cells L3 and L10 of *Aplysia californica*. The behavior exhibited here is that of negative association; L10's firing is inhibiting the firing of L3 (for approximately 0.5 s). If one asks whether the values at negative lags differ from the level of no-association by more than sampling fluctuations, one finds they do not. This result is consistent with the cell L10 driving the cell L3. The middle graph corresponds to positive association. It is for a cell in the right visceropleural connective (RVP) and cell R15. The first cell tends to excite the second for about 0.25 s. The final graph represents a more complicated (polyphasic) situation. These data sets were also analyzed in Bryant, Marcos and Segundo (1973), where further details may be found. The approximate sampling distributions of such statistics were developed in Brillinger (1975b). It was found, for example, that it could be more convenient to graph the square root of the estimate in some circumstances.

The cross-intensity function, being a point process analog of covariance, may be expected to be an inadequate measure of relationship (as usual, correlation does not imply causation). In the case of elementary statistical data, it is usual

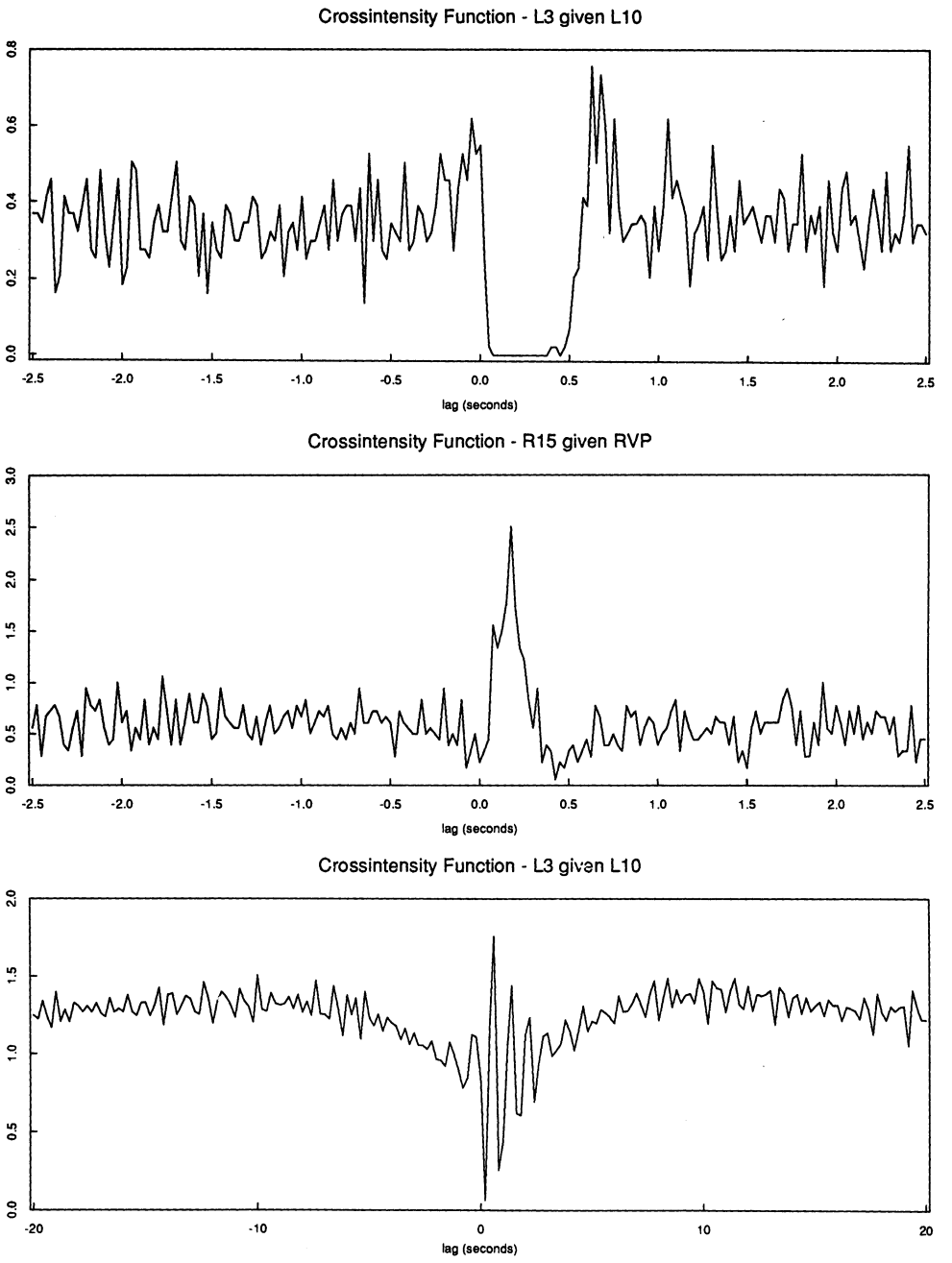

FIG. 15. *Estimates of the cross-intensity functions for three pairs of Aplysia neurons. The estimates are based on* (1746,302), (1101,288), (1019,993) *spikes in the pairs of trains, respectively.*

to turn to regression as a better technique. In the point process case it is possible to carry out regression-type analyses. For example, one may fit the following form of model:

$$\lim_h \text{Prob}\{N \text{ spike in } (t, t+h) | M \text{ spike train}\}/h = \mu + \sum_m a(t - \sigma_m),$$

as h tends to 0. The function $a(t)$ appearing in this model is referred to as the impulse response. This model may be fit as follows. Set

$$d_M^T(\lambda) = \sum_{m=1}^{M} \exp\{-i\lambda\sigma_m\},$$

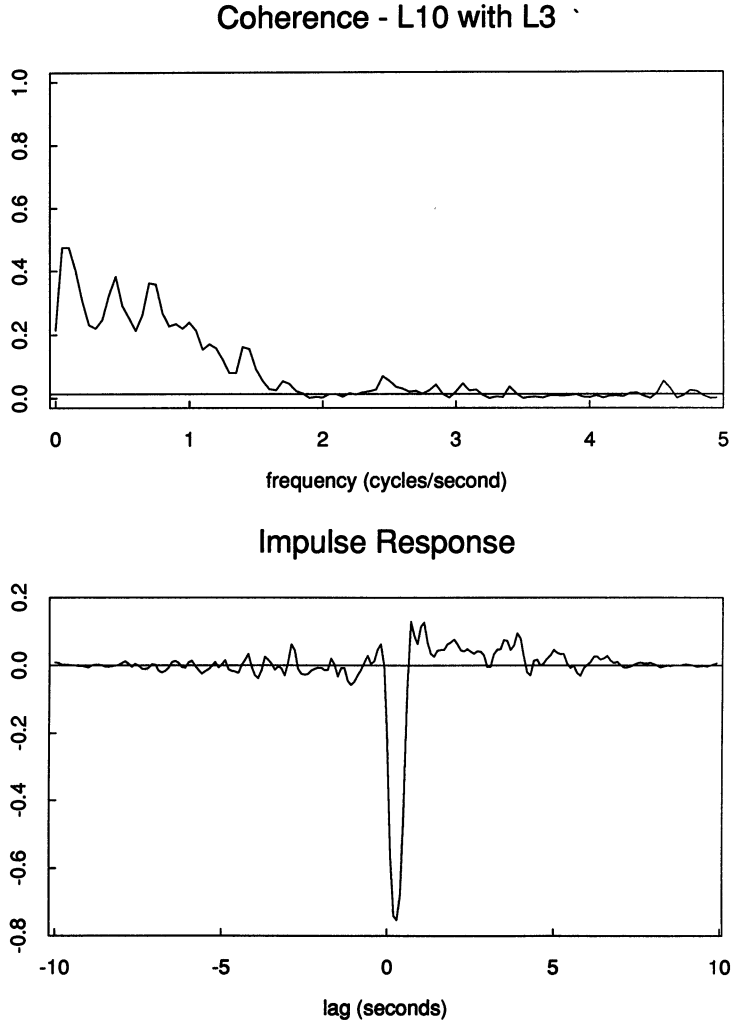

FIG. 16. *The estimated coherence and impulse response for the data of the upper graph of Figure 15. The horizontal line gives an estimate of the level exceeded by chance only 5% of the time when the spike trains are independent.*

with a similar definition for $d_N^T(\lambda)$. These are point process analogs of the empirical Fourier transform (3) of time-series data. The cross-periodogram of the given data at frequency λ is defined as

$$I_{NM}^T(\lambda) = (2\pi T)^{-1} d_N^T(\lambda)\overline{d_M^T(\lambda)}.$$

If the cross-periodogram is smoothed to obtain $f_{NM}^T(\lambda)$, then $f_{NM}^T(\lambda)$ is an estimate of the cross-spectrum in the case that $\{M, N\}$ is a bivariate stationary point process. Now $A(\lambda)$, the Fourier transform of the impulse response $a(t)$, may be estimated by $f_{NM}^T(\lambda)f_{MM}^T(\lambda)^{-1}$. The impulse response itself may be

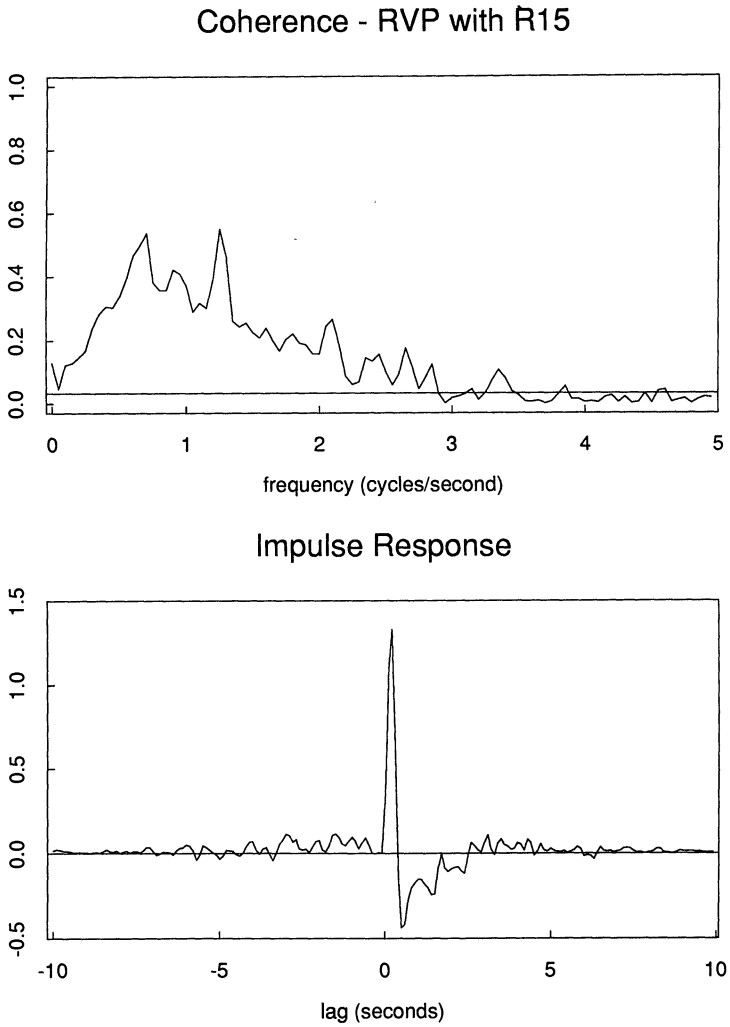

FIG. 17. *The estimated coherence and impulse response for the data of the middle graph of Figure 14. The horizontal line in the upper graph gives the approximate upper 95% null point of the distribution of the sample coherence.*

estimated by back Fourier transforming A^T. The strength of the relationship proposed in the model may be measured, at frequency λ, by the sample coherency function $R^T_{NM}(\lambda) = f^T_{NM}(\lambda)/\sqrt{f^T_{MM}(\lambda) f^T_{NN}(\lambda)}$. Its modulus squared is called the sample coherence. The coherence lies between 0 and 1, being nearer to 1 the stronger the relationship. More details of these computations may be found in Brillinger (1975b) and Brillinger, Bryant and Segundo (1976). [We here follow the use of the terms "coherency" and "coherence" in Wiener (1930).]

Figures 16 and 17 provide the results of such an analysis for the first two data sets of Figure 15. In each case the first graph is of the sample coherence. The coherences are at some distance from the value 1.0, but above the 95% null significance level (given by the horizontal lines in the figures). The relationship is inherently nonlinear, so it could have been anticipated that the coherence estimate would not be close to 1.0. Further discussion of these and similar analyses may be found in Brillinger, Bryant and Segundo (1976).

12. Assessing connectivities. Questions that can arise with small networks of neurons include; is one neuron driving the rest and if one apparently is, which one is it? The next data analysis to be presented addresses this question for three *Aplysia* cells L2, L3 and L10. From other experiments the neurophysiologists knew that cell L10 was driving cells L2 and L3. It was not known if there were any direct connections between L2 and L3. The first three graphs of Figure 18 present estimates of the three coherences, L10 with L2, L2 with L3 and L10 with L3. As might have been anticipated, these suggest relationship exists in each case.

It is possible to address the question of the direct connection of cells L2 and L3, in the presence of L10, by partial coherence analysis. Suppose that $\{A, B, C\}$ is a trivariate stationary point process. Let $R_{AB}(\lambda)$ denote the coherency function of processes A and B, with similar definitions of R_{AC} and R_{BC}. Then the partial coherency of the processes B and C, having removed the (linear time invariant) effects of process A, is given by

$$R_{BC|A} = \frac{R_{BC} - R_{BA}R_{AC}}{\sqrt{(1 - |R_{BA}|^2)(1 - |R_{CA}|^2)}}, \tag{13}$$

suppressing the dependence on λ. This definition may be motivated several ways. For example, it is the coherency between the processes resulting when their best linear predictors based on A are removed. Or, it is given by

$$\lim_{T\to\infty} \left| \operatorname{corr}\left\{ d^T_B - \frac{f_{BA}}{f_{AA}} d^T_A, d^T_C - \frac{f_{CA}}{f_{AA}} d^T_A \right\} \right|^2 .$$

Here corr denote the (complex) correlation coefficient and f_{BA}/f_{AA}, f_{CA}/f_{AA} are approximate regression coefficients. An estimate may be determined by substituting estimates for the quantities appearing on the right-hand side of expression (13). If there is no connection between the processes B and C beyond their

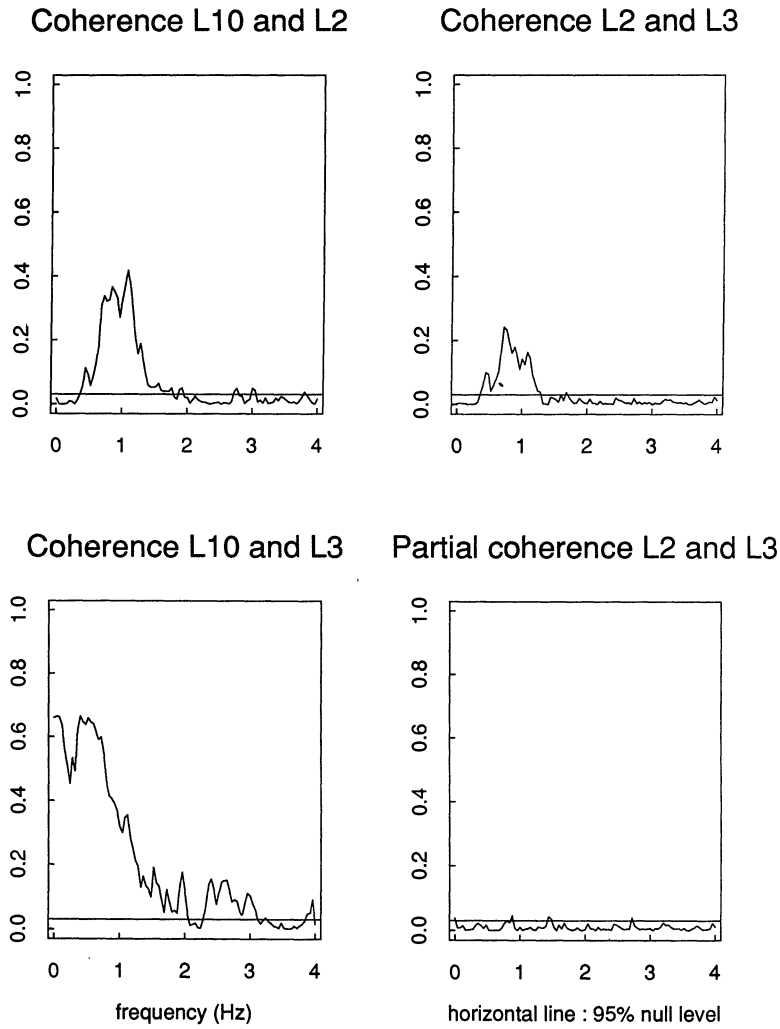

FIG. 18. *Data for a network of three Aplysia neurons. The partial coherence estimate is based on expression* (13) *of Section* 12. *In each case the horizontal solid line gives the approximate upper* 95% *null level.*

joint dependence on A, then the sample partial coherence $R^T_{BC|A}$ may be expected to be near zero.

The final graph of Figure 18 provides the results of the computation for the cells L2, L3 and L10. There is no suggestion of a direct connection being present. On reflection, it is quite astonishing the degree to which the linear models and quadratic statistics have apparently captured the dependency in a highly nonlinear situation. Further discussion and other examples of partial coherence computations may be found in Brillinger, Bryant and Segundo (1976).

13. A structural stochastic model. The analyses of neuronal firing, so far presented, are of the correlation and regression type. Parameters with direct biological interpretations have not been introduced. In Brillinger and Segundo (1979), a conceptual model is constructed and fitted by the method of maximum likelihood. The model involves the following elements.

Input to a nerve cell leads to electrical-current genesis. This current flows to a trigger zone, being filtered in the course of its passage. When the voltage level at the current zone exceeds a threshold value, the nerve cell fires. The neuron remembers back only to the time of previous firing. This process may be specified analytically as follows. Let $U(t)$ denote the voltage (membrane potential) at the trigger zone at time t. Let $B(t)$ denote the time elapsed since the neuron last fired. Let $X(t)$ denote the (measured) input to the cell. Then, assuming linearity and time invariance, one can write

$$U(t) = \int_0^{B(t)} a(s)X(t-s)\,ds,$$

for some summation function (impulse response) $a(\cdot)$. The neuron fires when the process $U(t)$ crosses a threshold level $\theta(t)$. Depending on the level at which the threshold is set and the internal mechanics of the nerve cell, the input will either accelerate (excite) or slow (inhibit) the firing. In Brillinger and Segundo (1979), this mechanism was completed and discretized as follows. Input to the cell was written X_t, $t = 0,\ldots,T-1$. Corresponding output was Y_t, $t = 0,\ldots,T-1$, with $Y_t = 1$ if there was a firing in the (small) interval immediate to t and with $Y_t = 0$ otherwise. With B_t denoting the time elapsed at t since the preceding time that $Y = 1$, they set

$$U_t = \sum_{s=0}^{B_t-1} a_s X_{t-s}.$$

The presence of B_t in the model had the effect of introducing a form of feedback. Finally, they assumed that the threshold function had the form $\theta_t = \theta + \varepsilon_t$ with the ε's independent normals having mean 0, variance 1 and cumulative distribution function $\Phi(\cdot)$.

The likelihood function of the given data and the model then took the form

$$\prod_{t=0}^{T-1} \Phi(U_t - \theta)^{Y_t}(1 - \Phi(U_t - \theta))^{1-Y_t}.$$

Parameter estimates were determined by maximizing this likelihood with respect to θ and the a_s. Approximate standard errors were determined by procedures traditional to maximum likelihood.

Figure 19 presents the results of one such analysis. In this case fluctuating current $X(t)$ was injected directly into the cell R2 of *Aplysia*. The current level was taken to have marginal distribution that was approximately uniform (but that is not crucial to the technique). The sampling rate was 50 samples/s. The upper graph of the figure gives a stretch of the noise signal injected and the corresponding times at which the neuron fired. Details of the experiments are given in Bryant and Segundo (1976). It is very difficult, if not impossible, to see a

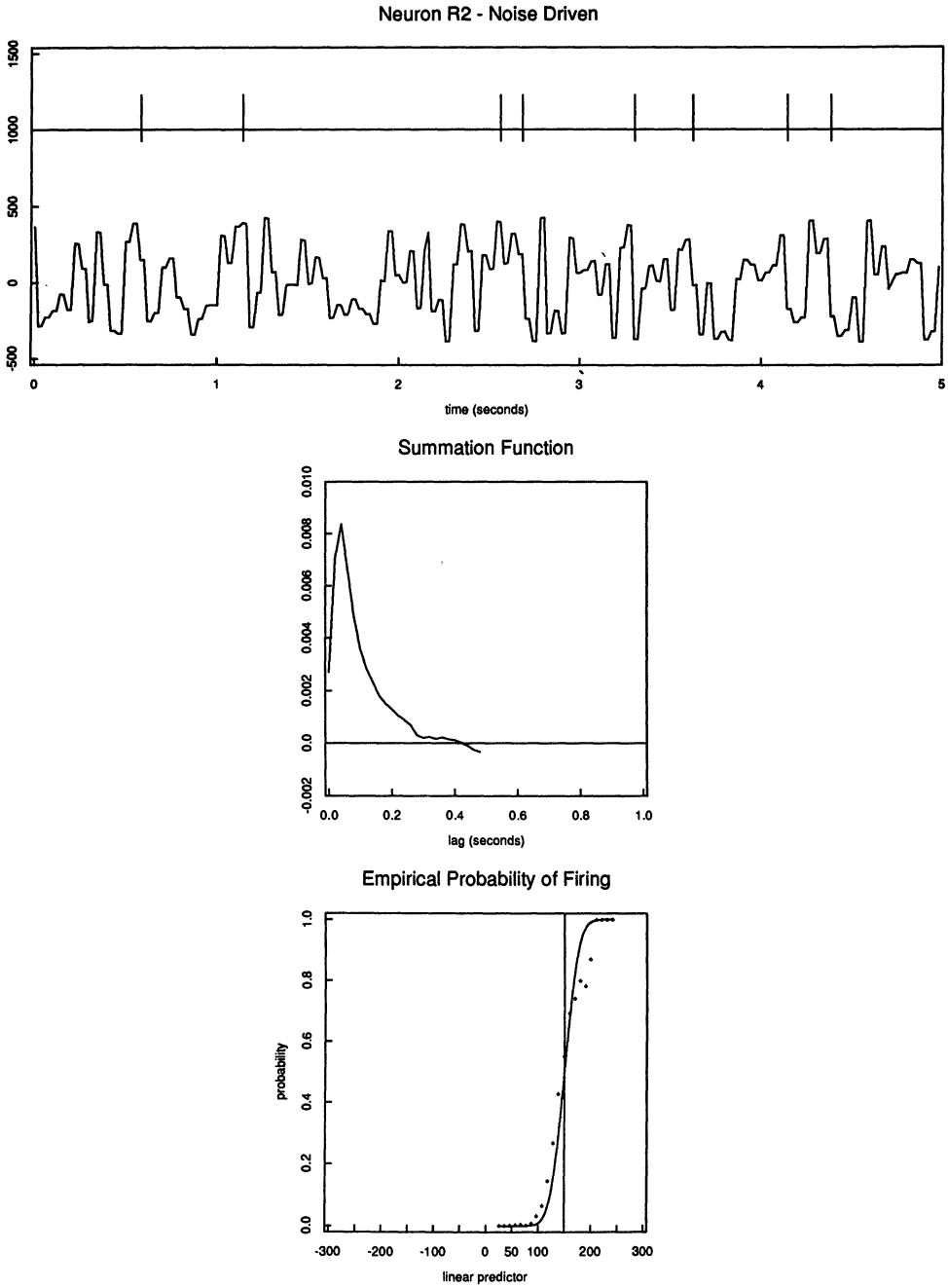

FIG. 19. *The results of fitting the neuron model of Section* 13 *to data obtained in an experiment with the cell* R2 *of Aplysia. The upper graph is a segment of the data. Noise (lower trace) is injected into the cell. The upper trace gives corresponding observed firing times. The middle graph gives the maximum likelihood estimate of the summation function* $a(\cdot)$, *estimated at* 25 *lags. The lower graph provides the statistic* (14) *of Section* 13 *and the curve* $\Phi(U - \hat{\theta})$ *with* $\hat{\theta}$ *the estimated mean threshold. The vertical line is at* $U = \hat{\theta}$.

connection between these two stretches of data. The middle graph gives the estimated summation function $\hat{a}_s$. The lower graph is one means of assessing the fit of the model. It is analogous to expression (9) of Section 3, and given by

$$(14) \quad \#\{Y_t = 1 \text{ with } U - h < \hat{U}_t < U + h\} / \#\{t \text{ with } U - h < \hat{U}_t < U + h\},$$

for small h, plotted versus U. Here

$$\hat{U}_t = \sum_{s=0}^{B_t - 1} \hat{a}_s X_{t-s}$$

is the fitted linear predictor. The smooth curve is the corresponding $\Phi(U - \hat{\theta})$. The fit may be described as adequate. The computations were carried out by a variant of the program developed for handling the seismic first-motion data of Section 3. Further examples and discussion may be found in Brillinger and Segundo (1979). Other types of input are employed and alternate estimating procedures compared there.

The large-sample properties of such estimates may be studied as in Sagalovsky (1982). A great advantage of the model-building approach, of this section, is that the parameters introduced and estimated have biological interpretations. A further advantage of the maximum likelihood approach, over that of partial coherency, is that the spike trains involved can be highly nonstationary.

14. Analysis of evoked responses. A traditional means of studying the nervous system involves applying sensory stimuli to a subject and examining the ongoing electroencephalogram for an evoked response. The stimulus may be auditory, visual (e.g., light flash, checkerboard pattern), olfactory, somatosensory (e.g., an electrical shock), gustatory or a task. Generally, the stimulus is applied for a time interval that is brief in comparison to the duration of the response. Evoked-response experiments play an essential role in quantitative biology. Because the experimenter is able to choose which stimuli to apply and when to apply them, conclusions can pass beyond associations noted, to formal inferences concerning causal mechanisms. These experiments are formally the same as the seismological reflection experiments described in Section 6.

Some dramatic success stories of the technique may be mentioned. One is presented in Bergamini, Bergamasco, Fra, Gandiglio and Mutani (1967). Siamese twins were joined in such a way that it was not possible to determine by traditional means if the peripheral nervous pathways were interconnected. Before operating, it was crucial to determine the interconnections of the twins. Ongoing EEGs were recorded for each. A series of trials were carried out in which each of the twins' legs was stimulated in turn by electrical shocks. What was found was that when a leg of one twin was stimulated, response was noted only in her EEG. On the basis of this information, the twins were separated—successfully. A second notable example of the use of the evoked response technique is provided by hearing exams for newborn infants (including infants asleep.) EEGs are recorded. These are examined for responses after loud clicks are made near the infants' ears. Rapin and Graziani (1967) present an

example for an infant with hearing difficulties, both wearing and not wearing a hearing aid. The hearing aid is found to have an objectively measurable effect.

Figure 20 presents an example of evoked-response data recorded at a 4×4 array of sensors implanted in a rabbit. In this case the stimulus was an odor and the sensors were implanted in order to study the rabbit's olfactory system. These responses were recorded concurrently. A second example is given in Figure 21. It gives the 20 successive responses evoked by applying a current pulse to the lateral olfactory tract of a rabbit and recording from a sensor implanted in the depth of the pre-piriform cortex. The signal is fairly pronounced in Figure 20. In Figure 21 the strength of the stimulus was weak and the signal is not apparent. Both of these data sets were collected in the laboratory of W. J. Freeman, University of California, Berkeley. Some details of his experiments may be found in Freeman and Schneider (1982).

Crucial to many evoked response experiments is the fact that it is generally insufficient to apply a stimulus once. Rather it must be applied repeatedly (perhaps thousands of times) and the responses averaged. (This is also true in the case of reflection seismology as mentioned earlier.) In the twins and infant examples discussed previously, M equaled 250 and 100, respectively. Formally, if $Y(t)$ denotes the measured EEG and the stimulus is applied at times σ_m, $m = 1, \ldots, M$, then it is usual to take as the basic statistic, the average evoked response

$$\bar{Y}(s) = \frac{1}{M} \sum_{m=1}^{M} Y(s + \sigma_m).$$

The left-hand column in Figure 22 presents the results of averaging the data of Figure 21 with $M = 3$, 5, 10, 20 and 38. With increasing averaging a signal is slowly appearing from the noise. Some alternate evidence for the presence of a signal is provided by the results of the right-hand column. These are averages of 38 responses, where the stimulus has been applied at a succession of increasing strengths.

A variety of questions, which have statistical formulations, arise in the course of work with evoked responses. (1) Does an applied stimulus elicit a response? (2) Do two different stimuli elicit the same response? (3) Is the same response elicited at two different sensor locations? (4) Is the response stationary? (5) If the order of stimuli application is altered, are the corresponding responses altered? (6) Are the effects of different stimuli additive? (7) How does the response depend on the stimulus intensity? (8) How do the responses depend on exogenous variables? To go with answers to these questions, researchers seek quick efficient data collection, precise estimates and indications of variability. Difficulties that commonly arise include small response, large noise, variability in response, artifacts present and superposed effects. Next in this section, two formal set-ups will be presented that may be employed to address the situation.

Suppose, to begin, that there is a single stimulus and that it is applied at times σ_m. Let $a(\cdot)$ denote the response in a single-shock experiment. If the system is time invariant and the effects of the various shocks additive

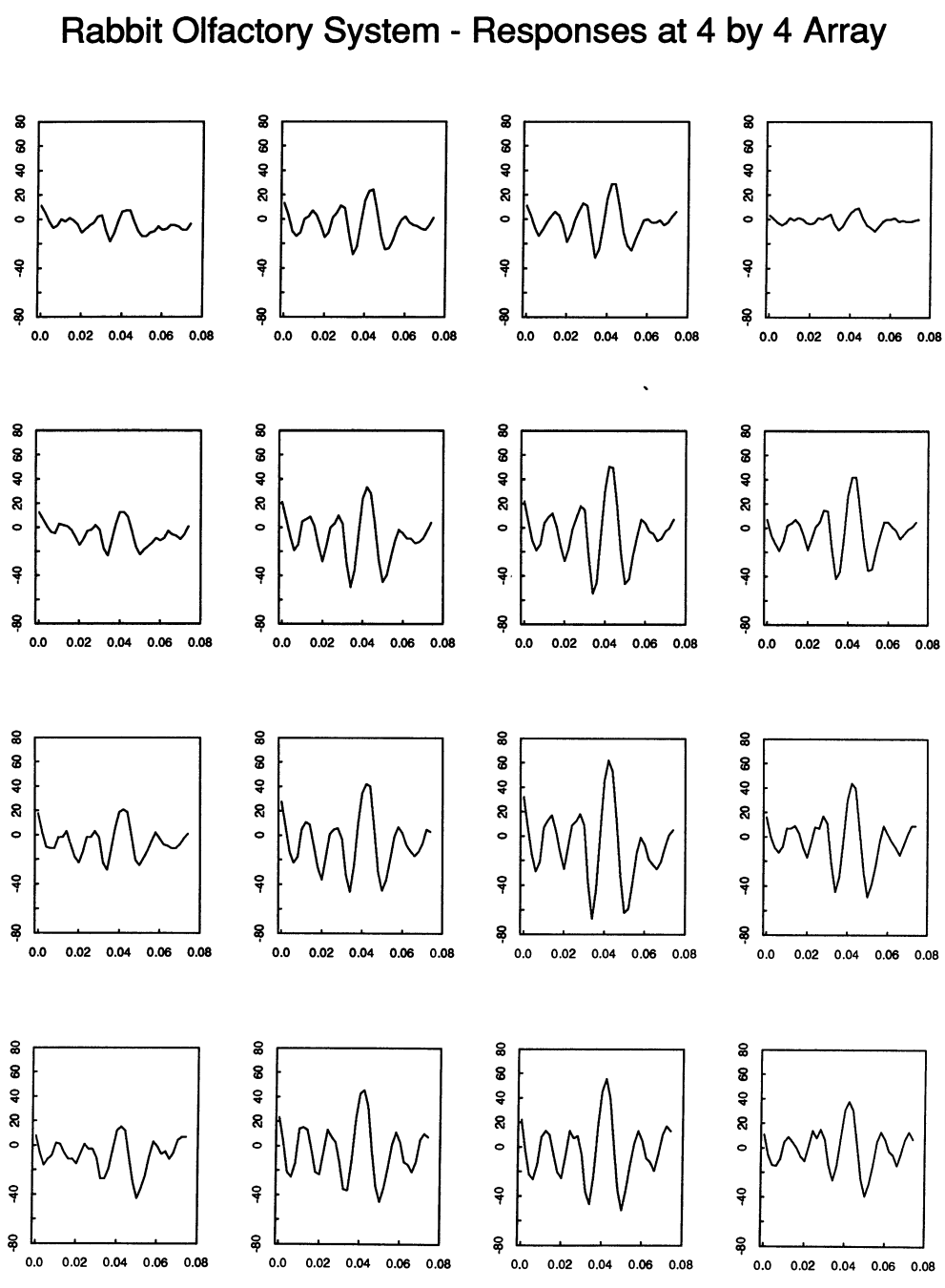

FIG. 20. *The bursts of electroencephalographic activity recorded at the* 16 *sensors of a* 4 × 4 *array implanted in a rabbit, following the stimulation of the rabbit by an odor. The units of the x-axis are in seconds.*

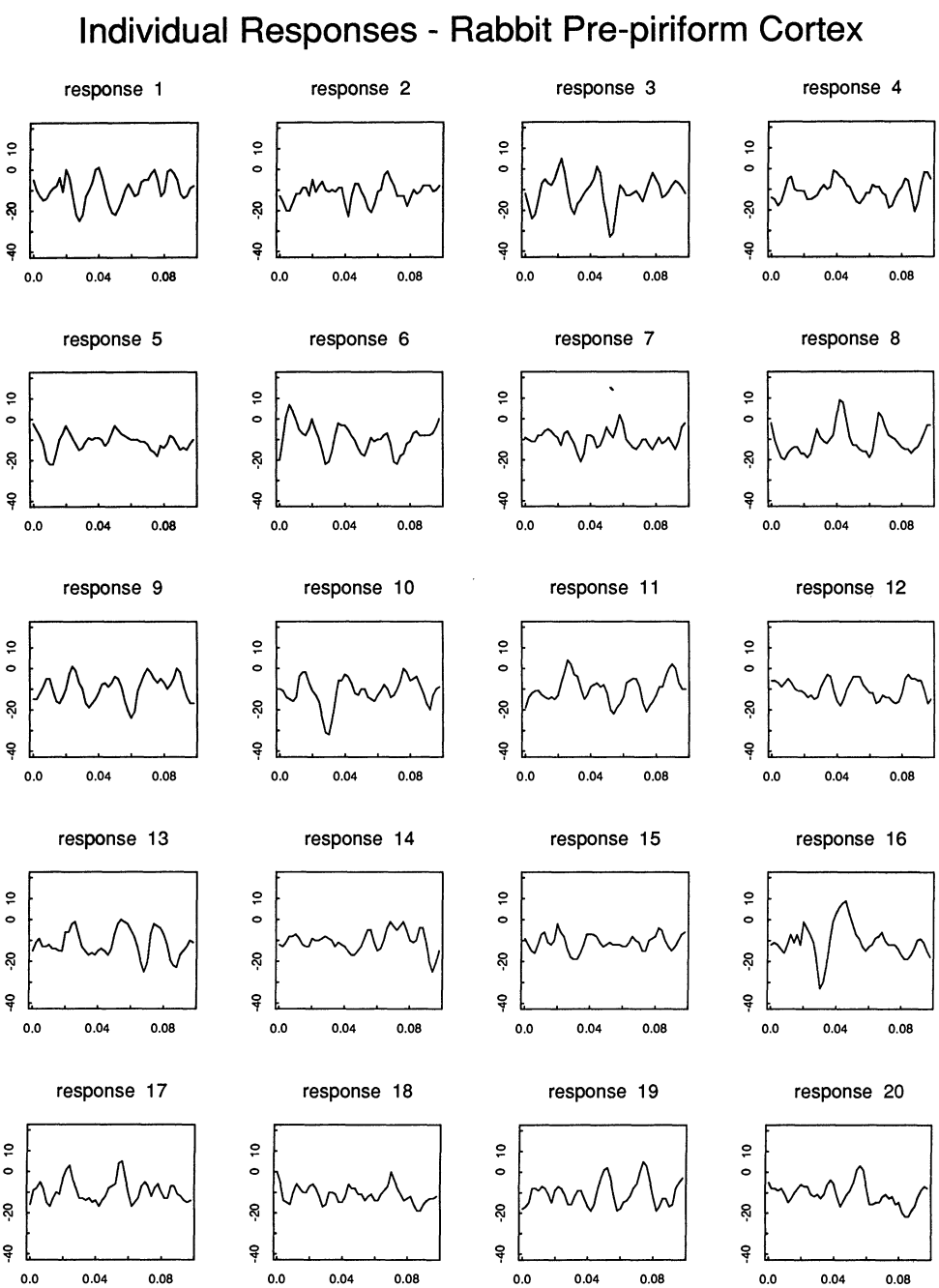

FIG. 21. *Twenty successive responses evoked in the pre-piriform cortex by (electrically) stimulating a rabbit. The x-axis units are in seconds.*

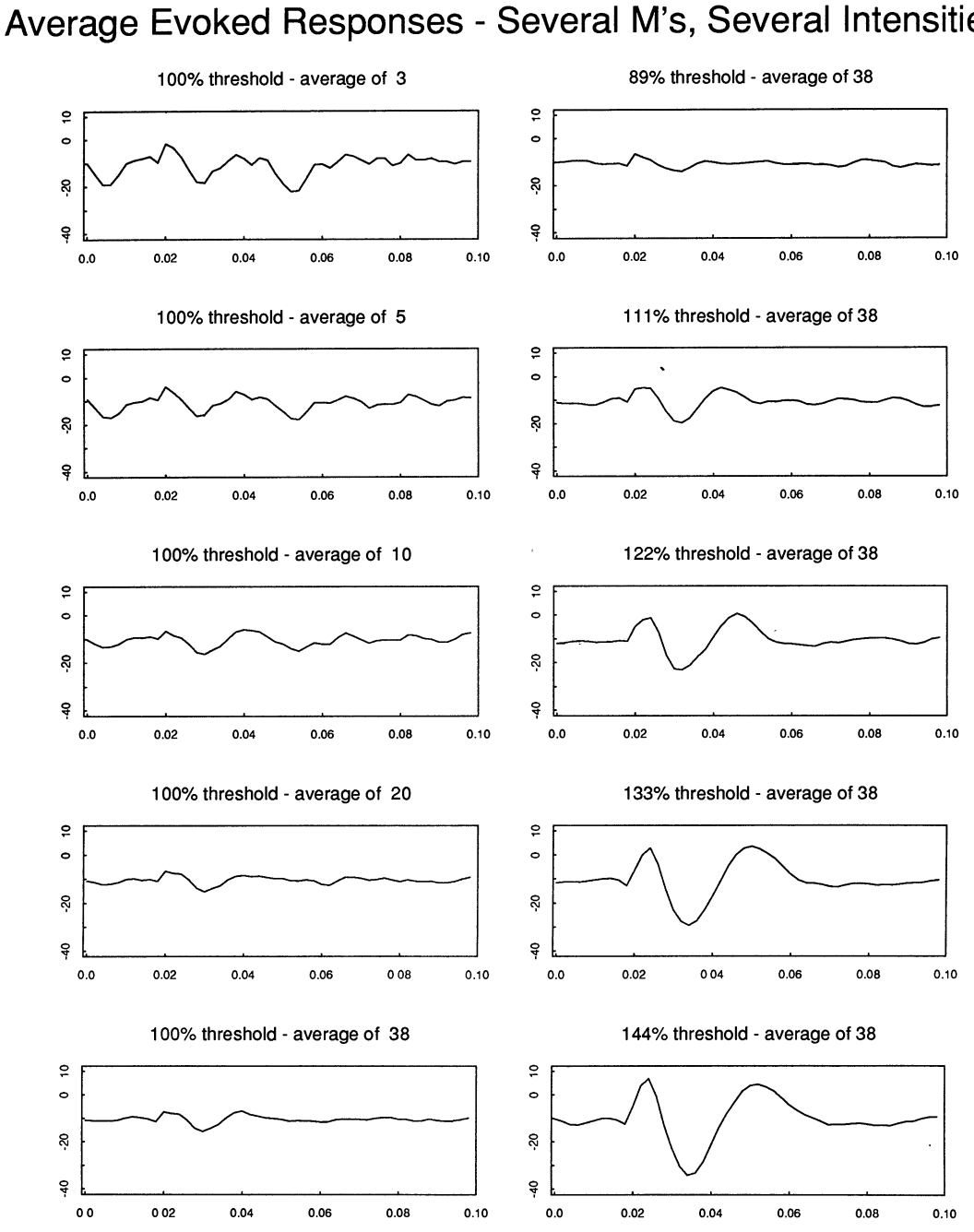

FIG. 22. *The various graphs here are meant to show the effects of changing the number of responses averaged (left column) and the strength of stimulus applied (right column) for data such as that of Figure* 21.

(superposable), then a model for consideration is

$$Y(t) = \mu + \sum_m a(t - \sigma_m) + \varepsilon(t), \tag{15}$$

with $Y(\cdot)$ denoting the ongoing EEG and $\varepsilon(\cdot)$ denoting noise. In the case of the EEG this model seems to have to be empirically verified, rather than being an implication of basic biology. (In the seismological case it came out of a conceptual framework.) For example, the assumption of superposability may be examined as follows for the animal studied. To begin, carry out some single-shock experiments, i.e., apply the shocks at times far enough apart that their individual effects seem likely to have died off. Let $\hat{a}(s)$ denote the average of the responses evoked, with s lag since stimulus application. Now carry out some two-shock experiments, i.e., apply shocks say Δ time units apart. Let $\hat{b}(s, \Delta)$ denote the average of the responses evoked. To examine the assumption of superposability compare $\hat{a}(s) + \hat{a}(s - \Delta)$ with $\hat{b}(s, \Delta)$. The results of carrying out such a check, in an experimental situation, are given in Biedenbach and Freeman (1965). They form averages of $M = 150$ responses and do not note departure from superposability.

We now turn to one formal analysis of the model (15). If one writes

$$X(t) = \sum_m \delta(t - \sigma_m),$$

then (15) takes the form

$$Y(t) = \mu + \int a(t - s)X(s)\, ds + \varepsilon(t),$$

i.e., it is seen to be the model of cross-spectral analysis. Taking Fourier transforms, one has

$$d_Y^T(\lambda) \approx A(\lambda)\, d_X^T(\lambda) + d_\varepsilon^T(\lambda),$$

for $\lambda > 0$, with $A(\lambda)$ denoting the Fourier transform of $a(\cdot)$. Consider a number of frequencies $\lambda_k = 2\pi k/T$ near λ. Then, assuming $A(\cdot)$ smooth, one has the approximate linear model

$$Y_k \approx A(\lambda)X_k + E_k,$$

with

$$Y_k = \sum_{t=0}^{T-1} Y(t)\exp\left\{-i\frac{2\pi kt}{T}\right\},$$

and similar definitions of X_k, E_k. Next, via a central limit theorem for empirical Fourier transforms, the noise variates E_k may be approximated by independent (complex) normals having mean 0 and variance $2\pi T f_{\varepsilon\varepsilon}(\lambda)$. All the inference procedures for the linear model become available. For example, as an estimate of the transfer function $A(\lambda)$, one has

$$\hat{A}(\lambda) = \sum_k Y_k \overline{X}_k \Big/ \sum_k X_k \overline{X}_k$$

and this variate will be approximately distributed as complex normal with mean $A(\lambda)$ and variance $2\pi T f_{\varepsilon\varepsilon}(\lambda)/\Sigma_k |X_k|^2$. This formulation has a variety of convenient properties. It directly extends to the cases of multiple stimuli and multiple responses. It handles stimuli of varying intensity. It allows the individual responses of the separate shocks to overlap. Formal inference procedures, such as tests, are available. Complex experiments may be designed and analyzed—complexities handled such as blocking, rotation, factorial treatment structure, measured covariates. Formal checks for interaction are available. Finally, one can turn to the question of optimal design.

It is sometimes convenient to adopt a different viewpoint for the problem. Suppose that the shocks are applied at times such that $\sigma_{m+1} - \sigma_m > V$ with $a(s) = 0$ for $s > V$ and $s < 0$. Write

$$Y_m(s) = Y(s + \sigma_m).$$

Then $Y_m(s) = \mu + a(s) + \varepsilon_m(s)$ for $0 \le s \le V$. The average evoked response is now conveniently denoted $\bar{Y}(s)$. As an example of the use of this formulation, suppose there are I different stimuli and that each are applied J times, then one is led to set down the model

$$Y_{ij}(s) = \mu_{ij} + a(s) + b_i(s) + \varepsilon_{ij}(s),$$

with i indexing stimuli and j indexing replicates. Other methodologies, such as grown curves and discriminant analysis, are seen to become available with this formulation.

It was mentioned that evoked-response data may be contaminated by artifacts. It is perhaps worth noting that robust/resistant estimates are directly available. Suppose one has a measure of distance, such as

$$\|Y - a\|^2 = \int_0^V [Y(s) - a(s)]^2 \, ds$$

and an estimate of scale $\hat{\rho}$. Then a family of robust/resistant estimates is provided by

$$\hat{a}(s) = \sum_m W_m Y_m(s) \Big/ \sum_m W_m,$$

with $W_m = W(\|Y_m - \hat{a}\|/\hat{\rho})$ and $W(\cdot)$ a univariate set of multipliers for robust/resistance. The estimate will need to be computed recursively. An elementary example is provided by the "trimmed mean"

$$\hat{a}(s) = \sum{}' Y_m(s)/\beta M,$$

with Σ' over the βM smallest $\|Y_m - \hat{a}\|$. This class of estimates was proposed in Brillinger (1979, 1981a) and investigated in Folledo (1983). The upper graph of Figure 23 provides an example of this estimate with 50% trimming ($\beta = 0.5$), in the case of data like that of Figure 21 (but with a stimulus of strength 122% of the threshold stimulus). The solid curve denotes the average evoked response, the dashed one the trimmed statistic. The two curves are nearly identical, although when examined, the individual responses are found to differ noticeably. Fifty-percent trimming was employed, because this is usually considered a highly

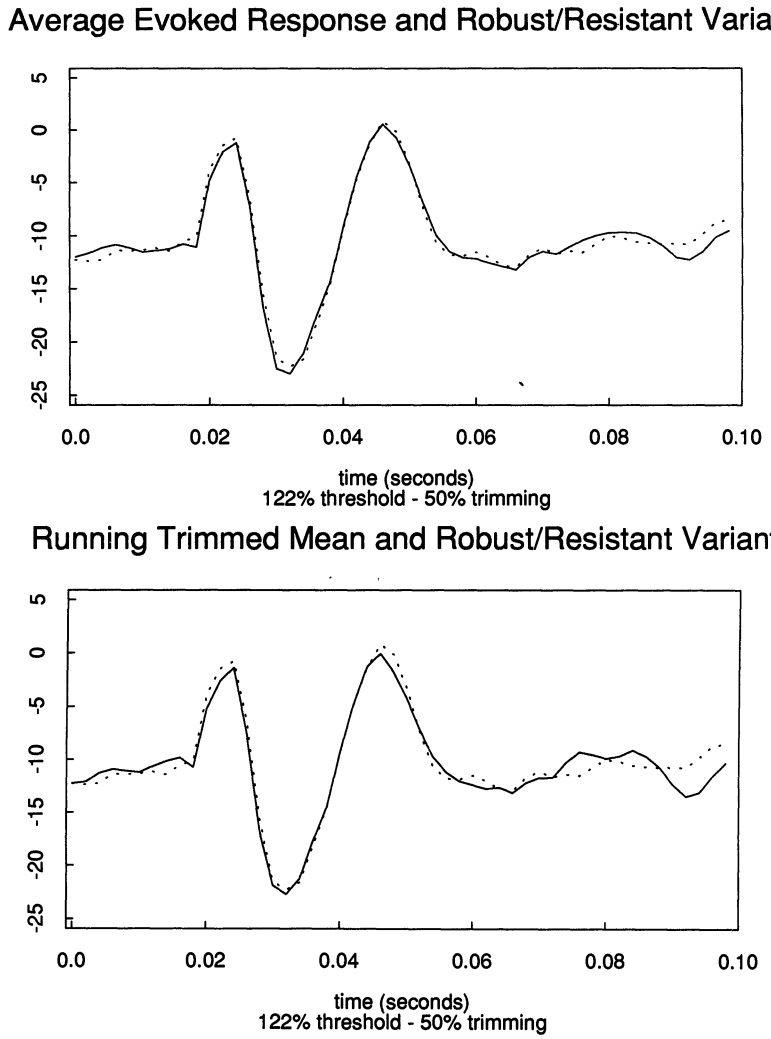

FIG. 23. *The upper graph compares the average evoked response with the* 50% *trimmed mean for the data taken at* 122% *of a threshold stimulation value. The lower graph contrasts the* 50% *trimmed mean statistic with a value computed recursively.*

resistant level in the case of elementary statistics. The fact that the trimming had such little effect on the final answer suggests that there were no substantial outlying curves in the data set. Had a curve been far removed from the rest, then it would have been rejected from the average. It is to be remarked that in this case of present concern, whole curves are being eliminated from the average, not just outlying points that some curves might have.

It is to be noted that a "real-time" version of such a trimmed mean may be computed; see Brillinger (1981a). This statistic is given recursively for $m =$

$1, 2, \ldots$ by

$$\hat{\rho}_{m+1} = \hat{\rho}_m - \frac{L}{m}\left(\frac{1}{\beta} - 1\right),$$

if $\|Y_{m+1} - \hat{a}_m\| \leq \hat{\rho}_m$, and

$$\hat{\rho}_{m+1} = \hat{\rho}_m + L/m$$

otherwise, and by

$$\hat{a}_{m+1}(s) = \hat{a}_m(s) + \frac{1}{\beta m}(Y_{m+1}(s) - \hat{a}_m(s)),$$

if $\|Y_{m+1} - \hat{a}_m\| \leq \hat{\rho}_m$, and

$$\hat{a}_{m+1}(s) = \hat{a}_m(s)$$

otherwise. (In preparing a worked example, it was found more convenient in the choice of L to replace ρ by its logarithm.) The lower graph of Figure 23 gives the result for the same data as that of the upper graph. The algorithm was run setting $\hat{a}_1(s) = Y_1(s)$ and $L = 0.15$. The real-time estimate, given by the solid line, has performed virtually as well as the dead-time estimate in this case. One can remark again that had there been some highly dissimilar curves present, then this estimate would have differed from the sample average. Following the advice sometimes given in connection with resistant regression estimates, it would seem sensible to compute both the ordinary and the resistant forms. If the two are similar, then there is no difficulty. If the two differ noticeably, then the situation should be examined in some detail.

Brillinger (1979) proposed the preceding techniques and various others. Brillinger (1981a, b) were based on that lecture and cover some other statistical problems arising from evoked-response methods. Tukey (1978) also addresses statistical issues and proposes some procedures.

15. A confirmed (Fourier) inference. Muscle cells are electrochemical devices. If the chemical acetylcholine is applied at the neuromuscular junction, measurable voltage fluctuations result. Specifically, acetylcholine release causes postsynaptic membrane channels to open leading to voltage fluctuations. Katz and Miledi (1971, 1972) measured voltage fluctuations associated with this phenomenon and found that the power spectrum could be approximated by the functional form $\alpha/(\beta^2 + \lambda^2)$. [An example of the fit of this function to such data and a description of a fitting procedure may be found in Bevan, Kullberg and Rice (1979).] They proposed the model

$$Y(t) = \sum_m a(t - \sigma_m),$$

with the σ_m points of a Poisson process and with $a(t) = \exp\{-\beta t\}$. This $a(\cdot)$ function corresponds to the effectiveness of an open channel decaying exponentially and leads to a power spectrum of the indicated form. Katz and Miledi mentioned that the pulses might actually be rectangular of random duration, but they preferred to deal with the exponential form. Stevens (1972) proposed the

specific model

$$Y(t) = \sum_m a_m(t - \sigma_m),$$

also with $\{\sigma_m\}$ Poisson, but now with $a_m(t) = 1$ for $0 < t < T_m$ and $a_m(t) = 0$ otherwise. The T_m are independent exponentials of mean $1/\beta$ and correspond to the lengths of time that the channels are open. Stevens noted that this model also led to a power spectrum of the form $\alpha/(\beta^2 + \lambda^2)$. The models were indistinguishable with the data collected.

The problem was later resolved by improved experimental technique. Neher and Sakman (1976) developed a technique that allowed the opening and closings of individual channels to be seen. They found that the channels remained equally effective and open for time periods of varying lengths. The two proposed models could be distinguished.

Examples of single-channel data and the corresponding estimated power spectrum may be found in Lecar (1981). Jackson and Lecar (1979) present results confirming the exponential duration of the openings.

16. Other topics. Spatial–temporal data are commonly collected by neuroscientists. One form is the electroencephalogram recorded by an array of sensors on the scalp. Figure 19 presented an example of data collected for the olfactory system of the rabbit. The stimulus was release of the odor ethylacetate. An 8×8 array of electrodes was imbedded in the animal. The data, already presented in Figure 19, give the responses for the sensors at the positions with x-coordinates 2, 4, 6 and 8 and y-coordinates 1, 3, 5 and 7 of Figure 23. One procedure that Freeman has found helpful for understanding this type of data is the computing of empirical orthogonal functions; see Freeman (1980). Figure 24 gives an example. These results are derived by stacking the responses into a matrix $\mathbf{X}$ with rows corresponding to sensor and columns to time, and then computing the singular value decomposition $\mathbf{X} = \mathbf{UDV}^\tau$, of that matrix. The $\mathbf{U}$ for a particular component, say the first, are then plotted versus sensor location as in the upper graph of Figure 23. The $\mathbf{V}$ values are similarly plotted versus time and appear in the lower graph. The results of Figure 23 are based on 64 series, not just the 16 of Figure 19. The contour plot suggests the presence of a focus of activity. The time-series component elicited may be seen lurking in the individual responses of Figure 19. (It may be mentioned that meteorologists have long computed empirical orthogonal functions for spatial–temporal data and used them in forecasting; see Lorentz (1956), for example. A number of other references are given in Jolliffe (1986).]

Childers has also made use of array data in studying the neural system. In Childers (1977), he estimates the frequency–wavenumber spectrum for responses evoked by visual stimuli (light flashes) in the human EEG. He was concerned with estimating the speed and direction of propagating waves. In the paper cited he first notes an apparent high-velocity wave. After this wave has been "removed," he notes the presence of a pair of waves moving in opposite directions. His research is directed at developing a diagnostic procedure for various visual disorders and obtaining insight concerning how the visual system functions.

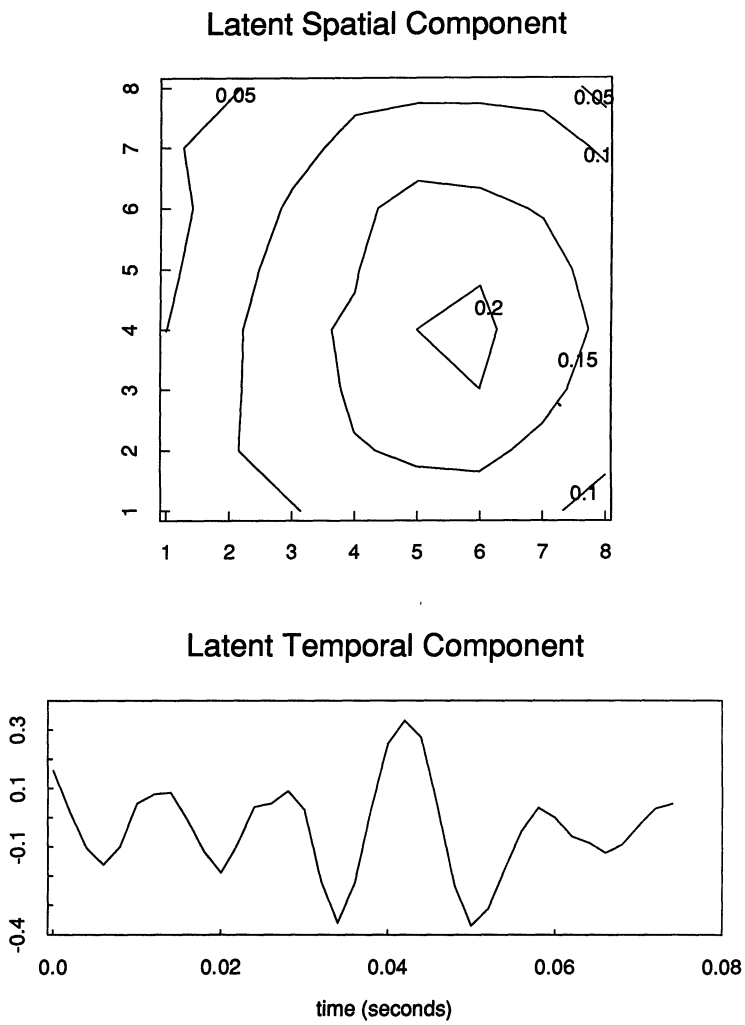

FIG. 24. *The results of a singular value decomposition of the full set of the data from which the bursts of Figure* 20 *were taken. The values graphed are for the first components. The axes in the upper graph give spatial location.*

The decaying cosine model of Section 2 has also found a use in neurophysiology. In his work with the olfactory system, Freeman (1972, 1975, 1979) found that the average evoked response could be well fitted by the sum of a few decaying cosine terms. He developed a model involving spike-to-wave conversion, involving collections of constant coefficient second-order differential equations, involving feedforward and feedback and involving wave-to-spike conversion. He employed nonlinear regression in the time domain to estimate the unknowns. In one case, involving two cosines, he was led to view the stronger wave as representing intracortical negative feedback and the weaker as representing a

second feedback loop. Of interest in this type of work is what happens to the frequencies and the decay rates when the experimental conditions are altered. A second reference to decaying cosines is Childers and Pao (1972). They consider the model

$$Y(t) = \sum_k \alpha_k t \exp\{-\beta_k t\}\cos(\gamma_k t + \delta_k) + \varepsilon(t), \qquad t > 0,$$

for visual evoked responses monitored over the occipital region. In particular, they study the data by complex demodulation.

Brief reference will be made to several other topics. Dumermuth, Huber, Kleiner and Gasser (1971) estimate the bispectrum of human EEGs. de Weerd and Kap (1981) discuss the computation of some time-varying quantities. Marmarelis and Naka (1974) consider the case of biological systems with several inputs. An extreme case of this occurs when the input is varying in both time and space. This circumstance is considered in Yasui, Davis and Naka (1979). The book by Marmarelis and Marmarelis (1978) goes into great detail concerning the identification of systems that are polynomial and time invariant in the input. They emphasize the advantages resulting from employing a Gaussian white-noise input. The dedication of the book is worth mentioning—"To an ambitious new breed: SYSTEMS PHYSIOLOGISTS".

Another area of research activity has been that of control. The works by Poggio and Reichardt (1981) and Wehrhahn, Poggio and Bulthoff (1982) may be noted. They are concerned with data that are three-dimensional trajectories.

17. Discussion. As the examples presented indicate, a broad range of data types arise in the neurosciences. Furthermore, data are collected at both the micro and macro level. The procedures developed often have the opportunity to move on to direct clinical use.

It is particularly interesting to note the evolution of the analysis in the case of the neuronal signaling analysis as presented in Sections 11 and 13. One can recognize the stages of (1) (feature) description; (2) correlation/association; (3) (ad hoc) regression; (4) conceptual model. These stages are usual in many elementary situations.

The field of neurophysiology has the satisfying aspect that in many cases controlled laboratory experiments are possible and repeatable. Furthermore, there are opportunities for the design of experiments. In the field, statistics has been seen to provide techniques for model formation and validation, for measuring uncertainty in conclusions and for addressing questions of causality. Statistical techniques have led to insight concerning the underlying physiology. In this connection it seems important to note the following proviso of my collaborator J. P. Segundo, however, "... . The maxim of all of the above is that the power of available mathematics (and of the instrumentation that implements them) should be used exhaustively, guided by an unflagging biological realism, mistrustful and stubborn, and keeping in mind that the ultimate goal is understanding in strictly biological terms." [See Segundo (1984), page 294.]

It seems likely that in the neurosciences, more often than not, notable advances will come from the carrying out of novel experiments, rather than from

novel analytic methods. Experiments will be carried out measuring things at new orders of smallness. More complex stimuli will be invoked. Nonlinear systems will be the norm. Neural networks will be a major concern. Luckily, for us statisticians, digital computers have become common in the laboratory and this seems to be bringing a move toward quantization of other aspects of the work beyond the simple recording of the data.

18. Update. The analysis of single ion-channel data, briefly referred to in Section 15, has become a whole industry. Models with several states are now routinely fitted. References include Coloquhoun and Hawkes (1983), Labarca, Rice, Fredkin and Montal (1986) and Milne, Edeson and Madsen (1986). Extending the work of Section 13, Brillinger (1986) presents a number of examples of the maximum likelihood fitting of a neural model employing corresponding spike train input and output data. Smith and Chen (1986) study a more complicated neural model. The chirp signal was propounded as being of substantial importance in seismic exploration. Some use of it has been made recently in physiological studies. In Norcia and Tyler (1985), a 10-s spatial frequency sweep stimulus is employed and the corresponding visual evoked potential measured. Th. Gasser and collaborators have now carried out a substantial number of statistical and substantive analyses of evoked responses. We mention in particular the papers by Mocks, Tuan and Gasser (1984), Gasser, Mocks, Kohler and de Weerd (1986) and Gasser, Mocks and Kohler (1986). Finally, we note that Grajski, Breiman, di Prisco and Freeman (1986) apply modern classification procedures to study the effects of applying different odors on the olfactory bulb EEGs of rabbits and that Gevins, Morgan, Bressler, Cutillo, White, Illes, Greer, Doyle and Zeitlin (1987) relate human performance accuracy to brain electrical patterns just before a task.

IV. Concluding remarks. In this article we have presented a number of examples, drawn mainly from our personal experience, showing the use of the same statistical technique in the rather separate sciences of seismology and neurophysiology. It now seems appropriate to ask what, if anything, have the *three* sciences—statistics, seismology, neurophysiology—gained from each other as a result of connections even though they are indirect? Having in mind a broader class of examples than those discussed in this paper, one can say that: (i) statistics is richer for having been led to develop and study various novel methods to handle specific problems arising in seismology or neurophysiology; (ii) both seismology and neurophysiology are the richer for the other's field having generated a problem for the statistician to abstract sufficiently that the result's applicability to their field became apparent; (iii) either seismology or neurophysiology benefit from a statistical formulation because various of their problems seem necessarily to need to be stated in terms of probabilities (e.g., neither neuron firings nor earthquakes seem deterministic) and because these fields need procedures to validate results and to fit conceptual models. That the methods of statistics can lead to important insight and understanding in substantive problems seems agreed.

It may be remarked that the applicability of statistical procedures to these two substantive fields has further grown in direct consequence of their move to greater quantification and digital data collection. The data sets analyzed were of high quality. The fact that the analyses were informative to an extent here bodes well for the use of such techniques in fields with data of lesser quality. I need to remark how crucial, in working with the data sets discussed, I have found it to be to plot the data in its original form. Something special seemed to be learned in each case from doing so. This is why for the various analyses, I have sought to provide data plots as parts of the presentation.

The reader will have noted that some of the analyses were time-side and some were frequency-side. Each domain has its advantages. It seems worth pointing out specifically that stationarity was not required for some of the frequency-side procedures. It would seem that most time series and point process situations would benefit from carrying out simultaneous time-side and frequency-side analyses.

On review it may be seen that the techniques employed for time-series data and for point process data in many cases are not that different. Brillinger (1978) presents some comparative discussion of the techniques for the two cases. Our presentation is somewhat remiss in the seismological case in not presenting some worked examples of auto- and cross-intensity estimation. Examples could have been provided.

It should be apparent that major data management and computational efforts were required in the derivation of all results presented. I have been impressed by the way that the neuroscientists could turn to their lab book kept during the experiments and pull out crucial details, sometimes many years after the experiments had been completed. My analyses also extend over many years now. I have found it very useful recently to maintain a "Readme" file in the various computer directories for the data sets, wherein I list what the various programs do, future wishes concerning the programs and all of the things that I think I will never forget.

Turning to thoughts concerning developments to come, it seems that the future will see many of the traditional statistical techniques extended to apply to datum of more complicated forms—specifically, to curves, moving surfaces, point clouds and the like. It seems that techniques developed in one field will continue to be transferred (by statisticians?) to other fields. For example, I expect to see the results developed by neuroscientists for arrays on a curved surface (the skull) to be taken up by the seismologists as they need to take specific note of the Earth's curvature. If I have found anything lacking in our current toolkit of statistical methods and devices, it is a collection of techniques that suggest what to do next when a model fails a validation check. Perhaps the future will see such techniques developed in an organized fashion.

Acknowledgments. One wonders if it is possible to work meaningfully on problems from substantive areas unless one is in close contact with researchers of those areas. I doubt it. I can say that these lectures would never have come about, but for the help and encouragement that my collaborators B. A. Bolt and

J. P. Segundo have provided through the years. I cannot thank them enough. The influence of my doctoral supervisor, J. W. Tukey, should be readily apparent as well. In particular, I remember his saying, when I was a student, "One cannot consult with a chemist, unless one becomes a chemist." This has certainly been my own experience with respect to the parts of neurophysiology and seismology on which I have worked. The reverse has also proved to be the case, namely, collaborators from those fields have had to learn parts of statistics as well as I knew them. I also must thank my students who have worked with me on various of the problems described. In particular, I mention M. Folledo and R. Ihaka, whose theses have been detailed.

Finally, I would like to thank J. Rice and R. Zucker for providing some neurophysiological references. I would like to thank N. Abrahamson and R. Daragh for providing the Smart 1 data sets. I would like to thank W. J. Freeman and K. Grajski for providing the rabbit olfactory system data. I would like to thank Finbarr O'Sullivan, the Associate Editor, and the referees for their helpful suggestions for improvement of the initial draft paper.

REFERENCES

ABRAHAMSON, N. A. (1985). Estimation of seismic wave coherency and rupture velocity using the Smart 1 motion array recordings. Ph.D. thesis, Univ. California, Berkeley.

AIDLEY, D. J. (1978). *The Physiology of Excitable Cells*. Cambridge Univ. Press, Cambridge.

AKI, K. and RICHARDS, P. G. (1980). *Quantitative Seismology*. Freeman, San Francisco.

BERGAMINI, L., BERGAMASCO, B., FRA, L., GANDIGLIO, G. and MUTANI, R. (1967). Diagnostic recording of somato-sensory cortical evoked potentials in man. *Electroenceph. Clin. Neurophysiol.* **22** 260–262.

BEVAN, S., KULLBERG, R. and RICE, J. (1979). An analysis of cell membrane noise. *Ann. Statist.* **7** 237–257.

BIEDENBACH, M. A. and FREEMAN, W. J. (1965). Linear domain of potentials for the prepyriform cortex with respect to stimulus parameters. *Experimental Neurology* **11** 400–417.

BOLT, B. A. (1976). Abnormal seismology. *Bull. Seismol. Soc. Amer.* **66** 617–623.

BOLT, B. A. (1982). *Inside the Earth*. Freeman, San Francisco.

BOLT, B. A. and BRILLINGER, D. R. (1979). Estimation of uncertainties in eigenspectral estimates from decaying geophysical time series. *Geophys. J. Roy. Astron. Soc.* **58** 593–603.

BOLT, B. A. and MARUSI, A. (1962). Eigenvibrations of the Earth observed at Trieste. *Geophys. J. Roy. Astron. Soc.* **6** 299–311.

BOLT, B. A., TSAI, Y. B., YEH, K. and HSU, M. K. (1982). Earthquake strong motions recorded by a large near-source array of digital seismographs. *Earthquake Engin. Structural Dynamics* **10** 561–573.

BOORE, D. M. (1977). The motion of the ground in earthquakes. *Scientific American* **237** 68–78.

BRILLINGER, D. R. (1975a). *Time Series: Data Analysis and Theory*. Holt, Rinehart and Winston, New York.

BRILLINGER, D. R. (1975b). Statistical inference for stationary point processes. In *Stochastic Processes and Related Topics* (M. L. Puri, ed.) **1** 55–99. Academic, New York.

BRILLINGER, D. R. (1978). Comparative aspects of the study of ordinary time series and of point processes. In *Developments in Statistics* (P. R. Krishnaiah, ed.) **1** 33–133. Academic, New York.

BRILLINGER, D. R. (1979). Some statistical methods in neurophysiology. Special Invited Lecture, Western Regional Meeting, Institute of Mathematical Statistics, Los Angeles.

BRILLINGER, D. R. (1981a). Some aspects of the analysis of evoked response experiments. In *Statistics and Related Topics* (M. Csörgő, D. A. Dawson, J. N. K. Rao and A. K. Md. E. Saleh, eds.) 155–168. North-Holland, Amsterdam.

BRILLINGER, D. R. (1981b). The general linear model in the design and analysis of evoked response experiments. *J. Theoret. Neurobiol.* **1** 105–119.

BRILLINGER, D. R. (1983). The finite Fourier transform of a stationary process. In *Handbook of Statistics* (D. R. Brillinger and P. R. Krishnaiah, eds.) **1** 21–37. North-Holland, Amsterdam.

BRILLINGER, D. R. (1985). A maximum likelihood approach to frequency–wavenumber analysis. *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-33** 1076–1085.

BRILLINGER, D. R. (1986). Maximum likelihood analysis of spike trains of interacting nerve cells. Technical Report 80, Dept. Statistics, Univ. California, Berkeley.

BRILLINGER, D. R., BRYANT, H. L. and SEGUNDO, J. P. (1976). Identification of synaptic interactions. *Biol. Cybernetics* **22** 213–228.

BRILLINGER, D. R. and IHAKA, G. R. (1982). Maximum likelihood estimation of source parameters. *Earthquake Notes* **53** 39–40.

BRILLINGER, D. R. and SEGUNDO, J. P. (1979). Empirical examination of the threshold model of neuron firing. *Biol. Cybernetics* **35** 213–220.

BRILLINGER, D. R., UDIAS, A. and BOLT, B. A. (1980). A probability model for regional focal mechanism solutions. *Bull. Seismol. Soc. Amer.* **70** 149–170.

BRUNE, J. N. (1970 / 1971). Tectonic stress and the spectra of seismic shear waves from earthquakes. *J. Geophys. Res.* **75** 4997–5009. Correction **76** 5002.

BRYANT, H. L., MARCOS, A. R. and SEGUNDO, J. P. (1973). Correlations of neuronal spike discharges produced by monosynaptic connections and by common inputs. *J. Neurophysiol.* **36** 205–225.

BRYANT, H. L. and SEGUNDO, J. P. (1976). Spike initiation by transmembrane current: A white noise analysis. *J. Physiol.* **260** 279–314.

BUFORN, E. (1982). Estudio estadistico de la direccion de esfuerzos principales en terremotos. Doctoral thesis, Universidad Complutense de Madrid.

BULLEN, K. E. and BOLT, B. A. (1985). *An Introduction to the Theory of Seismology*. Cambridge Univ. Press, Cambridge.

BURNS, B. D. (1968). *The Uncertain Nervous System*. Arnold, London.

BYERLY, P. (1926). The Montana earthquake of June 28, 1925. *Bull. Seismol. Soc. Amer.* **16** 209–265.

CAPON, J. (1969). High-resolution frequency–wavenumber spectrum analysis. *Proc. IEEE* **57** 1408–1418.

CHILDERS, D. G. (1977). Evoked responses: Electrogenesis, models, methodology, and wavefront reconstruction and tracking analysis. *Proc. IEEE* **65** 611–626.

CHILDERS, D. G. and PAO, M. (1972). Complex demodulation for transient wavelet detection and extraction. *IEEE Trans. Audio Electroacoust.* **AU-20** 295–308.

CHIU, S.-T. (1986). Statistical estimation of the parameters of a moving source from array data. *Ann. Statist.* **14** 559–578.

COLOQUHOUN, D. and HAWKES, A. G. (1983). The principles of the stochastic interpretation of ion-channel mechanisms. In *Single-Channel Recording* (B. Sakman and E. Neher, eds.) 135–175. Plenum, New York.

COOK, C. E. and BERNFIELD, M. (1967). *Radar Signals*. Academic, New York.

COPAS, J. B. (1983). Plotting p against x. *Appl. Statist.* **32** 25–31.

DAHLEN, F. A. (1982). The effect of data windows on the estimation of free oscillation parameters. *Geophys. J. Roy. Astron. Soc.* **69** 537–549.

DE WEERD, J. P. C. and KAP, J. I. (1981). Spectro–temporal representations and time-varying spectra of evoked potentials. *Biol. Cybernetics* **41** 101–117.

DONOHO, D. L. (1981). On minimum entropy deconvolution. In *Applied Time Series Analysis* (D. F. Findley, ed.) **2** 565–608. Academic, New York.

DONOHO, D. L., CHAMBERS, R. E. and LARNER, K. L. (1986). Robust estimation of residual statics corrections. Preprint, Dept. Statistics, Univ. California, Berkeley.

DUMERMUTH, G., HUBER, P. J., KLEINER, B. and GASSER, TH. (1971). Analysis of the interrelations between frequency bands of the EEG by means of the bispectrum. A preliminary study. *Electroenceph. Clin. Neurophysiol.* **31** 137–148.

FOLLEDO, M. (1983). Robust/resistant methods in the estimation of the evoked response curve. Ph.D. thesis, Univ. California, Berkeley.

FREEMAN, W. J. (1972). Linear analysis of the dynamics of neural masses. *Ann. Rev. Biophys. Bioeng.* **1** 225–256.

FREEMAN, W. J. (1975). *Mass Action in the Nervous System*. Academic, New York.

FREEMAN, W. J. (1979). Nonlinear dynamics of paleocortex manifested in the olfactory EEG. *Biol. Cybernetics* **35** 21–37.

FREEMAN, W. J. (1980). Measurement of corticalevoked potentials by decomposition of their wave forms. *J. Cybernetics Inform. Sci.* **2** 44–56.

FREEMAN, W. J. and SCHNEIDER, W. (1982). Changes in spatial patterns of rabbit olfactory EEG with conditioning to odors. *Psychophysiology* **19** 44–56.

GASSER, TH., MOCKS, J. and KOHLER, W. (1986). Amplitude probability distribution of noise for flash-evoked potentials and robust response estimates. *IEEE Trans. Biomed. Eng.* **BME-33** 579–583.

GASSER, TH., MOCKS, J., KOHLER, W. and DE WEERD, J. P. C. (1986). Performance and measures of performance estimators of brain potentials using real data. *IEEE Trans. Biomed. Eng.* **BME-33** 949–956.

GEVINS, A. S., MORGAN, N. H., BRESSLER, S. L., CUTILLO, B. A., WHITE, R. M., ILLES, J., GREER, D. S., DOYLE, J. C. and ZEITLIN, G. M. (1987). Human neuroelectric patterns predict performance accuracy. *Science* **235** 580–585.

GIANNAKIS, G. B. and MENDEL, J. M. (1986). Tomographic wavelet estimation via higher-order statistics. *Proc. Fifty-Sixth Internat. Conf. Soc. Explor. Geophys.* 512–514. Houston, Texas.

GLASER, E. M. and RUCHKIN, D. S. (1976). *Principles of Neurobiological Signal Analysis*. Academic, New York.

GRAJSKI, K. A., BREIMAN, L., DI PRISCO, G. V. and FREEMAN, W. J. (1986). Classification of EEG spatial patterns with a tree-structured methodology: CART. *IEEE Trans. Biomed. Eng.* **BME-33** 1076–1086.

HANSEN, R. A. (1982). Simultaneous estimation of terrestrial eigenvibrations. *Geophys. J. Roy. Astron. Soc.* **70** 155–172.

HASAN, T. (1982). Nonlinear time series regression for a class of amplitude modulated cosinusoids. *J. Time Series Anal.* **3** 109–122.

HAUBRICH, R. A. (1965). Earth noise, 5 to 500 millicycles per second. *J. Geophys. Res.* **70** 1415–1427.

HOLDEN, A. V. (1976). *Models for the Stochastic Activity of Neurones*. Springer, New York.

HOPPER, M. J. (1980). *Harwell Subroutine Library*. Atomic Energy Research Establishment, Harwell, England.

HUDSON, J. A. (1981). Mathematics for seismology. *J. Inst. Math. Appl.* **17** 34–39.

IHAKA, R. (1985). Ruaumoko. Ph.D. thesis, Dept. Statistics, Univ. California, Berkeley.

JACKSON, M. B. and LECAR, H. (1979). Single postsnyaptic channel currents in tissue cultured muscle. *Nature* **282** 863–864.

JEFFREYS, H. (1977). *Collected Papers of Sir Harold Jeffreys on Geophysics and Other Sciences*. Gordon and Breach, London.

JEFFREYS, H. and BULLEN, K. E. (1940). *Seismological Tables*. British Association, Gray-Milne Trust.

JOLLIFFE, I. T. (1986). *Principal Component Analysis*. Springer, New York.

KATZ, B. and MILEDI, R. (1971). Further observations on acetylcholine noise. *Nature* **232** 124–126.

KATZ, B. and MILEDI, R. (1972). The statistical nature of the acetylcholine potential and its molecular components. *J. Physiol. (London)* **224** 665–700.

KITAGAWA, G. and GERSCH, W. (1985). A smoothness priors time-varying AR coefficient modeling of nonstationary covariance time series. *IEEE Trans. Automat. Control* **AC-30** 48–56.

LABARCA, P., RICE, J. A., FREDKIN, D. R. and MONTAL, M. (1986). Kinetic analysis of channel gating: Application to the cholinergic receptor channel and the chloride channel from *Torpedo californica*. *Biophysical J.* **47** 469–478.

LAPWOOD, E. R. and USAMI, T. (1981). *Free Oscillations of the Earth*. Cambridge Univ. Press, Cambridge.

LECAR, H. (1981). Single-channel conductances and models of transport. In *The Biophysical Approach to Excitable Systems* (W. J. Adelman and D. E. Goldman, eds.) 109–121. Plenum, New York.

LII, K. S. and ROSENBLATT, M. (1982). Deconvolution and estimation of transfer function phase and coefficients for nongaussian linear processes. *Ann. Statist.* **10** 1195–1208.

LINDBERG, C. R. (1986). Multiple taper spectral analysis of terrestrial free oscillations. Ph.D. thesis, Univ. California, San Diego.

LORENTZ, E. H. (1956). Empirical orthogonal functions and statistical weather prediction. Report 1, Statistical Forecasting Project, Dept. Meteorology, MIT, Cambridge, Mass.

MARMARELIS, P. Z. and MARMARELIS, V. Z. (1978). *Analysis of Physiological Systems*. Plenum, New York.

MARMARELIS, P. Z. and NAKA, N.-I. (1974). Identification of multi-input biological systems. *IEEE Trans. Biomed. Eng.* **BME-21** 88–101.

MENDEL, J. M. (1983). *Seismic Deconvolution: An Estimation-Based Approach*. Academic, New York.

MENDEL, J. M. (1986). Some modeling problems in reflection seismology. *IEEE ASSP Magazine* 4–17.

MILNE, R. K., EDESON, R. O. and MADSEN, B. W. (1986). Stochastic modeling of a single ion channel: Interdependence of burst length and number of openings per burst. *Proc. Roy. Soc. London Ser. B* **227** 83–102.

MOCKS, J., TUAN, P. D. and GASSER, TH. (1984). Testing for homogeneity of noisy signals evoked by repeated stimuli. *Ann. Statist.* **12** 193–209.

MOORE, G. P., PERKEL, D. H. and SEGUNDO, J. P. (1966). Statistical analysis and functional interpretation of neuronal spike data. *Ann. Rev. Physiol.* **28** 493–522.

NEHER, E. and SAKMAN, B. (1976). Single-channel currents recorded from membrane of denervated frog muscle fibres. *Nature* **260** 799–801.

NEYMAN, J. (1974). A view of biometry: An interdisciplinary domain concerned with chance mechanisms operating in living organisms; illustration: Urethan carcinogenis. In *Reliability and Biometry* 183–201. SIAM, Philadelphia.

NORCIA, A. M. and TYLER, C. W. (1985). Spatial frequency sweep VEP: Visual acuity during the first year of life. *Vision Res.* **25** 1399–1408.

OGATA, Y. (1983). Likelihood analysis of point processes and its applications to seismological data. *Bull. Internat. Statist. Inst.* **50** (2), 943–961.

OGATA, Y. and KATSURA, K. (1986). Point-process models with linearly parameterized intensity for application to earthquake data. In *Essays in Time Series and Allied Processes* (J. Gani and M. B. Priestley, eds.) 291–310. Applied Probability Trust, Sheffield.

O'SULLIVAN, F. (1986). A statistical perspective on ill-posed inverse problems (with discussion). *Statist. Sci.* **1** 502–527.

POGGIO, T. and REICHARDT, W. (1981). Visual fixation and tracking by flies: Mathematical properties of simple control systems. *Biol. Cybernetics* **40** 101–112.

RAPIN, I. and GRAZIANI, L. J. (1967). Auditory-evoked responses in normal, brain-damaged and deaf infants. *Neurology* **17** 881–891.

ROBINSON, E. A. (1983). *Migration of Geophysical Data*. International Human Resources Development Corp., Boston.

SAGALOVSKY, B. D. (1982). Maximum likelihood and related estimation methods in point processes and point process systems. Ph.D. thesis, Univ. California, Berkeley.

SCHEIMER, J. and LANDERS, T. E. (1974). Short-period coda of a local event at LASA. *Semiannual Technical Summary* 42–44. Lincoln Lab., MIT, Cambridge, Mass.

SEGUNDO, J. P. (1984). La neurofisiologica: algunos supuestos y bases, recovecos e implicaciones. UNAM, Mexico.

SHUMWAY, R. H. and DER, Z. A. (1985). Deconvolution of multiple time series. *Technometrics* **27** 385–393.

SMITH, C. E. and CHEN, C.-L. (1986). Serial dependency in neural point processes due to cumulative after hyperpolarization. Statistics Mimeo Series 1691, North Carolina State Univ.

STEVENS, C. F. (1972). Inferences about membrane properties from electrical noise measurements. *Biophys. J.* **12** 1028–1047.

TJØSTHEIM, D. (1981). Multidimensional discrimination techniques—theory and application. In *Identification of Seismic Sources—Earthquake or Underground Explosion* (E. S. Husebye and S. Mykkeltveit, eds.) 663–694. Reidel, Dordrecht.

TUKEY, J. W. (1978). A data analysts's comments on a variety of points and issues. In *Event-Related Brain Potentials in Man* (E. Callaway, P. Tueting and S. H. Koslow, eds.) 139–154. Academic, New York.

UDIAS, A., MUNOZ, D. and BUFORN, E. (1985). Mecanismo de los terremotos y tectonica. Editorial de la Universidad complutens, Madrid.

VERE-JONES, D. and SMITH, E. G. C. (1981). Statistics in seismology. *Comm. Statist. A—Theory Methods* **10** 1559–1585.

WALD, A. (1952). On the principles of statistical inference. *Notre Dame Mathematical Lectures*, No. 1. Notre Dame, Ind.

WATERS, K. H. (1978). *Reflection Seismology*. Wiley, New York.

WEHRHAHN, C., POGGIO, T. and BULTHOFF, H. (1982). Tracking and chasing in houseflies. (*Musca*) *Biol. Cybernetics* **45** 123–130.

WIENER, N. (1930). Generalized harmonic analysis. *Acta Math.* **55** 117–258.

WOOD, L. C. and TREITEL, S. (1975). Seismic signal processing. *Proc. IEEE* **63** 649–661.

YASUI, S., DAVIS, W. and NAKA, K.-I. (1979). Spatio–temporal receptive field measurement of retinal neurons by random pattern stimulation and cross correlation. *IEEE Trans. Biomed. Eng.* **BME-26** 263–272.

ZADRO, M. B. and CAPUTO, M. (1968). Spectral, bispectral analysis and Q of the free oscillations of the Earth. *Nuovo Cimento Suppl.* **6** 67–81.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720