Ann. Inst. Statist. Math.
Vol. 44, No. 1, 13–25 (1992)

SIGNAL ESTIMATION USING STOCHASTIC VELOCITY MODELS AND IRREGULAR ARRAYS

P. J. THOMSON

Institute of Statistics and Operations Research, Victoria University of Wellington, P.O. Box 600, Wellington, New Zealand

(Received January 17, 1989; revised January 7, 1991)

Abstract. Consider the situation where a plane wave signal is received by a spatial arrangement of recorders. Information derived from observations on such a process can be used to determine the speed and direction of the signal together with properties of the medium through which the signal is being propagated. Certain models for the case where the signal velocity can be regarded as stochastic and where the array is irregular are investigated and estimation procedures proposed. A major practical property of these models is that, unlike their deterministic counterparts, coherence decays to zero as distance between recorders increases.

Key words and phrases: Fourier analysis, array estimation, velocity estimation, space-time spectral analysis, random media, random arrays, frequency wavenumber processing.

1. Introduction

The estimation of the speed, direction and other properties of a common signal from observations collected by an array of recorders is a problem of general interest relevant to many areas of science. The recorders could be seismometers measuring earth tremors, radio telescopes measuring the level of activity of a distant star, tide gauges used in oceanographic studies etc. As a consequence the problem has been extensively discussed by many authors and in many different contexts; see, for example, Capon (1969), Hinich and Shaman (1972), Cameron and Hannan (1978), Hinich (1981), Thomson (1982), Shumway (1983), Brillinger (1985), Cameron and Thomson (1985), Hannan and Thomson (1988), and Ziskind and Wax (1988) to name but a few. The references listed in these papers provide a more extensive bibliography of the field.

Typically the signal is assumed to be a plane waveform, as would be the case if the signal source was located far from the array relative to its size. What is received at one recorder will, modulo noise and the effects of the medium, have been received at other recorders before or will be received some time lag or delay later. Knowledge of these delays enables the speed and direction of the wave to be determined.

The array could be three dimensional, but the more usual type of array would consist of a linear or two-dimensional arrangement of recorders with a known coordinate structure receiving a signal propagating along the line or in the plane respectively. Situations where a signal is received by a lower dimensional arrangement of recorders are clearly possible, but in that case more will need to be known about the incoming signal, such as its speed, if both speed and direction are to be resolved.

In the case where the medium is dispersive and the various frequency components that make up the signal travel at different velocities, velocity is typically measured as a deterministic function of frequency. Such information together with relative attenuations provides information about the medium through which the signal is propagating.

This paper considers the case where the velocity of the signal is no longer a deterministic function of frequency, but is better regarded as stochastic in nature. This will typically come about because of varying sample paths and scattering due to inhomogeneities in the medium. In such cases, non-constant point velocities will result and travel times will vary about their expected value. A model for stochastic velocities of this kind is developed in Section 2. See also Sato (1984) and the references therein for related work concerning stochastic velocity models and random media. In the subsequent sections an estimation criterion for a regular array is proposed and the effects of various irregular array structures on the model and its estimation investigated.

2. A stochastic velocity model

For ease of exposition we restrict attention throughout to the simplest case where the space variable is one dimensional and the recorders are located on a line. Let $Y(x,t)$ be the observation at time t of a recorder located at position x and assume that

$$Y(x,t) = S(x,t) + N(x,t) \qquad (-\infty < x,t < \infty)$$

where $S(x,t)$ is the plane wave signal and $N(x,t)$ is noise. Set $S(t) = S(0,t)$ and assume that $S(t)$ is a stationary process with spectral density $f(\omega)$. Then $S(t)$ has spectral representation

$$S(t) = \mu + \int_{-\infty}^{\infty} e^{-it\omega} dZ(\omega)$$

where the spectral amplitudes $dZ(\omega)$ have mean zero and are orthogonal with mean square $E\{|dZ(\omega)|^2\} = f(\omega)d\omega$.

Consider the deterministic velocity case. If there is no dispersion and the signal propagates through the medium with velocity V, then $S(x,t) = S(t - L(x))$ where $L(x) = x/V$ is the time lag between the receipt of the signal at location 0 and location x. If the medium is dispersive and V is a function of ω then

$$S(x,t) = \mu + \int_{-\infty}^{\infty} e^{-i(t-L(x,\omega))\omega} dZ(\omega)$$

where $L(x,\omega) = x/V(\omega)$. Furthermore, if we allow for the possibility of attenuation then the model becomes

$$(2.1)\quad Y(x,t) = \mu + \int_{-\infty}^{\infty} e^{-i(t-L(x,\omega))\omega} a(x,\omega)dZ(\omega) + N(x,t) \qquad (-\infty < x, t < \infty)$$

where $a(x,\omega)$ describes the attenuation of the component at frequency ω.

Given sampled data from $Y(x,t)$ and suitable assumptions about the noise (typically $N(x,t)$ is assumed to be stationary in t and uncorrelated in x) various estimation procedures have been proposed for (2.1) and related models. Standard methods include delay and sum beamforming, Capon's high resolution method and the Hamon and Hannan method (see Capon (1969) and Hamon and Hannan (1974)). Of these, only Capon's method was designed for the case where $N(x,t)$ is stationary in space as well as time which restricts the $N(x,t)$ to have the same spectra, but does allow for spatial correlation. Moreover, as noted in Hinich (1981), delay and sum beamforming is equivalent to finding peaks in the frequency wavenumber spectrum (see also Hannan and Thomson (1988)). As the references given in Section 1 indicate, there are many other methods that have been developed, most tailored to particular variants of the model (2.1). However all these procedures have in common the assumption that $L(x,\omega) = x/V(\omega)$ where $V(\omega)$ is the deterministic velocity of the component of the signal at frequency ω.

Now consider the stochastic velocity case. Here we need a model to account for varying sample paths, scattering etc. We choose to model $L(x,\omega)$ as a process of stationary, independent increments where $L(x'+x,\omega) - L(x',\omega)$, the time taken for the component at frequency ω to travel a distance x from x', has mean $x/V(\omega)$ and characteristic function

$$E\{e^{i(L(x'+x,\omega)-L(x',\omega))\theta}\} = \psi(\theta,\omega)^x \qquad (x > 0).$$

(For a discussion of such processes and their properties see, for example, Ash and Gardner (1975), p. 195.) Thus $\psi(\theta,\omega)$ is the characteristic function of the time taken to travel unit distance. This assumption implies that travel times over disjoint space intervals are independent and are identically distributed with mean $x/V(\omega)$ when these intervals have the same length x. Moreover $L(x,\omega)$ is assumed to be independent of $S(t)$ and $N(x,t)$ with $S(t)$, $N(x,t)$ mutually uncorrelated. In addition, it is assumed that $N(x,t)$ is not necessarily uncorrelated in space, but is stationary in space and time with spectral density $f_n(\lambda,\omega)$, and the attenuation $a(x,\omega)$ does not vary with x (equals $a(\omega)$ say) across the array. The latter means that $a(\omega)$ and $f(\omega)$ cannot be identified from the spectrum without further information. Hence we set $a(\omega) = 1$ or, equivalently, absorb $a(\omega)$ into the signal spectrum $f(\omega)$.

With these assumptions $Y(x,t) = S(x,t) + N(x,t)$ is now stationary in space and time. Moreover, $S(x,t)$ has mean μ and autocovariance function

$$\begin{aligned}(2.2)\qquad & E\{(S(x',t') - \mu)(S(x'+x,t'+t) - \mu)\} \\ &= \int_{-\infty}^{\infty} e^{it\omega} E\{e^{-i(L(x'+x,\omega)-L(x',\omega))\omega}\} f(\omega)d\omega \\ &= \int_{-\infty}^{\infty} e^{i(t\omega - x\theta(\omega))} \rho(\omega)^{|x|} f(\omega)d\omega \qquad (-\infty < x, t < \infty)\end{aligned}$$

where $\rho(\omega)$ and $\theta(\omega)$ are the modulus and argument respectively of $\psi(\omega, \omega)$ with $\rho(\omega) = \rho(-\omega)$, $\theta(\omega) = -\theta(-\omega)$. Note that $\theta(\omega) = \omega/V(\omega)$ if the distribution corresponding to $\psi(\theta, \omega)$ is symmetric about its mean. In particular, observe that the stochastic velocity model has an autocovariance function that decays to zero as the distance between recorders increases. Such a situation typically occurs in practice. However the autocovariance function for the deterministic velocity model never decays to zero, irrespective of how far apart the recorders are placed. Thus this property is a major practical advantage of the stochastic velocity model over the deterministic velocity model.

The space-time spectral density of $Y(x,t)$ is now given by $f_s(\lambda,\omega) + f_n(\lambda,\omega)$ where

$$f_s(\lambda,\omega) = f(\omega)\frac{g(\omega)}{\pi(g(\omega)^2 + (\lambda + \theta(\omega))^2)} \qquad (-\infty < \lambda, \omega < \infty) \tag{2.3}$$

and $g(\omega) = -\log \rho(\omega)$. For $\rho(\omega) < 1$, $f_s(\lambda,\omega)$ is a continuous time AR(1) spectral density in λ centred at $\lambda = -\theta(\omega)$. Note that $f_s(\lambda,\omega)$ becomes $f(\omega)\delta(\lambda+\omega/V(\omega))$, $\delta(\cdot)$ the Dirac delta function, in those cases where $\rho(\omega) = 1$. This includes the important special case when the velocity function is deterministic with $L(x,\omega) = x/V(\omega)$ and $f_s(\lambda,\omega) = f(\omega)\delta(\lambda + \omega/V(\omega))$ for all λ, ω. Thus the stochastic velocity model has replaced the delta function of the deterministic velocity model by an AR(1) density. This spreading and attenuation of the peak along the line $\lambda = -\theta(\omega)$ is a direct consequence of the fact that travel time is no longer exactly proportional to distance, but fluctuates about a mean value.

3. Estimation for linear equispaced arrays

In this section we consider the stochastic velocity model introduced in Section 2 and restrict attention to the important case of a linear array of equispaced recorders. Irregular arrays will be dealt with in the later sections. Suppose that $Y(x,t)$ has been sampled at unit intervals in time and space with the time unit chosen so that there is no spectral aliasing in time. Then $f(\omega)$, $f_n(\lambda,\omega)$ are zero for $|\omega| > \pi$ and $Y(x,t)$ has space-time spectral density $f(\lambda,\omega) = \bar{f}_s(\lambda,\omega) + f_n(\lambda,\omega)$ where

$$\bar{f}_s(\lambda,\omega) = \frac{f(\omega)}{2\pi} \cdot \frac{1-\rho(\omega)^2}{|1-\rho(\omega)e^{i(\lambda+\theta(\omega))}|^2} \qquad (-\pi < \lambda, \omega < \pi) \tag{3.1}$$

and $f_n(\lambda,\omega)$ has (possibly) been replaced by its aliased form in λ. Analogous comments to those made following (2.3) can also be made here. In particular $\bar{f}_s(\lambda,\omega)$ is a discrete time AR(1) spectral density in λ centred at $\lambda = -\theta(\omega)$ (mod2π) for $\rho(\omega) < 1$ and is $f(\omega)\delta(\lambda + \omega/V(\omega))$ when $\rho(\omega) = 1$.

Now consider the simple case of fitting the stochastic velocity spectral model (3.1) over a narrow band of time frequencies B centered at some frequency of interest ω_0 $(0 < \omega_0 < \pi)$. This case arises in many practical situations where the data has been band-pass filtered by virtue of the frequency response of the recorder, or because the signal is band-limited and essentially monochromatic.

Assume that $L(x+1,\omega) - L(x,\omega)$, the time taken to travel unit distance, has a symmetric distribution about its mean so that $\theta(\omega) = \omega/V(\omega)$. Also assume that $f(\omega)$, $\rho(\omega)$ and $V(\omega)$ are (approximately) constant over the band B, and $f_n(\lambda,\omega)$ is (approximately) constant over $\omega \in B$ and all λ. The latter is equivalent to assuming that the noise process is uncorrelated in space, but stationary and not necessarily uncorrelated in time. The more general case where $N(x,t)$ is stationary and correlated in space and time is briefly discussed at the end of this section.

The space-time Fourier transform of the observations $Y(x,t)$ ($x = 1,\ldots,S$; $t = 1,\ldots,T$) yields the statistics

$$W(\lambda,\omega) = \frac{1}{2\pi\sqrt{ST}} \sum_{x=1}^{S} \sum_{t=1}^{T} Y(x,t) e^{i(x\lambda + t\omega)}$$

evaluated over frequencies λ, ω of the form $2\pi j/S$, $2\pi k/T$ respectively. These are key quantities in any space-time Fourier analysis. Their importance in the current context arises because, for large S and T, they are approximately independent, complex Gaussian random variables with zero means and variances $f(\lambda,\omega)$ ($(\lambda,\omega) \neq (0,0)$) (see Brillinger (1974) for example). Moreover, as is well known, the $W(\lambda,\omega)$ can be computed very efficiently.

Since the exact likelihood will be difficult to obtain in general, an obvious technique is to use maximum likelihood based on the asymptotic likelihood of the $W(\lambda,\omega)$, $\omega \in B$. After some manipulation, the log likelihood that results is, apart from a constant, proportional to

$$-\sum_{\lambda} \sum_{\omega \in B} \{\log \tilde{f}(\lambda,\omega_0) + |W(\lambda - \omega/V(\omega_0),\omega)|^2/\tilde{f}(\lambda,\omega_0)\} \tag{3.2}$$

where

$$\tilde{f}(\lambda,\omega_0) = \frac{f(\omega_0)}{2\pi} \cdot \frac{1-\rho(\omega_0)^2}{|1-\rho(\omega_0)e^{i\lambda}|^2} + f_n(\omega_0). \tag{3.3}$$

Equivalently, estimates of $f(\omega_0)$, $\rho(\omega_0)$, $f_n(\omega_0)$ and $V(\omega_0)$ are obtained by minimising

$$\sum_{\lambda} \{\log \tilde{f}(\lambda,\omega_0) + Q(\lambda, V(\omega_0))/\tilde{f}(\lambda,\omega_0)\} \tag{3.4}$$

where

$$Q(\lambda, V(\omega_0)) = \frac{1}{m} \sum_{\omega \in B} |W(\lambda - \omega/V(\omega_0),\omega)|^2$$

and m is the number of frequencies ω of the form $2\pi k/T$ in B. Given $V(\omega_0)$, minimising (3.4) is equivalent to fitting an AR(1) plus noise model in space to periodogram ordinates given by $Q(\lambda, V(\omega_0))$.

Let $\hat{V}(\omega_0)$, $\hat{\boldsymbol{\theta}} = (\hat{f}(\omega_0), \hat{\rho}(\omega_0), \hat{f}_n(\omega_0))^T$ be the estimates of $V(\omega_0)$, $\boldsymbol{\theta} = (f(\omega_0), \rho(\omega_0), f_n(\omega_0))^T$ obtained by minimising (3.4) and allow m, S, T to increase in such a way that $\|B\| = 2\pi m/T \downarrow 0$. Then, subject to certain regularity conditions, it can be shown that $\hat{V}(\omega_0)$, $\hat{\boldsymbol{\theta}}$ are strongly consistent estimators of $V(\omega_0)$, $\boldsymbol{\theta}$ and $\sqrt{mS}(\hat{V}(\omega_0) - V(\omega_0))$, $\sqrt{mS}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ are asymptotically independent, zero-mean, normal variates with

$$\begin{aligned}
&\lim_{m,S,T\to\infty} \{\mathrm{Var}\{\sqrt{mS}(\hat{V}(\omega_0) - V(\omega_0))\}\} \\
&\qquad = \frac{V(\omega_0)^4}{\omega_0^2} \cdot \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \frac{\partial \log \tilde{f}(\lambda, \omega_0)}{\partial \lambda} \right\}^2 d\lambda \right\}^{-1}, \\
&\lim_{m,S,T\to\infty} \{\mathrm{Var}\{\sqrt{mS}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\}\} \\
&\qquad = \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial \log \tilde{f}(\lambda, \omega_0)}{\partial \boldsymbol{\theta}} \cdot \frac{\partial \log \tilde{f}(\lambda, \omega_0)}{\partial \boldsymbol{\theta}^T} d\lambda \right\}^{-1}.
\end{aligned} \tag{3.5}$$

The method of proof follows conventional lines and is an amalgam of techniques given for stationary time series in standard works such as Hannan (1970) and Brillinger (1975). In addition to the assumptions already made, $Y(x,t)$ is typically required to satisfy a mixing condition in addition to smoothness conditions on the various spectral components. Appropriate conditions are given in Brillinger (1974).

As $\rho(\omega_0)$ approaches 1, minimising (3.4) becomes equivalent to maximising

$$\begin{aligned}
Q(0, V(\omega_0)) &= \frac{1}{m} \sum_{\omega \in B} |W(-\omega/V(\omega_0), \omega)|^2 \\
&= \frac{1}{2\pi m S} \sum_{\omega \in B} \left| \sum_{x=1}^{S} W_x(\omega) e^{-ix\omega/V(\omega_0)} \right|^2
\end{aligned}$$

where the $W_x(\omega)$ are the time Fourier transforms of the data with

$$W_x(\omega) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} Y(x,t) e^{it\omega}. \tag{3.6}$$

Thus, as $\rho(\omega_0)$ approaches 1, the method reduces to delay and sum beamforming.

Since $\hat{V}(\omega_0)$ and $\hat{\boldsymbol{\theta}}$ are asymptotically independent, a computationally efficient algorithm to obtain these estimates is given in Table 1. Note that the delay and sum estimate $\hat{V}^{(0)}$ is a consistent, but not efficient, estimator of $V(\omega_0)$ irrespective of the value of $\rho(\omega_0)$.

Table 1. Algorithm for fitting stochastic velocity model to equispaced arrays.

(1) Form an initial estimate $\hat{V}^{(0)}$ of $V(\omega_0)$ using the delay and sum method.
(2) Form an initial estimate $\hat{\boldsymbol{\theta}}^{(0)}$ of $\boldsymbol{\theta}$ by fitting an AR(1) plus noise model to the $Q(\lambda, \hat{V}^{(0)})$.
(3) Update $\hat{V}^{(0)}$ to $\hat{V}^{(1)}$ using one step of the Newton-Raphson algorithm that minimises $\sum_\lambda Q(\lambda, V)/\hat{f}(\lambda, \omega_0)$ where $\hat{f}(\lambda, \omega_0)$ is $\tilde{f}(\lambda, \omega_0)$ evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}^{(0)}$.
(4) Iterate from (2) until convergence.

To illustrate the practicality of the algorithm given in Table 1, we applied it to data from a seismic reflection-refraction survey undertaken by the Geophysics Division of the New Zealand Department of Scientific and Industrial Research (see Davey and Smith (1982)). The data subset considered comprised 512 time observations from 16 seismometers. These were spaced 1 km apart along a radial line from a shotpoint 6 km from the closest recorder where an explosive device was detonated to generate the data. The resulting traces were digitised at 200 samples per second and the data subset considered was selected from the stationary coda. Since the response functions of the recorders were band limited, a narrow band B of 11 time frequencies was considered centred about $\omega_0 = 2\pi f_0$ where $f_0 = 0.09$ cycles per unit time or 18 cycles per second. A contour plot of the space-time periodogram function $|W(\lambda, \omega)|^2$ over the region of interest is given in Fig. 1. Since the data contain a mix of refracted and reflected arrivals, the model introduced in Section 2 is not strictly applicable. However, for the data subset considered, the one reflected arrival is the dominant signal and so, for the purposes of illustration and to a first degree of approximation, the refracted arrivals are subsumed in the noise. (The extension of the model to handle a superposition of delayed signals is currently under study.) The conventional delay and sum estimate yielded $\hat{V}^{(0)} = 8.240$ km/sec and the algorithm described in Table 1 converged after two iterations to give a final value of $\hat{V} = 8.241$ km/sec. The line $\lambda = -\omega/\hat{V}$ is superimposed for reference on the contour plot given in Fig. 1. Although the velocity estimates are much the same, their standard errors are not. The standard error for $\hat{V}$ using the stochastic velocity model was 0.057 km/sec whereas that for $\hat{V}^{(0)}$ using the deterministic velocity model was 0.036 km/sec. Thus, as would be expected, the stochastic velocity model gives more conservative error bounds for the velocity estimates. It should be noted that the estimate of $\rho(\omega_0)$ was 0.91 with standard error 0.06 indicating that in this case it may be difficult to discriminate between the two competing models.

Finally we consider the situation where $f_n(\lambda, \omega)$ can only be regarded as approximately constant over $\omega \in B$ and $\lambda \in C$. Here C is a non-degenerate subset of $(-\pi, \pi]$ which should contain $-\omega_0/V(\omega_0)$ if sensible results are to be obtained. This allows $N(x, t)$ to be stationary and correlated in both time and space. Following the same arguments as before we find that estimates of $f(\omega_0)$, $\rho(\omega_0)$, $f_n(\omega_0)$ and $V(\omega_0)$ are obtained by minimising

$$\sum_{\lambda \in C} \sum_{\omega \in B} \{\log \tilde{f}(\lambda + \omega/V(\omega_0), \omega_0) + |W(\lambda, \omega)|^2 / \tilde{f}(\lambda + \omega/V(\omega_0), \omega_0)\}.$$

Moreover, under similar regularity conditions to those given before, the estimators are strongly consistent and $\sqrt{mS}(\hat{V}(\omega_0) - V(\omega_0))$, $\sqrt{mS}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ are asymptotically zero-mean, normal variates with variances and covariances given by (3.5), but with the range of integration now restricted to $C + \omega_0/V(\omega_0)$. The asymptotic independence of $\hat{V}(\omega_0)$ and $\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}$ is also preserved provided C is (asymptotically) symmetric about $-\omega_0/V(\omega_0)$.

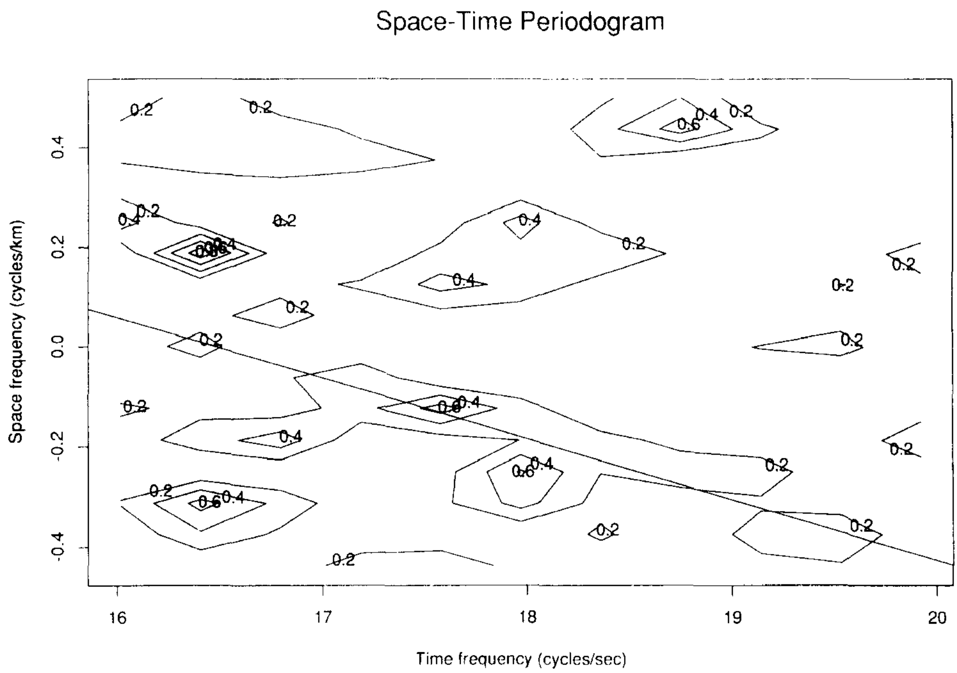

Fig. 1.

4. Deterministic irregular arrays

In this section we consider linear arrays of irregularly spaced recorders where the locations x_j $(x_1 < x_2 < \cdots < x_S)$ are given and regarded as deterministic rather than stochastic. Now consider the S-dimensional vector stationary process $\boldsymbol{Y}(t)$ with j-th component $Y(x_j, t)$ given by (2.1) where $a(x, \omega) = 1$ and $N(x, t)$ is stationary in time, uncorrelated in space and uncorrelated with $S(x, t)$. Denote the associated vector of time Fourier transforms by $\boldsymbol{W}(\omega)$ where $\boldsymbol{W}(\omega)$ has j-th component given by (3.6) with $x = x_j$. Assume as before that $L(x, \omega)$ is independent of both $S(t)$ and $N(t, x)$ with $L(x+1, \omega) - L(x, \omega)$ having a symmetric distribution, and that $Y(x, t)$ has been sampled at unit time intervals with the time interval chosen so that aliasing is not a problem. Then, for T large, the $\boldsymbol{W}(\omega)$ $(0 < \omega < \pi)$ may be regarded as approximately independent complex Gaussian random vectors each with mean $\mathbf{0}$ and covariance matrix

$$\boldsymbol{\Lambda}(\omega)\boldsymbol{f}(\omega)\boldsymbol{\Lambda}(\omega)^* \tag{4.1}$$

where

$$\boldsymbol{f}(\omega) = f(\omega)\boldsymbol{R}(\omega) + f_n(\omega)\boldsymbol{I}.$$

Here $\boldsymbol{I}$ denotes the identity matrix, $\boldsymbol{R}(\omega)$ the matrix with typical element $\rho(\omega)^{|x_j - x_k|}$, $\boldsymbol{\Lambda}(\omega)$ the diagonal matrix with j-th diagonal entry $\exp ix_j\omega/V(\omega)$ and the * denotes conjugate transpose. This result follows from the form of (2.2) and will be true in quite general circumstances (see Hannan (1970) and Brillinger (1975) for example).

Consider, in particular, the case of fitting the spectral model (4.1) over a narrow band B of m frequencies centred at ω_0 $(0 < \omega_0 < \pi)$ and assume that $f(\omega)$, $\rho(\omega)$, $f_n(\omega)$ and $V(\omega)$ are approximately constant over B. Proceeding as before we are led to estimate $f(\omega_0)$, $\rho(\omega_0)$, $f_n(\omega_0)$ and $V(\omega_0)$ by minimising

$$\log\det \boldsymbol{f}(\omega_0) + \frac{1}{m}\sum_{\omega\in B} \boldsymbol{U}(\omega)^* \boldsymbol{f}(\omega_0)^{-1}\boldsymbol{U}(\omega) \tag{4.2}$$

where $\boldsymbol{U}(\omega) = \boldsymbol{\Lambda}_0(\omega)^* \boldsymbol{W}(\omega)$ and $\boldsymbol{\Lambda}_0(\omega)$ is a diagonal matrix with typical element $\exp ix_j\omega/V(\omega_0)$. Note that the $\boldsymbol{U}(\omega)$ are vectors of re-phased Fourier transforms. The special case when $\rho(\omega_0) = 1$ has been considered by many authors and assumes that the coherence between any two recorders in the array is the same regardless of location. However the more general model given here assumes that the coherence between recorders is an exponentially decreasing function of the distance between them which would seem to be a reasonable assumption in practice.

Observe that $\boldsymbol{f}(\omega_0)$ is the covariance matrix of an irregulary sampled, continuous parameter AR(1) process plus noise. Such models can typically be fitted by maximising the likelihood directly or, by using the Kalman filter. Taking the latter approach, the re-phased Fourier transforms $\boldsymbol{U}(\omega)$ can be rearranged as m-dimensional vectors $\boldsymbol{U}_j$ $(j = 1, \ldots, S)$ where $\boldsymbol{U}_j$ has typical element $\boldsymbol{U}_j(\omega)$ with ω varying over the m frequencies in B. An appropriate state space model for the $\boldsymbol{U}_j$ is

$$\begin{aligned} \boldsymbol{U}_j &= \boldsymbol{S}_j + \boldsymbol{N}_j \\ \boldsymbol{S}_j &= \boldsymbol{R}_j \boldsymbol{S}_{j-1} + \boldsymbol{E}_j \end{aligned} \qquad (j = 1, \ldots, S) \tag{4.3}$$

where the (complex) vector processes $\{\boldsymbol{N}_j\}$, $\{\boldsymbol{E}_j\}$ are Gaussian and independent, $\boldsymbol{N}_j$ is white noise with covariance matrix $f_n(\omega_0)\boldsymbol{I}$, $\boldsymbol{R}_j$ is $\rho(\omega_0)^{x_j - x_{j-1}}\boldsymbol{I}$ with x_0 defined as $-\infty$, and the $\boldsymbol{E}_j$ are independent each with zero mean and $\operatorname{cov}(\boldsymbol{E}_j) = f(\omega_0)(1 - \rho(\omega_0)^{2(x_j - x_{j-1})}\boldsymbol{I}$. Note that the complex state space model (4.3) can be reformulated into one involving just the real and imaginary parts of $\boldsymbol{U}_j$, $\boldsymbol{S}_j$, $\boldsymbol{N}_j$, $\boldsymbol{E}_j$ and the model involves replicates over frequency. For a discussion on the fitting of such models to irregularly spaced data, see Robinson (1977) and Jones (1981).

As before, we now define $\hat{V}(\omega_0)$, $\hat{\boldsymbol{\theta}} = (\hat{f}(\omega_0), \hat{\rho}(\omega_0), \hat{f}_n(\omega_0))^T$ to be the estimates of $V(\omega_0)$, $\boldsymbol{\theta} = (f(\omega_0), \rho((\omega_0), f_n(\omega_0))^T$ that minimise (4.2), and let m and T increase in such a way that $\|B\| = 2\pi m/T \downarrow 0$. Given suitable regularity conditions (see Hannan (1975) for example), it can be shown that $\hat{V}(\omega_0)$, $\hat{\boldsymbol{\theta}}$ are strongly consistent estimators of $V(\omega_0)$, $\boldsymbol{\theta}$ and, moreover, $\sqrt{m}(\hat{V}(\omega_0) - V(\omega_0))$, $\sqrt{m}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ are asymptotically independent, zero-mean, normal variates with

$$\begin{aligned} &\lim_{m,T\to\infty}\{\operatorname{Var}\{\sqrt{m}(\hat{V}(\omega_0) - V(\omega_0))\}\} = \frac{V(\omega_0)^4}{\omega_0^2}\cdot\frac{1}{2}\{\boldsymbol{x}^T(\boldsymbol{\Phi} - \boldsymbol{I})\boldsymbol{x}\}^{-1}, \\ &\lim_{m,T\to\infty}\{\operatorname{Var}\{\sqrt{m}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\}\} = \boldsymbol{\Omega}^{-1}. \end{aligned} \tag{4.4}$$

Here the vector $\boldsymbol{x}$ has typical element x_j $(j = 1, \ldots, S)$, the matrix $\boldsymbol{\Phi}$ has typical element $\boldsymbol{f}(\omega_0)_{jk}\boldsymbol{f}(\omega_0)^{kj}$ ($\boldsymbol{f}(\omega_0)^{jk}$ denoting the typical element of $\boldsymbol{f}(\omega_0)^{-1}$) and $\boldsymbol{\Omega}$ has typical element

$$\boldsymbol{\Omega}_{jk} = \operatorname{tr}\left\{\boldsymbol{f}(\omega_0)^{-1}\frac{\partial \boldsymbol{f}(\omega_0)}{\partial\theta_j}\boldsymbol{f}(\omega_0)^{-1}\frac{\partial \boldsymbol{f}(\omega_0)}{\partial\theta_k}\right\}.$$

The asymptotic independence of $\hat{V}(\omega_0)$ and $\hat{\boldsymbol{\theta}}$, together with the observations made in the previous paragraph suggest the iterative fitting scheme given in Table 2. Again, $\hat{V}^{(0)}$ is a consistent, but not efficient, estimator of $V(\omega_0)$ irrespective of $\rho(\omega_0)$.

The model (4.1) can readily be generalised to include varying attenuations of the form indicated in model (2.1) and to include noise processes $N(x,t)$ which, although uncorrelated in time, have variances that depend on location.

Table 2. Algorithm for fitting stochastic velocity model to irregular arrays.

(1) Form an initial estimate $\hat{V}^{(0)}$ of $V(\omega_0)$ using the delay and sum method.
(2) Form an initial estimate $\hat{\boldsymbol{\theta}}^{(0)}$ of $\boldsymbol{\theta}$ by fitting a (replicated) continuous parameter AR(1) plus noise model to the $\boldsymbol{U}_j$.
(3) Update $\hat{V}^{(0)}$ to $\hat{V}^{(1)}$ using one step of the Newton-Raphson algorithm that minimises $(1/m)\sum_{\omega\in B}\boldsymbol{U}(\omega)^*\hat{\boldsymbol{f}}(\omega_0)^{-1}\boldsymbol{U}(\omega)$ where $\hat{\boldsymbol{f}}(\omega_0)$ is $\boldsymbol{f}(\omega_0)$ evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}^{(0)}$.
(4) Iterate from (2) until convergence.

5. Random irregular arrays

Consider the situation as described in Section 4, but where now the x_j may be regarded as a realisation of some stochastic process. In situations such as these the data is typically Fourier transformed as if the recorders were equispaced and the irregularities are taken account of by appropriate spatial modelling. However the resulting computational efficiencies, while clearly a virtue, will only apply if the number of recorders is large.

The case where the x_j can be thought of as independent observations from a common distribution with unknown mean and variance is given in Hinich (1982). Two other models of interest are

$$x_j = x_{j-1} + \varepsilon_j \tag{5.1}$$

and

$$x_j = j\Delta + \eta_j \tag{5.2}$$

where $\{\varepsilon_j\}$, $\{\eta_j\}$ are sequences of independent and identically distributed random variables with $E(\varepsilon_j) = \Delta$, $E(\eta_j) = 0$ and $\Delta > 0$ is an unknown space interval. To ensure that $x_j > x_{j-1}$ we shall require the ε_j to be distributed on $(0,\infty)$ and the η_j to be distributed on $(-\Delta/2, \Delta/2]$. Both situations relate to cases where the recorders are, on average, a distance Δ apart. However (5.1) allows measurement errors to accumulate as would happen if each recorder were located by being positioned relative to its neighbour. With (5.2) recorders are located, subject to error or jitter, at the points of a regular grid.

If model (5.1) is adopted then the process $S_j(t) = S(x_j, t)$ has mean μ and autocovariance function

$$(5.3) \qquad E\{S_j(s)S_{j+k}(s+t)\} = \int_{-\pi}^{\pi} e^{i(t\omega - k\theta_1(\omega))}\rho_1^{|k|}(\omega) f(\omega)d\omega$$

where $\rho_1(\omega)$, $\theta_1(\omega)$ are the modulus and argument respectively of $\phi(\omega) = E\{\psi(\omega,\omega)^{\varepsilon}\}$ with the random variable ε having the same distribution as ε_j. Thus $S_j(t)$ has space-time spectral density given by (3.1), but with $\theta(\omega)$, $\phi(\omega)$ replaced by $\rho_1(\omega)$, $\theta_1(\omega)$ respectively. This is formally equivalent to assuming that the unit increments of $L(x,\omega)$ have characteristic function $\phi(\omega)$ rather than $\psi(\omega,\omega)$. If the ε_j are identically equal to Δ so that the recorders are located at the points $j\Delta$, then (5.3) with $\Delta = 1$ reduces to (3.1) as expected. In the case of a deterministic velocity function, $\phi(\omega)$ is just the characteristic function of ε_j evaluated at $\omega/V(\omega)$. The important point here is that the spectral form (3.1) can arise purely through stochastic velocity, irregular spacing or a combination of both.

If model (5.2) is adopted then $S_j(t)$ has mean μ and space-time spectral density

$$(5.4) \quad \frac{f(\omega)}{2\pi} \cdot \frac{|1 + (\alpha(\omega) - 1)\rho(\omega)e^{i(\lambda+\theta(\omega))}|^2 - \rho(\omega)^2|\alpha(\omega)|^2}{|1 - \rho(\omega)e^{i(\lambda+\theta(\omega))}|^2} \qquad (-\pi < \lambda, \omega \le \pi)$$

where $\alpha(\omega) = E\{\psi(\omega,\omega)^{\eta}\}E\{\psi(\omega,\omega)^{-\eta}\}$ with the random variable η having the same distribution as η_j. Note that (5.4) reduces to (3.1) in the case where the η_j are zero and $\alpha(\omega) = 1$. However, in the case of a deterministic velocity function, (5.4) becomes

$$(5.5) \qquad f(\omega)\left\{|\beta(\omega/V(\omega))|^2\delta(\lambda + \omega/V(\omega)) + \frac{1}{2\pi}(1 - |\beta(\omega/V(\omega))|^2\right\}$$

where $\beta(\cdot)$ is the characteristic function of η_j. This is similar to (3.1) in the case $\rho(\omega) = 1$, but differs in that the delta function has been modified by the factor $|\beta(\omega/V(\omega))|^2$ and low order spectral mass added.

The spectral models arising from (5.1) and (5.2) can be fitted using essentially the same procedures as those described in Section 3. Note that the x_j will typically be known and estimates of $\phi(\omega)$, in the case of (5.1), and $\alpha(\omega)$, in the case of (5.2), can be obtained. Given the same assumptions as were made in Section 3, appropriate estimates are

$$(5.6) \quad \hat{\phi}(\omega) = \frac{1}{S-1}\sum_{j=2}^{S}\rho(\omega_0)^{\varepsilon_j}e^{i\varepsilon_j\omega/V(\omega_0)}$$

and

$$(5.7) \quad \hat{\alpha}(\omega) = \left\{\frac{1}{S}\sum_{j=1}^{S}\rho(\omega_0)^{\hat{\eta}_j}e^{i\hat{\eta}_j\omega/V(\omega_0)}\right\}\left\{\frac{1}{S}\sum_{j=1}^{S}\rho(\omega_0)^{-\hat{\eta}_j}e^{-i\hat{\eta}_j\omega/V(\omega_0)}\right\}$$

where $\hat{\eta}_j = x_j - j\Delta$ and $\hat{\Delta} = \sum jx_j / \sum j^2$ is the ordinary least squares estimator of Δ. However, as a result, the optimisation procedure is now more involved.

Acknowledgements

The author is greateful for the support provided by the Institute of Statistical Mathematics, Tokyo, Japan, where he completed the research for this paper, to Professor David Vere-Jones for his many helpful comments and constant encouragement, and to the Geophysics Division, New Zealand Department of Scientific and Industrial Research, for the use of their data.

References

Ash, R. B. and Gardner, M. F. (1975). *Topics in Stochastic Processes*, Academic Press, New York.

Brillinger, D. R. (1974). Fourier analysis of stationary processes, *Proceeding of the IEEE*, **62**, 1628–1643.

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

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

Cameron, M. A. and Hannan, E. J. (1978). Measuring the properties of plane waves, *Math. Geol.*, **10**, 1–22.

Cameron, M. A. and Thomson, P. J. (1985). Measuring attenuation, Time Series in the Time Domain, *Handbook of Statistics 5* (eds. E. J. Hannan, P. R. Krishnaiah and M. M. Rao), 363–367, North-Holland, Amsterdam.

Capon, J. (1969). High resolution frequency-wave number spectrum analysis, *Proceeding of the IEEE*, **57**, 1408–1418.

Davey, F. J. and Smith, E. G. C. (1982). The Wellington seismic reflection survey: Phase 1 1979, Report No. 187, Geophysics Division, Department of Scientific and Industrial Research, Wellington, New Zealand.

Hamon, B. V. and Hannan, E. J. (1974). Spectral estimation of time delay for dispersive and non-dispersive systems, *J. Roy. Statist. Soc. Ser. C*, **23**, 134–142.

Hannan, E. J. (1970). *Multiple Time Series*, Wiley, New York.

Hannan, E. J. (1975). Measuring the velocity of a signal, *Perspectives in Probability and Statistics* (ed. J. M. Gani), 227–237, Applied Probability Trust, Sheffield.

Hannan, E. J. and Thomson, P. J. (1988). Time delay estimation, *J. Time Ser. Anal.*, **9**, 21–33.

Hinich, M. J. (1981). Frequency-wave number array processing, *J. Acoust. Soc. Amer.*, **69**, 732–737.

Hinich, M. J. (1982). Estimating signal and noise using a random array, *J. Acoust. Soc. Amer.*, **71**, 97–99.

Hinich, M. J. and Shaman, P. (1972). Parameter estimation for an r-dimensional plane wave observed with additive independent Gaussian errors, *Ann. Math. Statist.*, **43**, 153–169.

Jones, R. H. (1981). Fitting a continuous time autoregression to discrete data, *Applied Time Series Analysis II* (ed. D. F. Findley), 651–682, Academic Press, New York.

Robinson, P. M. (1977). Estimation of a time series model from unequally spaced data, *Stochastic Process. Appl.*, **6**, 9–24.

Sato, H. (1984). Attenuation and envelope formation of three-component seismograms of small local earthquakes in randomly inhomogeneous lithosphere, *Journal of Geophysical Research*, **89**, 1221–1241.

Shumway, R. H. (1983). Replicated time-series regression: an approach to signal estimation and detection, Time Series in the Frequency Domain, *Handbook of Statistics 3* (eds. D. R. Brillinger and P. R. Krishnaiah), 383–408, North-Holland, Amsterdam.

Thomson, P. M. (1982). Signal estimation using an array of recorders, *Stochastic Process. Appl.*, **13**, 201–214.

Ziskind, I. and Wax, J. (1988). Maximum likelihood localisation of multiple sources by alternating projection, *IEEE Trans. Acoust. Speech Signal Process.*, **36**, 1553–1560.