Appl. Statist. (1974),
23, *No.* 2, p. 134

Spectral Estimation of Time Delay for Dispersive and Non-dispersive Systems

By B. V. HAMON and E. J. HANNAN

Commonwealth Scientific and Industrial Research Organization and the Australian National University

[Received February 1973. Revised June 1973]

SUMMARY

The problem considered is that of measuring the delay between the receipt of a signal at two sites where noise, incoherent between the two sites and with the signal, is also received at each site. In the case where the signal is dispersive the delay will be a function of frequency. The optimum solution involves a proper weighting of the phase shifts at each frequency, the weighting being dependent only on the coherence. The asymptotic distribution of the estimate is given. The method is worked out in detail for various special cases and is illustrated by an oceanographic example. The case where there are more than two sites is briefly discussed. A discussion is also given of the virtues of the method, of some possible variations to it and of some practical considerations.

Keywords: TIME SERIES ANALYSIS; SPECTRAL METHODS; TIME DELAYS; DISPERSION; WAVE VELOCITY

1. INTRODUCTION

WE shall discuss the problem of estimating the time delay between signals received at two recording sites, together with noise, incoherent with the signal. This has applications in many fields and we shall give an example using oceanographic data in Section 4. Other examples will be found in Fedor (1967) and Krause *et al.* (1972). The first of these deals with the determination of the velocity of drift of ionospheric clouds and the second with the speed of propagation of an impulse along a nerve fibre.

We use the term "non-dispersive" for any case where the time delay is expected to be independent of frequency. In such a case, only one parameter ("the" time delay) is to be estimated. More generally, time delay between two recording sites will be frequency dependent due, usually, to varying speed of propagation. This is the dispersive case and requires the estimation of more than one parameter. Interpretation of the parameters in the dispersive case will depend on the problem being studied and is outside the scope of this paper.

We consider the signals to be at least quasi-stationary. In particular, we do not consider the estimation of dispersion occurring in the path between a point source (in space and time) and a single recorder. In such cases, information on the dispersion is usually obtained by using the fact that different frequencies arrive at the recorder at different times (see, e.g., Dziewonski and Hales, 1972, for a review of this problem in the field of seismology).

It should be noted that we have in mind situations where the two records would be perfectly coherent were it not for uncorrelated noise at the two recorders. We discuss in Section 6 the complexities and possible misrepresentations that could arise with two or more sources whose signals might interfere at the two recorders.

We assume that the computations are digital so that the records are sampled at equal intervals of time. A standard method for estimating the lag, in the non-dispersive case, is to compute the cross-correlation (serial correlation) between the two records for a range of lags (necessarily integral multiples of the interval between sample points). This is used to locate the maximum of the serial correlation function which location validly estimates the delay, if the noise is incoherent between the two recorders. This method, which we further discuss below, is not usually efficient. We shall introduce an efficient procedure in Section 2 using the cross-spectrum between the two records and making optimum use of the signal-to-noise information, in each frequency band, as measured through the coherence function. The method extends to the dispersive case and has other advantages, which we shall discuss in Section 6. In Section 3 we discuss special cases of the method corresponding to special models for the nature of the dispersion. We shall also give simple procedures for determining the accuracy of the estimates and for testing hypotheses. These procedures are asymptotically valid under fairly general conditions where by this we mean that their validity increases as the length of record increases. We hope later to report simulations checking the rate of approach to the asymptotic situation. We shall discuss such questions further in Section 6. An asymptotic formula for the variance of the estimate of the velocity of propagation of a signal was given by Banerji (1960) and was used in Fedor (1967). The circumstances differ considerably from those of the present work and in particular commence from a very special model for the true serial correlation function.

Hoping that the techniques we introduce will be useful to research workers in many different fields, we have sought to make the present account widely accessible. We do not give proofs. A closely related but somewhat more special (and non-dispersive) situation is considered in Hannan and Robinson (1973), while in Hannan and Thomson (1973) proofs for the non-dispersive case are given, though there the concern is mainly with the estimation of group delay (see Section 6 of the present paper). The second author of the present paper will later publish a complete theoretical account including the general case of recorders at more than two sites. To indicate what is the nature of the situation with more than two sites we have, in Section 5 of this paper, included a brief account for the non-dispersive case and three sites. In the interests of wide accessibility, we do not assume any considerable (classical) statistical expertise but we do assume a knowledge of standard methods for the Fourier analysis of data.

2. Method

We assume that we have two records which have been sampled at equal time intervals. We take these intervals to be the time unit for convenience. Because of the nature of our estimation procedure we lose no generality in taking the series to have zero mean. We indicate the two series by $x(n)$, $y(n)$, $n = 1, 2, \ldots, N$. We assume the series to be stationary and to be composed of a signal and noise components, the latter being incoherent with each other and with the signals. Let $f_x(\omega), f_y(\omega), f_{yx}(\omega)$ be the true spectra and cross spectra, where ω is angular frequency so that $0 \leqslant \omega \leqslant \pi$. We begin by estimating these spectra at $M+1$ mid-band frequencies $\lambda_u = \pi u/M$, $u = 0, 1, \ldots, M$. We shall not use the estimates for $\lambda_u = 0, \pi$. In Section 6 we shall discuss other choices of these mid-band frequencies but continue to discuss these customary choices in this section, for simplicity. We call the estimates $\hat{f}_x(\lambda_u), \hat{f}_y(\lambda_u), \hat{f}_{yx}(\lambda_u)$. An advantageous method for the computation of these estimates is that based on the use of the discrete Fourier transform (DFT), namely

$$w_x(\omega_k) = \frac{1}{\sqrt{(2\pi N)}} \sum_1^N x(n) \exp(in\omega_k), \quad w_y(\omega_k) = \frac{1}{\sqrt{(2\pi N)}} \sum_1^N y(n) \exp(in\omega_k)$$

for $\omega_k = 2\pi k/N$, $k = 1, \ldots, [\frac{1}{2}N]$, where $[a]$ is the largest integer not greater than a. If $\sum_u$ is a sum over $m \simeq N/(2M)$ of the ω_k nearest to λ_u then we use

$$\hat{f}_x(\lambda_u) = m^{-1} \sum_u |w_x(\omega_k)|^2, \quad \hat{f}_{yx}(\lambda_u) = m^{-1} \sum_u w_y(\omega_k) \overline{w_x(\omega_k)}$$

with an analogous expression for $\hat{f}_y(\lambda_u)$. However, any of the standard methods for estimating spectra (or for that matter the use of a "fader" or "taper" in the definition of w_x, w_y) may be used. For a discussion of these methods we refer the reader to Hannan (1970, Chapter V). A value of M chosen, in the usual way, as a compromise between resolution and stability of the estimate will generally be satisfactory but the point is discussed in more detail later. We shall in fact be using only the estimates of phase, $\hat{\theta}(\lambda_u)$ and coherence, $\hat{\sigma}(\lambda_u)$, obtained from the estimates of spectra and cross-spectra. We define these by

$$\hat{f}_{yx}(\lambda_u) = |\hat{f}_{yx}(\lambda_u)| \exp\{i\hat{\theta}(\lambda_u)\}, \quad \hat{\sigma}(\lambda_u) = |\hat{f}_{yx}(\lambda_u)| / \{\hat{f}_x(\lambda_u) \hat{f}_y(\lambda_u)\}^{\frac{1}{2}}.$$

We use $\theta(\omega), \sigma(\omega)$ for the true quantities.

If the signal received at the y-recorder is obtained merely by delaying that received at the x-recorder by τ time units then $\theta(\omega) = \tau\omega$, i.e. the phase will be a linear function of frequency passing through the origin. However, at this stage, we assume a more general expression for $\theta(\omega)$ which is to be a known function of p parameters $\tau_1, \ldots, \tau_p$. We require $\theta(\omega)$ to be a reasonably regular function of these parameters, for example twice differentiable with continuous second derivative. We return to the simple time delay case later as a special case. We shall consider a band, B, of frequencies contained in $[0, \pi]$. This might be the whole interval but it might also be some smaller subinterval. (Indeed it need not be a continuous interval.) We return to this point later. It is essential, however, that there be no aliasing of frequencies, for the signal, with frequencies in the band B. In future we shall write $\sum_B$ for the sum over u such that $\lambda_u \in B$. We introduce a weighting function $W(\lambda_u)$, defined by $W(\lambda_u) = \hat{\sigma}^2(\lambda_u)/\{1-\hat{\sigma}^2(\lambda_u)\}$ and then the expression

$$M^{-1} \textstyle\sum_B [W(\lambda_u) \cos\{\hat{\theta}(\lambda_u) - \theta(\lambda_u)\}]. \tag{1}$$

As we have earlier indicated B will always exclude $\lambda_u = 0, \pi$, for which $\hat{\theta}(\lambda_u) \equiv \theta(\lambda_u) \equiv 0$. In (1) $\theta(\lambda_u)$ is regarded as a known function of $\tau_1, \ldots, \tau_p$. Our object is to give optimum estimates, $\hat{\tau}_1, \ldots, \hat{\tau}_p$, of these parameters, and to find their (asymptotic) distribution. The $\hat{\tau}_i$ are found as the values of τ_i that maximize (1).

Under rather general conditions on the nature of the random mechanisms generating $x(n), y(n)$ (and certainly without any assumption of a Gaussian distribution for these random variables) we may show that as $N \to \infty$, $M \to \infty$, $M/N \to 0$, the probability law of the $\hat{\tau}_i$ behaves in such a way that we may *treat the $\hat{\tau}_i$ as jointly normal with means $\hat{\tau}_i$ and with covariance matrix estimated as $N^{-1}V^{-1}$* where the $p \times p$ matrix V has v_{jk} in row j, column k and

$$v_{jk} = M^{-1} \textstyle\sum_B [\hat{\theta}'_j(\lambda_u)\, \hat{\theta}'_k(\lambda_u)\, W(\lambda_u)]. \tag{2}$$

Here $\hat{\theta}'_j(\lambda_u)$ is found by differentiating $\theta(\omega)$ with respect to τ_j and evaluating the result for $\omega = \lambda_u$ and $\tau_i = \hat{\tau}_i$, $i = 1, \ldots, p$.

The weight function $W(\lambda_u)$ is optimal in the sense that any essentially different choice would replace $\mathbf{V}^{-1}$ (or more properly the matrix $\tilde{\mathbf{V}}^{-1}$ that it estimates) by a larger matrix, A let us say, where by larger we mean that $\mathbf{x}'\mathbf{A}\mathbf{x} \geqslant \mathbf{x}'\tilde{\mathbf{V}}^{-1}\mathbf{x}$ for all vectors $\mathbf{x}$ and strict inequality holds for some. The choice of $W(\lambda_u)$ is a natural one since we are basically weighting inversely with the variances of the $\hat{\theta}(\lambda_u)$ (see Hannan, 1970, p. 256).

3. Special Cases

We shall illustrate the formulae of Section 2 in some special cases.

(i) *Simple time delay*: $\theta(\omega) = \tau\omega$

This is the most important case of a non-dispersive wave form. The maximization of (1) is now with respect to the single parameter τ. The variance of $\hat{\tau}$ is

$$N^{-1}[M^{-1} \textstyle\sum_B \lambda_u^2 W(\lambda_u)]^{-1}. \tag{3}$$

(ii) $\theta(\omega) = \tau_1 + \tau_2 \omega$

This model is difficult to justify physically in any strict sense, but might be a useful approximation for A.C. coupled recorders with slightly different low frequency responses. It is fairly easily checked that the maximization of (1) with respect to τ_1 and τ_2 is equivalent to first choosing τ_2 to maximize

$$|M^{-1} \textstyle\sum_B \exp i\{\hat{\theta}(\lambda_u) - \hat{\tau}_2 \lambda_u\} W(\lambda_u)| \tag{4}$$

and then using that maximizing value, $\hat{\tau}_2$, to estimate τ_1 as $\hat{\tau}_1 = \arctan(\hat{q}/\hat{p})$ with

$$M^{-1} \textstyle\sum_B [\exp i\{\hat{\theta}(\lambda_u) - \hat{\tau}_2 \lambda_u\} W(\lambda_u)] = \hat{p} + i\hat{q}.$$

The variance of $\hat{\tau}_2$ is increased compared to the case $\tau_1 = 0$ and is estimated as

$$N^{-1}[M^{-1} \textstyle\sum_B (\lambda_u - \bar{\gamma})^2 W(\lambda_u)]^{-1}$$

where $\bar{\lambda} = [\sum_B \lambda_u W(\lambda_u)]/[\sum_B W(\lambda_u)]$. To test $\tau_1 = 0$ we form

$$(\sqrt{N/M}) \textstyle\sum_B [\sin\{\hat{\theta}(\lambda_u) - \hat{\tau}_2 \lambda_u\} W(\lambda_u)]$$

and treat it as normal with zero mean and variance

$$M^{-1} \textstyle\sum_B W(\lambda_u).$$

(iii) $\theta(\omega) = \tau_1 + \tau_2 \omega + \tau_3 \omega^2$

This form might be expected to provide a useful fit when $\theta(\omega)$ is not quite linear in ω. Again it is not easy to justify $\tau_1 \neq 0$ on purely physical grounds. Again also we need to maximize an expression in τ_2, τ_3 only, even when $\tau_1 \neq 0$, the relevant expression being

$$|M^{-1} \textstyle\sum_B \exp i\{\hat{\theta}(\lambda_u) - \tau_2 \lambda_u - \tau_3 \lambda_u^2\} W(\lambda_u)|. \tag{5}$$

Then again

$$\hat{\tau}_1 = \arctan \hat{q}/\hat{p}$$

where $\hat{p}$ and $\hat{q}$ are the real and imaginary parts of the expression under the modulus sign in (5), evaluated at the maximizing values $\hat{\tau}_2, \hat{\tau}_3$. The matrix $\mathbf{V}$ now has entries

$$v_{jk} = M^{-1} \textstyle\sum_B \lambda_u^{j+k-2} W(\lambda_u), \quad j,k = 1,2,3.$$

The diagonal entries of $\mathbf{N}^{-1}\mathbf{V}^{-1}$ give the variances of the τ_j so that confidence intervals may be individually allocated to them in the standard fashion using these variances. If a joint (two-dimensional) confidence region for τ_2, τ_3 (for example) is required this may be constructed using the bottom right hand, 2×2, submatrix of $\mathbf{N}^{-1}\mathbf{V}^{-1}$ (see Johnson and Leone, 1964, p. 422). Of course $\tau_1 = 0$ may be tested using the first element in the diagonal of $\mathbf{N}^{-1}\mathbf{V}^{-1}$ as the variance of $\hat{\tau}_1$. If the test shows that τ_1 is not significantly different from zero it may be better to set $\tau_1 = 0$ and find τ_2, τ_3 by maximizing

$$M^{-1}\sum_B[\cos\{\hat{\theta}(\lambda_u)-\tau_2\lambda_u-\tau_3\lambda_u^2\}\,W(\lambda_u)]$$

since this will reduce the variances of $\hat{\tau}_2, \hat{\tau}_3$.

4. An Example

The spectral method discussed above has been applied to data on mean sea level (adjusted to fixed atmospheric pressure) for Sydney and Coff's Harbour (east coast of Australia). These data had previously been examined by cross-correlation, which showed that the Coff's Harbour record lagged the Sydney record by 1 time unit (1 day). The phase and coherence are shown in Fig. 5 of Hamon (1962), which is redrawn here as part of Fig. 1.

The same phase and coherence data have been re-examined by the present methods. We used the model $\theta(\omega) = \tau_2\omega+\tau_3\omega^2$, since it seemed reasonable in this case to assume $\theta(\omega)\to 0$ and $\omega\to 0$. N and M were 548 and 30 respectively. B consisted of the 25 estimates centred at $j\pi/30, j = 1, 2, \ldots, 25$, so that we excluded the five highest-frequency estimates, because of possible aliasing effects and low coherence.

Fig. 1 shows the phase and coherence data used and the fitted phase curve. The estimated values (± 1 standard deviation) were found to be $\hat{\tau}_2 = 0{\cdot}03 \pm 0{\cdot}13$ day, $\hat{\tau}_3 = 0{\cdot}76 \pm 0{\cdot}08$ rad^{-1} day^2. (The estimates were found in this case by hand contouring a field of values of (1), evaluated at suitable intervals in τ_2 and τ_3, but in general a computer subroutine for finding the maximum of a function of several variables would be used.) A separate calculation with the model $\theta(\omega) = \tau_1+\tau_2\omega$ showed $\hat{\tau}_1$ significantly different from zero, confirming that in this case a simple time delay is not an adequate fit to the data.

The shape of the fitted phase curve is consistent with the dispersion expected theoretically in this case for free waves (Buchwald and Adams, 1968), but more detailed physical interpretation is complicated by the fact that the waves are almost certainly not free (i.e. they are receiving energy from the wind stress) between Sydney and Coff's Harbour. Further discussion would be outside the scope of this paper.

5. Measurements from Three Recorders

All of the techniques of the earlier sections of this paper extend to the case where the velocity (speed and direction) is to be measured from recorders at more than two sites, each recorder receiving a (delayed) signal and noise. For simplicity here we shall briefly consider only the case of three recorders and only the case corresponding to a simple time delay where the signal is not dispersive.

Let $\boldsymbol{\xi}(1), \boldsymbol{\xi}(2), \boldsymbol{\xi}(3)$ be vectors of two components giving the positions of the three recorders, not on the same line. If a wave is being propagated in the direction of the vector, of unit length, $\boldsymbol{\phi}$ and with speed c then the delay at the jth recorder relative to the origin of co-ordinates will be $\{\boldsymbol{\phi}'\boldsymbol{\xi}(j)/c\}$, where we use a prime to indicate transposition of the vector $\boldsymbol{\phi}$. Then the phase shift between the jth and kth recorders will be $\omega c^{-1}\phi'\{\xi(j)-\xi(k)\}$ which we call $\theta_{jk}(\omega)$. We shall put $\tau = c^{-1}\phi$ in which case $\theta_{jk}(\omega) = \tau'\{\xi(j)-\xi(k)\}\,\omega$. We need therefore to estimate from the three records the value of c and the direction of the motion. We do this using the coherences $\hat{\sigma}_{jk}(\lambda_u)$ and phases $\hat{\theta}_{jk}(\lambda_u)$, $j,k = 1,2,3$, between the three pairs of recorders. We introduce the matrix $\mathbf{S}(\lambda_u)$ having $\hat{\sigma}_{jk}(\lambda_u)$ in row j, column k. (Of course $\hat{\sigma}_{jj}(\lambda_u) \equiv 1$.)

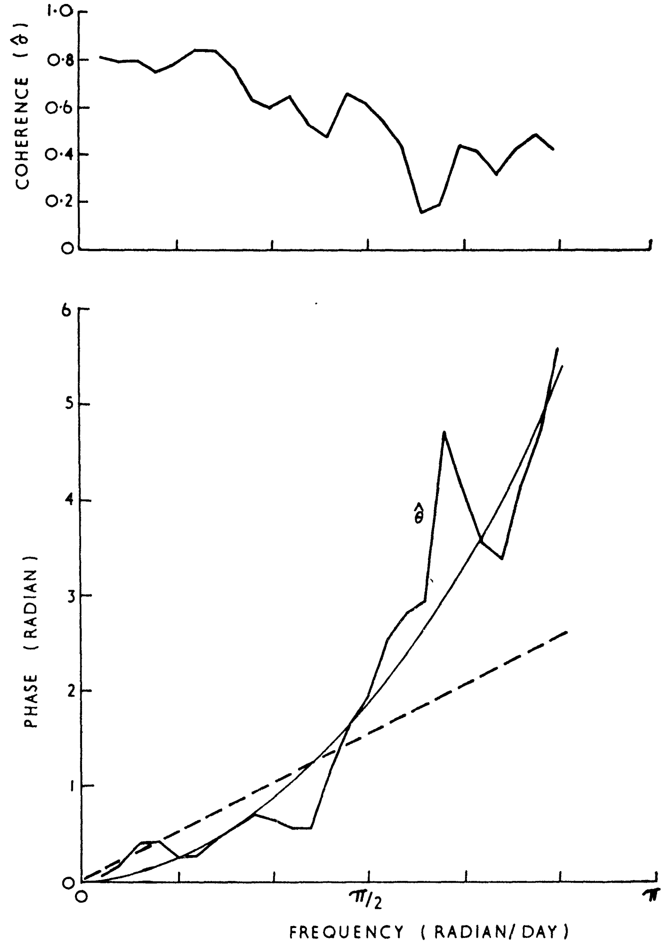

FIG. 1. Coherence (upper) and phase (lower) from adjusted mean sea level records at Sydney and Coff's Harbour. In the phase diagram, the smooth curve is the fitted phase curve (see text), and the dashed line the phase corresponding to a simple time delay of 1·0 day.

Let $\mathbf{S}(\lambda_u)^{-1}$ have $\hat{\sigma}^{jk}(\lambda_u)$ in row j, column k. We consider

$$-M^{-1}\sum_B\left[\sum_{j<k}\cos\{\theta_{jk}(\lambda_u)-\hat{\theta}_{jk}(\lambda_u)\}\,\hat{\sigma}^{jk}(\lambda_u)\,\hat{\sigma}_{jk}(\lambda_u)\right]. \tag{6}$$

It is easily seen that (6) reduces to (1) for two recorders. The expression (6) is a function of the two numbers required to specify the speed and the direction of the wave. We estimate these two numbers by choosing their values so as to maximize (6).

It will be simplest if we choose the parameters to be estimated as the two components τ_1, τ_2 of the vector $\boldsymbol{\tau} = c^{-1}\boldsymbol{\phi}$. Then as N increases, with M increasing also but so that $M/N \to 0$, the estimates $\hat{\tau}_1, \hat{\tau}_2$ may be treated as being normally distributed with mean τ_1, τ_2 and covariance matrix $\mathbf{N}^{-1}\mathbf{V}^{-1}$ where $\mathbf{V}$ has entries

$$v_{rs} = M^{-1}\Sigma_B \lambda_u^2 \left[\sum_{j<k} \{\xi_r(j) - \xi_r(k)\}\{\xi_s(j) - \xi_s(k)\} \hat{\sigma}^{jk}(\lambda_u)\hat{\sigma}_{jk}(\lambda_u) \right] \tag{7}$$

where $\xi_r(j)$ is the rth coordinate of $\boldsymbol{\xi}(j)$), $r = 1, 2$. Thus $\xi_1(j) - \xi_1(k)$ is the distance between the jth and kth sites in the direction of the first co-ordinate axis. The formula (7) is the direct analogue of (3).

6. Discussion

6.1. *Merits of the Method*

It is interesting to compare the variance of $\hat{\tau}$ for the simple time delay case discussed above with the variance of the corresponding quantity estimated by the more usual cross-correlation method. In the later case, the expression for the variance is more complicated than (3), being

$$N^{-1}[M^{-1}\Sigma_B\{\lambda_u^2(1-\hat{\sigma}^2(\lambda_u))\hat{f}_x(\lambda_u)\hat{f}_y(\lambda_u)\}] \Big/ \left[M^{-1}\sum_{\lambda_u \in B}\{\lambda_u^2 |\hat{f}_{xy}(\lambda_u)|\}\right]^2. \tag{8}$$

The value of (8) can never be less than that of (3). To illustrate this, consider the case where $f_x(\omega) = f_y(\omega) = \alpha\omega^{-2}$, $0 < a \leqslant \omega \leqslant b \leqslant \pi$ and $\sigma^2(\omega)$ is constant over the band $[a, b]$. If we define the efficiency, E, of the cross-correlation method as the ratio of (3) to (8) but taking the limit of that expression as $N \to \infty$, $M \to \infty$, $M/N \to 0$ we obtain

$$E = 3ab/(a^2 + b^2 + ab). \tag{9}$$

For example for $a = \pi/10$, $b = 9\pi/10$ this is 0·30. This, not unrealistic, example shows how considerable the loss in efficiency may be. The efficiency will be even less for these f_x, f_y if $\sigma^2(\omega)$ increases with ω. The cross correlation may be regarded, approximately, as being equivalent to the maximization of an expression of the form of (1) but with the (optimal) weight factor $\hat{W}(\lambda_u)$ replaced by

$$|f_{yx}(\lambda_u)| = \hat{\sigma}(\lambda_u)\{\hat{f}_x(\lambda_u)\hat{f}_y(\lambda_u)\}.$$

Of course if $(b-a)$ is small the expression (9) will be near to unity as it evidently must be since the band, B, is now very narrow and any fairly smooth weight function will give almost the same efficiency as the optimal one.

The variance formula (3) is much simpler than (8). If the accuracy of the cross-correlation method is needed, (8) must be evaluated and this requires the same computations as the method of Section 2. The computations are not very great, by modern standards, in any case. The method also has the virtue of flexibility since it is easy to omit frequency ranges which are, for example, contaminated by other sources of coherence. This is particularly easy if the DFT is used. It may also be observed that though the method uses the phase angles $\hat{\theta}(\lambda_u)$ it is independent of the arbitrariness of their determination up to a multiple of 2π.

Other advantages of the method are its generality (i.e. the possibility of extending it to the dispersive case) and the generality of the asymptotic results (these depend only on stationarity and properties of the noise and signal processes of a not very restrictive nature and not, for example, on their Gaussianity).

6.2. *Some Possible Variations*

It is not necessary to use $\lambda_u = \pi u/M$, $u = 1, \ldots, M-1$. In fact a better choice, if $B = (0, \pi)$, will be $\lambda_u = (2u-1)\pi/(2M)$, $u = 1, \ldots, M$ with $\omega_k = 0, \pi$ omitted from the calculations. In fact, it is not necessary to use bands of equal widths. If unequal bands are used and the band centred at λ_u is of "width" $2m_u \pi/N$, then in (1) and (3) we should eliminate the factor M^{-1} and insert an additional factor, $2m_u/N$, in each term. Though the evidence is not strong it seems possible that the weighting $\hat{\sigma}^2/(1-\hat{\sigma}^2)$ tends to overweight bands in which $\hat{\sigma}^2$ is low unless m_u is fairly large. Thus it may be wise to aggregate such bands into a smaller number of wider bands.

It is well known (Akaike and Yamanouchi, 1962) that if the group delay, $d\theta(\omega)/d\omega$, is large the estimate, $\hat{\sigma}$, of coherence will be biased downwards. To avoid this effect one might iterate the calculations described in Section 2 by using the estimate of $\theta(\omega)$ to rephase one series relative to the other before estimating coherence. Let us use $\tilde{\theta}(\omega)$ for the estimate of $\theta(\omega)$ obtained from the initial estimates of $\tau_1, \ldots, \tau_p$. One would now recompute the estimate of $f_{yx}(\lambda_u)$ as $\tilde{f}_{yx}(\lambda_u)$, let us say, where

$$\tilde{f}_{yx}(\lambda_u) = m^{-1} \textstyle\sum_u w_y(\omega_k)\overline{w_x(\omega_k)} \exp\{-i\tilde{\theta}(\omega_k)\}.$$

Then $\sigma(\lambda_u)$ is re-estimated using $\tilde{f}_{yx}(\lambda_u)$ in place of $\hat{f}_{yx}(\lambda_u)$. The resulting estimate is used to form the new weight function, $\tilde{W}(\lambda_u)$, and the estimation procedure for the τ_i is then repeated, the only change being the insertion of $\tilde{W}$ in place of W. In the simple delay case an alternative would be to delay the $y(n)$ series by $[\hat{\tau}]$ and then repeat all calculations. Since a large variation in $\theta(\omega)$ is required to bias badly the estimate of coherence, the replacement of $\hat{\tau}$ by $[\hat{\tau}]$ will not matter.

6.3. *Some Practical Considerations*

The expression (1) estimates

$$\frac{1}{\pi}\int_B \cos\{\theta_0(\omega) - \theta(\omega)\}\, \sigma^2(\omega)/\{1-\sigma^2(\omega)\}\, d\omega, \tag{10}$$

where now we find it convenient to indicate the true phase shift by a zero subscript in order to distinguish it from a trial function, $\theta(\omega)$. The expression (10) may have multiple maxima though the greatest of these will be at $\theta(\omega) \equiv \theta_0(\omega)$ (see Krause *et al.*, 1972, p. 261). There is therefore the possibility that (1) will have its greatest maximum not near to τ_0 (in the simple delay case) but rather, due to chance fluctuations, near to some other, smaller, maximum of (10). The formula for the variance of $\hat{\tau}$ takes no account of this possibility. Of course, as N increases the probability that such a gross error will occur becomes negligible, so that the accuracy of the determination of τ_0 is correctly described by our formulae. It is also true that in many cases prior information will be available that will severely restrict the possible range of values of τ_0. In this case, one will commence to search for the maximum of (1) in such a range and this will greatly reduce the chance of the kind of gross error we have here discussed.

A second practical problem closely related to the last one is that referred to in the introduction, namely the possibility that the signal may come from many sources (or even a diffuse source) so that the phase shift, $\theta(\omega)$, is a function not only of travel times between recorders but also of other spectral parameters. Thus, even if there were only two non-dispersive sources with spectra $f_1(\omega), f_2(\omega)$ and there was no noise, the cross-spectrum between the two records would be

$$f_{yx}(\omega) = f_1(\omega) \exp(i\tau_i \omega) + f_2(\omega) \exp(i\tau_2 \omega,$$

where τ_1 and τ_2 are the delays for the two signals, so that the coherence is not unity (unless $\tau_1 = \tau_2$) and $\theta(\omega)$ is a complicated function. It would be easy to construct cases in which such a situation could be mistaken for a dispersive signal from a single source. For a discussion of this problem see Gossard (1971). Munk *et al.* (1963, Section 7) discuss the estimation problem for the case of a signal with a discrete spectrum. The methods of the present work could be extended to deal with models such as we have discussed in this paragraph but the problem is more complicated than the one we have considered and in particular the computations would become more onerous.

Beyond a certain point elaboration of the form of $\theta(\omega)$ will probably cease to be profitable since increasing the value of p may merely result in an uninformative description of the variation in $\hat{\theta}(\omega)$. One might then prefer to examine the group delay, $d\theta(\omega)/d\omega$, estimating it over narrow bands, since this relates directly to the group velocity of the motion. We shall not discuss this further here but refer the reader to Hannan and Thomson (1973) where the estimation of group delay is considered.

References

Akaike, H. and Yamanouchi, Y. (1962). On the statistical estimation of frequency response function. *Ann. Inst. Statist. Math.*, **14**, 23–56.

Banerji, R. B. (1960). Radio-measurement of ionospheric drift as a problem in parameter estimation. In *Statistical Methods in Radio Wave Propagation* (Proc. Symp. UCLA) (W. C. Hoffman, ed.), pp. 40–48. Oxford: Pergamon.

Buchwald, V. T. and Adams, J. K. (1968). The propagation of continental shelf waves. *Proc. Roy. Soc.* A, **305**, 235–250.

Dziewonski, A. M. and Hales, M. (1972). Numerical analysis of dispersed seismic waves. In *Methods in Computational Physics*, Vol. 11: *Seismology. Surface Waves and Earth Oscillations* (Bruce A. Bolt, ed.), pp. 39–86. New York: Academic Press.

Fedor, L. S. (1967). A statistical approach to the determination of three dimensional ionospheric drifts. *J. Geophys. Res.*, **72**, 5401–5415.

Gossard, E. E. (1971). Use of coherence of time series in geo- and radiophysics. In *Statistical Methods and Instrumentation in Geophysics* (A. G. Kjelaas, ed.), pp. 41–66. Oslo: Tekhnologisk-Forlag.

Hamon, B. V. (1962). The spectrums of mean sea level at Sydney, Coff's Harbour, and Lord Howe Island. *J. Geophys. Res.*, **67**, 5147–5155. (Correction: *ibid.* **68**, 4635.)

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

Hannan, E. J. and Robinson, P. M. (1973). Lagged regression with unknown lags. *J. R. Statist. Soc.* B, **35**, 252–267.

Hannan, E. J. and Thomson, P. (1973). Estimating group delay. *Biometrika*, **60**, 241–253.

Johnson, N. L. and Leone, F. C. (1964). *Statistics and Experimental Design.* New York: Wiley.

Krause, D. J., Steadman, J. W. and Williams, T. W. (1972). Effect of record length on noise-induced error in the cross correlation estimate. *IEEE Trans. on Systems, Man and Cybernetics*, **SMC-2**, 255–261.

Munk, W. H., Miller, G. R. Snodgrass, F. E. and Barber, N. F. (1963). Directional recording of swell from distant storms. *Phil. Trans. Roy. Soc.* A, **255**, 505–584.