Radio Science, Volume 20, Number 6, Pages 1185–1195, November–December 1985

Spectral moment estimation in MST radars

Ronald F. Woodman

Instituto Geofísico del Perú, Lima, Peru

(Received February 20, 1985; revised June 25, 1985; accepted June 27, 1985.)

Processing techniques for the estimation of the power, the frequency shift, and the spectral width of mesosphere-stratosphere-troposphere (MST) radar returns are presented. The different techniques are compared in terms of statistical goodness of the estimators and in terms of computational convenience and simplicity.

1. INTRODUCTION

We are concerned in this paper with statistics of the returned signals in a mesosphere-stratosphere-troposphere (MST) radar. MST radar echoes are produced by fluctuations in the index of refraction of the atmosphere. In most cases, these are turbulence-induced fluctuations.

Because of the random nature of turbulence, radar returns from turbulence-induced fluctuations are stochastic processes and have to be characterized statistically. The returns from any one height form a random time series which, for the purpose of this work, will be considered quasi-stationary (stationary within an integration time) and Gaussian. Both assumptions are fair and very close to reality; one can always adjust the integration time so that the first assumption is true; the second is a consequence of the multiscattering nature of the radar return.

A Gaussian and stationary process is fully characterized by its autocorrelation function $\rho(\tau)$, or equivalently by its Fourier transform, the frequency power spectrum $S(\omega)$. In addition, in the case of scattering from turbulence induced fluctuations, the distribution of velocities in the turbulent scatter volume is Gaussian, and consequently the shape of $S(\omega)$ is also Gaussian. Thus the processes we will be discussing are Gaussian stationary processes and, in most cases, they have a Gaussian-shaped power spectrum. The first qualifier refers to the multivariant amplitude distribution of the signal proper, and the second to the distribution of the power at different frequencies, i.e., its spectral shape. They should not be confused. The autocorrelation function has also a (complex) Gaussian shape, since the Fourier transform of a Gaussian is also Gaussian.

We are aware that the assumption of homogeneity and Gaussian distribution of velocities may be violated for certain MST radar returns; for instance, multiple layers of turbulence are sometimes not resolved, and echoes are obtained at times from partial reflection, from quasi-systematic horizontal ledges in the index of refraction. Still the deviations from a Gaussian, or a Gaussian-looking bell-shaped spectrum, are seldom. Therefore, whenever we need to define the spectral shape in our discussion, we will assume it to be Gaussian. In most cases the conclusions we arrive at are valid, at least qualitatively, even when these are deviations from this assumption.

A Gaussian power spectrum has the form

$$S(\omega) = \frac{P}{(2\pi W^2)^{1/2}} \exp\left[-(\omega - \Omega)^2/2W^2\right] \tag{1}$$

where ω is the radian frequency. It is fully defined by the value of three parameters: P, Ω, and W. They correspond to the total power, the frequency shift, and the spectral width, respectively. Therefore, if the spectrum is Gaussian, these three parameters contain all the information we can obtain from the radar echoes, and they are all we need to know to characterize the process. They are a measure of three important physical properties of the medium: turbulence intensity, mean radial velocity, and velocity dispersion (turbulent velocity variance, $\langle\mu^2\rangle^{1/2}$, under certain conditions).

These three parameters correspond, also, to the three first moments of $S(\omega)$, defined as

$$P = \int S(\omega)\, d\omega \tag{2}$$

Paper number 5S0531.
0048-6604/85/005S-0531$08.00

$$\Omega = \frac{1}{P} \int \omega S(\omega)\, d\omega \tag{3}$$

$$W^2 = \frac{1}{P} \int (\omega - \Omega)^2 S(\omega)\, d\omega \tag{4}$$

It is preferable to take (2), (3), and (4) as the definition of the three parameters of interest, P, Ω, and W, since they are always well defined, even in the case when there are deviations from our assumptions and expectations about the nature of the process.

The scope of the present paper is to review signal processing techniques which have been used, or should be used, in MST radars, i.e., techniques which lead to a good estimation of the three first moments of the spectrum.

The number of possible estimators for P, Ω, and W^2 is unlimited. Therefore we cannot be exhaustive. In order to reduce the scope of the paper to bounds, we shall limit ourself to "good" estimators and to estimators that are presently in use.

We would like to talk about "best" estimators, but "best" is not easy to define, since there are two criteria for goodness one would like to satisfy: the estimator should be good from a statistical point of view; i.e., the variances of the estimated values should be as close to minimum as possible; but at the same time they should be practical. These two criteria are usually not compatible. As one improves the goodness of an estimator, one increases the complexity of the procedure. It is possible to talk about best estimators from a statistical view point, as we will see when we talk about the maximum likelihood (ML) estimators, but they are very difficult if not impossible to implement. In general, one would like a procedure which one can use in real time. This requirement can be very limitting, not so much because of the time scales of the processes, which are relatively slow, but because of the need to process a large number of parallel channels, especially if one is after the whole MST region with high-altitude resolution.

We can limit the scope of our paper if we limit ourselves to representative techniques which have been actually implemented in MST radars. We shall do this but include also some discussion about ML estimators since they give us a limit in performance with which we can compare other techniques.

Recently, *Zrnic'* [1979] has reviewed the subject of spectral moment estimation. Although the paper was motivated by weather radar applications and needs, it is fully applicable for MST radars. We shall take advantage of this review, avoiding repetition, unless we want to stress important conclusions. This includes the bibliography. The reader will find a very extensive list of references in the review by Zrnic'.

In the next section we shall describe straightforward power spectrum approaches; we shall then describe and discuss a correlation or covariance approach and finally the ML estimator concept and discuss the limits of performance they define.

2. MOMENT ESTIMATORS VIA POWER SPECTRUM

The most straightforward estimators of the three parameters of interest are suggested by their definition, through (2), (3), and (4). We should remember, though, that we cannot obtain in practice $S(\omega)$; we obtain instead statistically estimated values of it, $S'(\omega_i)$, at a finite discrete number N of frequencies.

The definitions suggest the following estimators, P', Ω', and W', for P, Ω, and W:

$$P' = \sum_{i=1}^{N} S'(\omega_i) \tag{5}$$

$$\Omega' = \frac{1}{P'} \sum_{i=1}^{N} \omega_i S'(\omega_i) \tag{6}$$

$$W'^2 = \frac{1}{P'} \sum_{i=1}^{N} (\omega_i - \Omega')^2 S'(\omega_i) \tag{7}$$

We need then, procedures, hopefully optimum, to find good estimated values of the power spectra. This is a very old and general problem for which there is extensive literature. The reader is referred to the book by *Blackman and Tukey* [1958] for an introduction, and to the section on spectral estimation in the IEEE book on signal processing for more modern approaches [*Rabiner and Rader*, 1976]. We would like to point out that unless the Nyquist frequency is significantly larger than the mean frequency plus the spectral widths, Ω' and W'^2, as given by (6) and (7), would be biased because of aliasing. This bias can be reduced if we assume periodicity with respect to the index i and period N and calculate the moments centered around a good guess of Ω. Let ω_j be a good guess of the actual value of Ω; then we evaluate a correction ω_ε such that $\Omega' = \omega_j + \omega_\varepsilon$, where ω_ε is evaluated from

$$\omega_\varepsilon = \frac{1}{P} \sum_{i=j-N/2}^{i=j+N/2} (\omega_i - \omega_j) S'(\omega_i) \tag{8}$$

This procedure can be iterated if desired. The spectral width is better estimated from

$$W' = \frac{1}{P} \sum_{i=j-N/2}^{i=j+N/2} (\omega_i - \omega_j + \omega_\varepsilon) S'(\omega_i) \tag{9}$$

In practice the problem is complicated by the fact that the signal is contaminated with noise and echoes from efficient targets on the ground (ground clutter). If we have an independent way of evaluating the noise power spectrum $N(\omega)$, the algorithms presented in (5)–(9) are still valid provided we replace $S'(\omega)$ by $S''(\omega) - N$, where $S''(\omega)$ is the power spectrum estimate including noise. Here the noise spectra have been taken as constant independent of frequency since usually the receiver bandwidth is much larger than the pulse repetition frequency (PRF), and there is no correlation between noise at two different sample pulses. The noise level can be estimated from an altitude where there is practically no signal, for instance from 45 km or from ionospheric altitudes, or from a few pulses with the transmitter off. The last approach requires a fraction, fortunately small, of the observing time, since the noise level is independent of altitude and one can use an average of the estimates from all the different altitudes.

The presence of ground clutter presents a source of bias and an additional problem. Different techniques have been used to cancel or minimize its effect. Ground clutter signals have a spectral signature which consists essentially of a single spectral line at the origin with a strength which depends on the ground shielding of the radar. At tropospheric and stratospheric heights it is at least comparable to the signal, and often many orders of magnitude larger. When the clutter is strong enough, it presents, in addition, a component of the spectrum with a spectral width comparable to the signal spectral width. This results from the slight propagation fading of the clutter and from echoes from vegetation moving with the wind. As in the case of noise, one should subtract the contribution of this interference before evaluating the moments. This contribution can be easily estimated in the case of nonfading clutter. The clutter adds a constant value to the signal, i.e., a spectral line, and can be estimated by integrating the returns for as long as the spectral estimation time (usually one or two minutes). One can then subtract the theoretical contribution of this constant component.

The fading component is difficult to estimate independently. One way to eliminate its biasing effect is to ignore the frequencies around zero (dc) frequency. This is only possible when the spectral offset Ω is larger than its width W. This occurs frequently, except when one is looking too close to vertical, or the medium velocity is too slow (horizontal velocities of the order of 1 m/s or less). Another technique takes advantage of the symmetry of the ground clutter component. The clutter contribution, being symmetric, disappears, and the signal contribution generates a negative image which is set to zero. One then evaluates the moments of what is obtained [*Sato and Woodman*, 1982]. This technique also has difficulties when the signal is too close to the center or to the Nyquist frequency.

In using the spectral moment technique the observer has some freedom in selecting the frequency spectrum estimating algorithm, the sampling frequency (or size of the spectral window), and the frequency resolution. This freedom has direct implications for the processing speed. Following are some criteria for its selection.

As far as the estimating algorithm, most modern procedures use a fast Fourier transform (FFT). This is an efficient way of doing it and should always be pursued, unless one has a hardwired autocorrelator. One should always use algorithms specially designed for 2^n samples and, if possible, specially designed for the particular exponent n selected; there can be considerable savings in time this way.

As far as the sampling frequency and maximum (Nyquist) frequency are concerned, the MST signals deserve some special considerations. The maximum duty cycle and maximum range of interest permit, in MST radars, pulse repetition frequencies which can be more than 2 orders of magnitude higher than the maximum frequency content of the signals. This produces high redundancy in the sampling and calls for some signal filtering: not so much to increase the system sensitivity, as one sometimes reads or hears, as to reduce the information rate input into the spectrum evaluating system and the amount of signal processing. As is well known, a FFT evaluation takes $N \ln N$ additions and computations. A reduction of, let us say, a factor of 256 in the number of sampled points speeds up the processing by a factor of 597, for a resulting spectrum of 64 points with the same frequency resolution.

The simplest and easiest filter to implement digitally is a boxcar integrator (coherent integration). It simply integrates N_f number of samples from a given altitude, takes the integrated value as a sample of the filtered output, and resets the integrator register to zero, ready for the next integration. The integration time should not be much larger nor shorter than half the period of the expected maximum Doppler frequency shift plus the expected spectral width. The integration time defines the sampling rate. Some un-

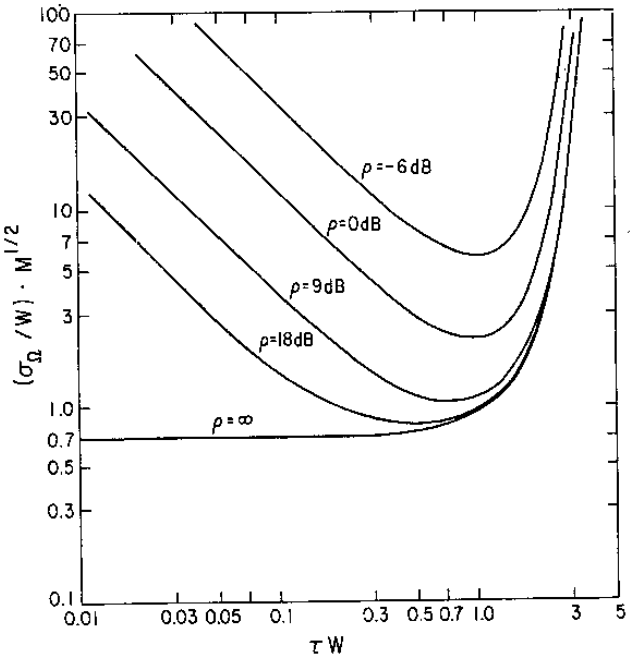

Fig. 1. Normalized standard deviation of mean-frequency estimator versus pulse pair spacing [*Miller and Rochwarger*, 1972].

dersampling and consequent aliasing can be allowed, if (8) and (9) are used for the evaluation of Ω' and W, but any oversampling is a waste of effort.

Another processing parameter that the observer has some freedom to choose is the frequency resolution. It is inversely proportional to the size of the time span taken in evaluating the discrete Fourier transform (DFT) or the time width of the weighting function (Hanning window, etc.). The latter should be longer but not much longer than the correlation width, say 2 or 4 times the half correlation time, since this will give us four or eight points to sample the spectral function shape, more than enough to determine the three parameters that define it. Higher resolution increases the processing effort without much gain in parameter accuracy.

In order to discuss the goodness of the spectral moment estimators we need to know the variances σ_P^2, σ_Ω^2, σ_W^2 of the estimated values with respect to their expectations. Here $\sigma_{P_i}^2 = \mathscr{E}(P_i - \langle P_i \rangle)^2$. This in general depends on the algorithm used for the evaluation of $S'(\omega)$. We will quote here the results obtained by *Denenberg* [1971].

He gives simple expressions for the variance σ_Ω^2 and σ_{W2}^2 for the case of a Gaussian spectrum with no additive noise. In terms of our notation and units (radians instead of hertz),

$$\sigma_\Omega^2 = \frac{\pi^{1/2} W^2}{2WT} \qquad (10)$$

and

$$\sigma_{W2}^2 = \frac{3\pi^{1/2} W^4}{4WT} \qquad (11)$$

Following his procedure, one can similarly find an expression for σ_W^2:

$$\sigma_W^2 = \frac{\sigma_{W2}^2}{4W^2} = \frac{3\pi^{1/2} W^2}{16WT} \qquad (12)$$

The expressions are valid for a single spectral estimate of a time series of length T. They can be generalized for the case where the spectral estimate $S'(\omega)$ is obtained from the average of M_s independent estimates $S_i'(\omega)$ of nonoverlapping segments each of length T. In this case, T should be replaced by $T_0 = M_s T$, the total length of the sequence.

In order to compare the goodness of different moment estimation techniques we have found it convenient to define the following figures of merit:

$$F_\Omega \equiv \frac{\sigma_\Omega (T_0 W)^{1/2}}{W} \qquad (13)$$

$$F_W \equiv \frac{\sigma_W (T_0 W)^{1/2}}{W} \qquad (14)$$

which normalize the variance of the estimates of Ω and W, σ_2 and σ_ω, with respect to the spectral width W, and remove the inverse dependence on the square root of the number of degrees of freedom (number of possible independent estimates), $(T_0 W)^{1/2}$, which should be common to all estimators. Here T_0 is the total observation time.

The corresponding figures of merit for (10) and (12) are

$$F_\Omega = (\pi/4)^{1/4} = 0.94 \qquad (15)$$

$$F_W = 3^{1/2}(\pi)^{1/4}/4 = 0.58 \qquad (16)$$

One can improve on the estimators (5), (8), and (9) with little additional effort. Following a rule that one should not use data that carry no information, one should use only those points in the spectrum, ω_i, for which there is a significant value for $S'(\omega_i)$, especially when the signal is contaminated with noise. This can be achieved with very little additional processing time once we have a reasonable estimate for the mean frequency and its width.

We can, in general, say that the spectral moment approach provides good estimators of the desired parameters. It involves the real time evaluation of DFT's for every altitude. This is a time-consuming operation, but fortunately MST echoes change slowly, especially at 50 MHz. With proper filtering (coherent integration) and the use of FFT processors, it should be possible to perform the necessary operations in real time, even in the case of high-resolution radars. The processing system at the Arecibo radar, for instance, is capable of processing in real time 32 point spectra, at 256 heights [*Woodman*, 1980]. It is actually capable of processing at least 4 times more information, being limited at present by the memory capacity of an array processor. It should be pointed out that the frequency of the Arecibo radar is 430 MHz, which produces time series (after proper filtering) close to 10 times faster than a 50-MHz radar and, therefore, 10 times more demanding. The coherent integration is performed by a special purpose preprocessor (a decoder). On the other hand, with the present state of the art, real-time full spectral processing of high-resolution radars is not possible with a simple minicomputer. One needs the help of a special purpose coherent integrator and FFT processors.

3. PARAMETER ESTIMATION BY NONLINEAR CURVE FITTING TECHNIQUES

The processing scheme described and discussed above implements the defining equations (2), (3), and (4) and does not take advantage of the knowledge about the spectral shape. There is a golden rule in detection theory that one should make use of as much a priori information as one has and ask only what one does not know. Equation (1) suggests another technique for evaluating the moments, or more properly, in this approach, the parameters P, Ω, and W^2. We can ask for a set of parameters such that $S(\omega) = S(\omega;\ P,\ \Omega,\ W)$ best approaches, in a least squares sense, the experimentally determined set $\{S'(\omega_i)\}$, for all i. This is a standard parameter estimation problem. This approach is more time demanding but should produce better estimates of P, Ω, and W. In fact we shall see later that with proper weighting, parameters obtained in this way are maximum likelihood estimates for a given set of experimental estimates, $\{S'(\omega_i)\}$.

The technique consists in minimizing an expression of the form

$$\varepsilon^2 = \sum_{i=1}^{N} A_i[S'(\omega_i) - S(\omega_i;\ P,\ \Omega,\ W)]^2 \tag{17}$$

The problem is nonlinear in the unknowns, P, Ω, and W, and involves special techniques. The reader is referred to the text by *Bard* [1974] for a comprehensive treatment.

This approach has been taken by *Sato and Woodman* [1982] to process ST spectra obtained with the 430-MHz radar at Arecibo. In fact, they used the technique to estimate up to eight additional parameters which define the noise, N, ground clutter interference, and if necessary, possible interference from strong turbulent layers from lower altitudes which leak to higher altitudes through code side lobes. The technique includes instrumental and signal processing sources of distortion and biases in the theoretical function. In this way the parameters of interest are evaluated free of all sources of biasing. Notice that an estimation of noise level and clutter characteristics are obtained simultaneously with the signal parameters.

This approach involves first the estimation of $S'(\omega)$, as in the previous case. The parameter information is obtained at the cost of additional processing.

Nonlinear automatic least squares parameter estimation involves nontrivial procedures. In the case of Arecibo, the additional processing is performed off line [*Sato and Woodman*, 1982]. This takes (making use of a floating-point array processor (AP-120)) a time equivalent to the time it took to obtain the data. Although it is feasible to perform this additional processing in real time by doubling the processing capacity, for many applications it is not necessary to perform the nonlinear estimation in real time.

4. THE AUTOCOVARIANCE OR AUTOCORRELATION APPROACH

One of the most efficient techniques, from the point of view of processing requirements, in obtaining P, Ω, and W is the single delay autocorrelation approach. In this approach the signal power and the autocovariance at a single delay are evaluated through the classical estimators

$$P' = \rho'(0) = \frac{1}{M} \sum_{i=1}^{M} x_i x_i^* \tag{18}$$

$$\rho'(\tau_j) = \frac{1}{M-j} \sum_{i=1}^{M-j} x_i x_{i+j}^* \tag{19}$$

where x_i is the ith complex sample corresponding to a given altitude and τ_j is the time displacement corresponding to j samples. The mean frequency shift and the velocity spread, Ω' and W', are then obtained from

$$\Omega' = \frac{\phi(\tau_1)}{\tau_1} \tag{20}$$

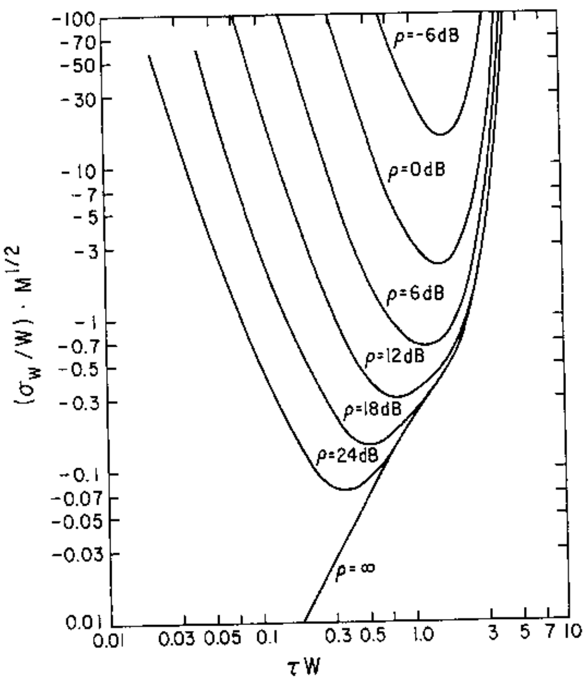

Fig. 2. Normalized standard deviation of the estimate of frequency spread versus pulse pair spacing [*Miller and Rochwarger*, 1972].

$$W^2 = 2\,\frac{1 - |\rho'(\tau_1)|/(\rho'(0) - \text{noise})}{\tau_1^2} \tag{21}$$

The technique takes advantage of the relationship that exists between the n'th derivative of the correlation function evaluated at the origin and the n'th moment of the frequency spectrum.

The technique was first used in 1968 by *Woodman and Hagfors* [1969] for estimating the electromagnetic drift of ionospheric plasmas at Jicamarca, and it was first used in 1972 by *Woodman and Guillén* [1974] for stratospheric and mesospheric applications. The technique is in much use today by the meteorological radar community, apparently as a consequence of some independent work by *Rummler* [1968] and by *Miller and Rochwarger* [1972], and has been subjected to much discussion and evaluation in the literature.

This technique involves only two complex multiplications and additions per altitude sample, as compared to ln N in the case of spectral moment estimation (where N is the number of spectral points). The variance of this approach is comparable to that obtained by integrating the moments of the frequency spectrum [*Rummler*, 1968; *Woodman and Hagfors*, 1969]. But this should not come as a surprise. After all, it is easily accepted that evaluating the power via the average of the square of the magnitudes (equation (18)) yields the same variance as that evaluated by integrating the area of the frequency spectrum (equation (5)). This is only a particular case, corresponding to the zeroth moment of a more general rule.

Woodman and Hagfors give us a simple expression for the variance of the mean angular frequency shift, valid for large values of M and small resultant values of $\tau^2\sigma_\Omega^2$ ($\ll 1$ rad):

$$\sigma_\Omega^2 = \frac{\rho^2(0) - \rho^2(\tau)}{2\tau^2 M \rho^2(\tau)} \tag{22}$$

Here M is the number of independent estimates. It is interesting to compare the figure of merit of this approach with that of (13). For large S/N ratios and Gaussian-shaped autocorrelation functions, (22) takes its best values at small τ. In this case,

$$\sigma_\Omega^2 = \frac{W^2}{2M} \tag{23}$$

For a given observation time T_0, the number of independent estimates is approximately $M \simeq T_0 W$. This is not quite equal, since contiguous sampled pairs, sampled at $1/W$ seconds apart, produce correlated estimates of the correlation function; therefore the number of independent estimates is somewhat less. Hence the figure of merit, F_Ω, would be somewhat larger than $1/2^{1/2}$, but in any case comparable to the spectral moment approach.

Later on, when we consider the case of using autocorrelation values at multiple delays, we shall see that the variances of the estimate using the single delay technique are close to optimum only when the signal to noise ratio is high. This relatively good performance deteriorates as the signal to noise ratio goes down. But it should be mentioned that the same happens with the spectrum moment approach represented by (5), (6), and (7), but not with the more sophisticated algorithm which includes weighting the spectral density by zero in the regions where there is no signal (match filter approach), or with the parameter estimation technique we have previously discussed.

Another limitation of this technique is the difficulty in discriminating against fading ground clutter or any other kind of interference. Fortunately, in many MST installations there are only nonfading clutter and white noise to worry about, and the biasing effect they produce can be eliminated by subtracting independent estimates of their contributions to $\rho(0)$ and $\rho(\tau)$. These estimates can be obtained by

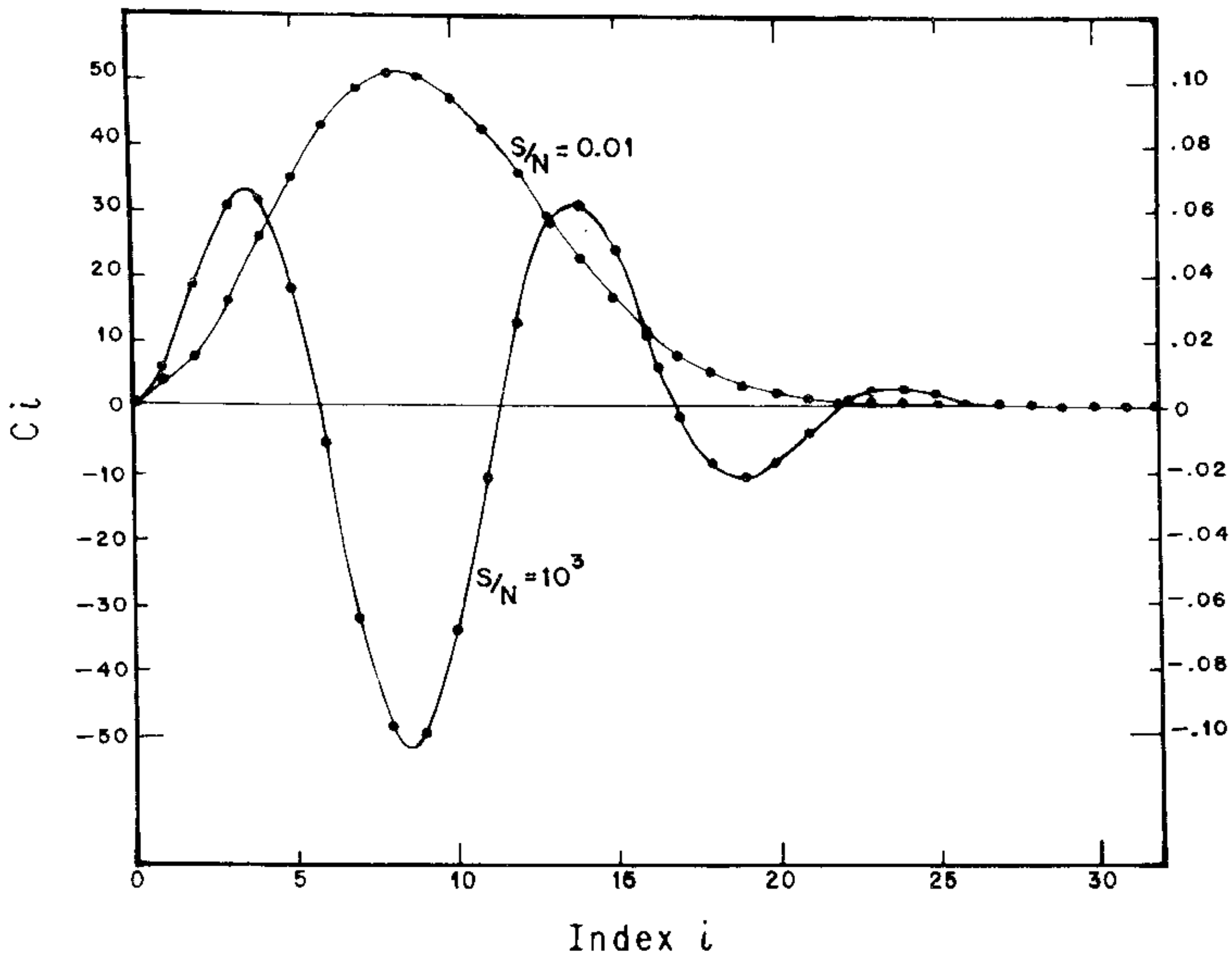

Fig. 3. Optimum weighting coefficients C_i to minimize $\sigma_\Omega = \langle(\Omega_a - \Omega)^2\rangle$, where $\Omega_a = \sum_{i-1}^{n} C_i \Omega_i$ and Ω_i is obtained from an estimate of the correlation function at a single delay τ_i. Two cases are considered, one for low (right ordinate) and the other for high (left ordinate) signal to noise ratio. The correlation function is assumed to have a Gaussian shape $\rho(\tau_i) = \exp(-0.693\tau_i^2/\tau_c^2)$. Here $n = 32$, and the sampling time $\tau_s = 0.1\tau_c$.

the same methods described before for the spectral moment approach.

Going from (3) and (4) to (20) and (21) involves approximating the derivative of $\rho(\tau)$ by finite differences between $\rho(0)$ and $\rho(\tau)$. This presents a bias which could become significant for relatively large values of τ [*Miller*, 1974]. Fortunately, in the case of a symmetric spectrum, equation (20) is an equality, and the bias disappears [*Woodman and Guillén*, 1974]. This is important since optimum values of τ, for noisy signals, are not close to the origin.

We are reproducing two graphs here (Figures 1 and 2) from *Miller and Rochwarger* [1972] which depict the performance of the single delay autocorrelation technique, by plotting the standard deviation of the estimates for Ω and W as a function of the sample separation τ ($\equiv h$, in their notation).

From Figures 1 and 2 we can see that the best separation for τ is that around a characteristic width of the correlation function $1/W$ and that for noisy signals the standard deviations of the estimates Ω' and W' are inversely proportional to the S/N ratio. Similar plots were produced by *Woodman and Hagfors* [1969] but for a typical incoherent scatter autocorrelation function shape.

It should be mentioned that the single delay autocorrelation approach, in contrast to the frequency spectrum approach, is very sensitive to the prefiltering of the time series. Filtering of the signal in this case does improve the signal to noise ratio and hence reduces the variance of the estimates. As is to be expected, optimum results are obtained using a matched filter, matched to the shape of the signal spectrum. But a boxcar integrator (coherent integration) produces similar results and is much easier to implement. It should be kept in mind, in any case,

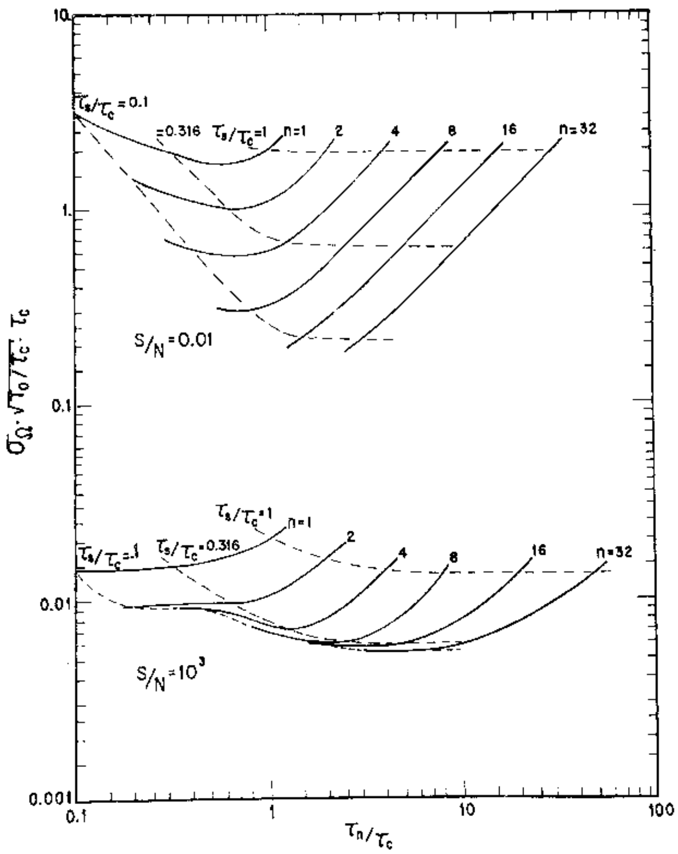

Fig. 4. Normalized standard deviation σ_Ω, defined as in Figure 3, as a function of the normalized delay τ_n/τ_c, the last lag corresponding to $i = n$. Here n, labeling the solid lines, is the number of correlation points taken in the average. Two signal to noise ratio are considered. The dashed lines join points of equal sampling time τ_s/τ_c.

that filtering could be a source of systematic biasing of Ω and W. This bias can be computed theoretically and should be correlated for.

5. COVARIANCE APPROACH AT MULTIPLE DELAYS

If the covariance approach was so efficient at a single delay, it is natural to ask how much improvement can be obtained by using more than one delay, τ. Let Ω_i and σ_i be estimates of Ω and σ_i obtained on the basis of equations (20) and (21) for different values τ_i of τ. We can always obtain a new estimate Ω_a and W_a through

$$\Omega_a = \sum_{i=1}^{n} C_i \Omega_i \quad (24)$$

$$W_a = \sum_{j=1}^{n} C_j W_j \quad (25)$$

where C_i and C_j are weights properly selected to minimize the variances of Ω_a and W_a and normalized such that $\sum C_i = \sum C_j = 1$. *Woodman* [1975] has treated the problem for the frequency shift Ω_a. He found an optimum set of values C_i, such that $\langle(\Omega_a - \Omega)^2\rangle$ is a minimum, and discussed numerically the effect of averaging for different signal to noise ratios, sampling spacing, number of averaged estimation, and correlation shapes. Figure 3 depicts the optimum set of weights C_i for two signal to noise ratios, for a Gaussian-shaped autocorrelation function, $\rho(\tau) = \exp(-0.693\tau^2/\tau_c^2)$. Normalization is such that $\rho(\tau_c) = \frac{1}{2}$. Lag spacing in this case is $0.1\tau_c$, and the number of lags averaged, n, is 32. The set of weights for low signal to noise is as expected; it corresponds to the normalized inverse of the variances of Ω_i, a well-known result for optimum averaging of independent samples. The resultant set for high S/N ratio is somewhat surprising; it has negative as well as positive signs, with absolute values which are larger than unity. This is a consequence of the fact that the difference estimates are not independent of one another.

Figure 4 shows the standard deviation Ωa, normalized with respect to a given observational time T_0, and characteristic width τ_c, as a function of the delay of the last lag, τ_n, used in the evaluation of Ω_n. The number of different Ω_i averaged, $i = 1, 2, 4, \cdots, 32$, are labeling the curves. Two cases are considered, one for high and the other for low signal to noise ratios, S/N. For every point in the graph there is a set of weighting coefficients such as the ones shown in Figure 3 for $n = 32$ and $0.1\tau_c$ spacing. The first conclusion we can draw from these results is that, indeed, for high signal to noise ratios there is not much difference between the standard deviation with 32 points at optimum delay and a single point close to the origin. There is a 60% difference in going from one to two points, and an additional 50% in going from two to 32. This last improvement is certainly not worth the effort. The increase from one to two could be justified, especially if the redundancy is used to check the existence of unexpected interference.

On the other hand, we see that for low S/N ratios and if optimum sampling time ($\tau_n/\tau_c \approx 1$) is used, the standard deviation improves faster than $1/n^{1/2}$, almost like $1/n$. In this case, every estimate Ω_i is almost completely independent of every other. The reason for an increase faster than $1/n^{1/2}$ comes from the fact that the number of independent estimates is increased like n^2, first because of the increase in n proper, and second as a consequence of an increase in M. This latter increase is due to the decrease in sampling time as n increases. Note that optimum spacing calls for a relatively constant τ_n, within a characteristic width, avoiding large uncorrelated lags which do not contribute information.

In practice, the sampling time is related to the coherent integration time. The sampling time cannot be less than a given coherent integration time. If the latter is reduced to decrease the sampling time, and increase M, the S/N increases, canceling any expected reduction in σ_Ω. There could still be an improvement due to a larger possible n, but only proportional to its square root. Figure 4 assumes constant S/N ratios and does not show changes due to different possible coherent integration times.

If for any reason the sampling and incoherent integration time is fixed at a small value, e.g., $\tau_s = 0.1\tau_c$, an increase in n improves σ_Ω significantly as shown by the corresponding dashed line in Figure 4. The improvement is fast, as $1/n$, until it levels off when τ_n becomes larger than the correlation width. The fast improvement is a consequence of an increase in n and the involvement of longer delays which give a better estimate of Ω_i. The leveling off comes about when the new delays involved are no longer correlated.

We can conclude then that Ω', as defined in (18), is a good estimator from a statistical as well as from a practical point of view when the S/N is better than 1, but it is not as good when the signal to noise ratio is low, and it is far from optimum if the coherent integration time is much smaller than a correlation characteristic time (τ_c or $1/W$).

It should be mentioned that the deterioration with noise of the single delay covariance approach with respect to optimum should not be held as an argument in favor of the simple spectral moment approach. Unless some more sophisticated processing is performed with the spectra, the single delay autocorrelation approach yields the same performance as the straight spectral approach, including the case of noise signals, as was quoted before [*Rummler*, 1968].

Similar computations have not been performed for the variance of the spectral width estimate, but we can expect that, qualitatively at least, the same conclusions will hold.

6. MAXIMUM LIKELIHOOD ESTIMATORS AND BOUNDS

Given a set or sequence (random process) of M observables x_i with a joint probability function $F(\mathbf{X}; \{A\})$ such that

$$F(\mathbf{X}; \{A\}) = \mathscr{P}(\mathbf{X} < \mathbf{x} < \mathbf{X} + d^N\mathbf{X}/\{A\})$$

where $\{A\}$ is a set of parameters. It is possible to find practically an innumerable number of estimators of $\{A\}$ and estimates $\{A_i\}$ as a function of the observables. For these estimates to be of practical use these must meet the condition that $\langle A_i' = A_i$ and that the variances $\sigma_a^2 = \langle (A_i' - A_i)^2 \rangle$ be small. For a given sample $\mathbf{x}$ of the process, we can form a function $F(\mathbf{x}; \{A\})$ and let $\{A\}$ vary. There will be a value of $\{A\}$ for which $F(\mathbf{x}; \{A\})$ is a maximum in $\{A\}$ space. This value is called the maximum likelihood (ML) estimate of $\{A\}$. It can be shown that such an estimate produces minimum variance among all possible estimators [*Cramer*, 1946].

Usually, it is not possible to find explicit solutions or practical algorithms for the ML estimators; on the other hand, the theory gives us formal expressions for the ML variances, which can be used to compare the "efficiency" of a given estimator. It is possible, in the case of large M processes with a Gaussian-shaped spectrum plus white noise, and using justifiable approximations, to obtain explicit expressions for these bounds. *Zrnic'* [1979], for instance, using a ML approach, finds the following lower bounds:

$$\sigma_\Omega^2 = \frac{12W^2(WT_s/2\pi)^2}{M[1 - 12(WT_s/2\pi)^2]} \tag{26}$$

when the noise level is zero, and

$$\sigma_\Omega^2 \geq \frac{W^2}{M}\, 4(\pi)^{1/2}(WT_s/2\pi)(N/s)^2 \tag{27}$$

when the S/N $\ll 1$ and $WT_s/2\pi \ll 1$. He assumes a continuous sequence of M complex samples spaced by T_s.

We should state, though, that we find (26) disturbing, since for a given observation time $T_0 = T_s M$ we can make the variance arbitrarily small by making T_s as small as possible. This is contrary to our expectations, since for a given W and no noise, sampling times smaller than W^{-1} give redundant information and should not improve the variance of any estimator. There is no explicit indication in the reference for the expression not to be valid for small WT_s.

If the sequence of observables $\{x_i\}$ is given by M pairs of independent complex values $\{x_{1i}, x_{2i}\}$ but correlated in between, the ML estimator can be found explicitly for large values of M [*Miller*, 1974]. It turns out to be the same as the covariance approach heuristically described by *Rummler* [1968] and *Woodman and Hagfors* [1969] and discussed in section 4.

It is also possible to use a ML approach starting with sample estimates of $\rho'(\tau_i)$ or $S'(\omega_i)$, of either the autocovariance function $\rho(\tau)$ or the spectrum $S(\omega)$, as the set of random variables to be used in a ML estimate of the parameters P, Ω, W, and N that we are interested in. The procedure, then, starts with the set of observables $\{x_i\}$, from which we obtained an estimate $\rho'(\tau_i)$ or $S'(\omega_i)$, of $\rho(\tau_i)$ or $S(\omega_i)$, using any of the available algorithms. These estimates, which hopefully contain all the desired information about the process, are then used in a ML approach to obtain the desired parameters. *Levine* [1965] has taken this approach starting with an estimate $S'(\omega)$ of the frequency spectrum. We refer the reader to the original reference, or to the review by *Zrnic'* [1979], for the solution algorithm. There is no explicit formula for the estimates. They involve the solution of some nonlinear simultaneous equations. The lower bounds for the variances are given by

$$\sigma_\rho^2 \geq \frac{P^2}{M}\left(2.25 - \frac{30}{(2\pi)^2}(WT_s)^2 + \frac{180}{(2\pi)^4}(WT_s)^4\right) \tag{28}$$

$$\sigma_\Omega^2 \geq \frac{3}{\pi^2}\frac{W^2}{M}(WT_s)^2 \tag{29}$$

$$\sigma_W^2 \geq \frac{45}{2\pi^4}\frac{W^2}{M}(WT_s)^4 \tag{30}$$

These bounds are valid for large signal to noise ratios.

We should notice that (29) gives about the same lower bounds as (26), which means that at least for the frequency shift variance this approach can be as good as the ML approach which starts with the observational time series $\mathbf{x}_i$. In terms of our figures of merit we can write

$$F_\Omega = \frac{3^{1/2}}{\pi}(WT_s)^{3/2} \tag{31}$$

$$F_W = 45^{1/2}\frac{1}{2\pi^2}(WT_s)^2 \tag{32}$$

We can see that for sampling times comparable to a correlation time, i.e., for $WT_s \simeq 1$, the performance of the spectrum and the single delay frequency shift estimators is comparable to both ML estimators. According to (29) and (30) both estimators improve as we reduce the sampling time spacings, eventually becoming much better than the simple estimators we have mentioned. Again, we find this behavior in the limit, as $T_s \to 0$, disturbing, since redundant high sampling rates should eventually produce redundant oversampling, which should not decrease the variance of our estimates. Figure 4, for instance, despite its sophistication, definitely does not show this improvement; it shows instead some leveling off, as we expect.

The corresponding variances for the case of small S/N ratios are

$$\sigma_\rho^2 = \frac{3}{2\pi^{1/2}}(WT_s)\frac{N^2}{M} \tag{33}$$

$$\sigma_\Omega^2 = \frac{2W^2}{\pi^{1/2}M}(WT_s)\left(\frac{N}{S}\right)^2 \tag{34}$$

$$\sigma_W^2 = \frac{2}{\pi^{1/2}}\frac{W^2}{M}(WT_s)\left(\frac{N}{S}\right)^2 \tag{35}$$

The corresponding figure of merit for the first moment is

$$F_\Omega = \frac{2}{\pi^{1/2}}(WT_s)(N/S) \tag{36}$$

which behaves in the same way, as far as it dependence on T_s and N/S, as the multiple delay autocovariance approach we have discussed previously (Figure 4).

7. LEAST SQUARES PARAMETER ESTIMATION TECHNIQUE AS A ML ESTIMATION TECHNIQUE

We have mentioned before that parameter estimation by a least squares fitting of the theoretical shape of the spectrum is a ML technique. It is indeed a ML estimator which starts with the frequency spectrum estimates $s(\omega_i) \equiv s_i$ as the original set of random variables. Let $F(\mathbf{S}; \{P_i\})$ be the multivariate distribution function, where $\mathbf{S}$ is the set of spectral values $S(\omega_i) \equiv S_i$ in vector form. If $s(\omega_i)$ is obtained by averaging a sufficiently large number M_a of DFT's of weighted sections of the original time series, $F(\mathbf{S}, \{P_i\})$ is a Gaussian joint probability distribution function, and the logarithm of the likelihood function is given by

$$L(\{P_i\}; s_i) = -\ln|Q|^{-1} - \sum_{i,j}(s_i - S_i)(Q_{ij})^{-1}(s_j - S_j) + \text{const}$$

where $S_i = S(\omega_i; \{P_k\})$ is a known function of the unknown parameters P_k. We shall consider the covariance matrix Q known. Maximizing the likelihood function L is equivalent to minimizing the quadratic expression, namely to solve the set

$$\frac{\partial}{\partial P_k}\sum_{i,j}(s_i - S_i)(Q)_{ij}^{-1}(s_j - S_j) = 0 \qquad \forall\ k$$

It is known that if the size of the time window in the DFT is large with respect to the correlation time, the variances of $(S_i - s_i)$ are independent, and Q_{ij} is diagonal with elements σ_{ii}^2. The problem is then reduced to solve the set

$$\frac{\partial}{\partial P_k} \sum_i [s_i - S_i(\{P_m\})]^2 \frac{1}{\sigma_{ii}^2} = 0 \qquad \forall\, k$$

But this is exactly the starting point of a least squares estimation technique provided that each element $(s_i - S_i)^2$ in the quadratic expression is weighted by the inverse of their expected variance.

Note that the set of parameters is not limited to P, ω, and W. The parameter estimation procedure used for the Arecibo ST data [*Sato and Woodman*, 1982], for instance, fits up to 11 parameters.

8. CONCLUSIONS

The single delay autocorrelation approach is a very simple and statistical efficient estimator for MST radars and should be used for real-time processing of MST radar signals, whenever the complexity and cost of the installation are to be kept low. A coherent integrator is indispensable, since this reduces the processing capacity requirements and improves the S/N and final estimated variances. Nonfading clutter and noise should be estimated concurrently and accounted for. The technique does not allow for correcting other sources of interference.

If the complexity of the installation allows for the inclusion of an FFT processor, the full spectrum or correlation function should be evaluated, and the parameters evaluated using existing sophisticated algorithms. Parameters can be evaluated in this way with much improvement over the single delay correlation technique, especially under conditions of low S/N ratio and existing sources of interference like fading ground, ocean, or self-clutter. A least squares parameter estimation technique appears to be the best approach. Normally, only the estimation of the spectrum or correlation needs to be evaluated in real time.

REFERENCES

Bard, Y., *Nonlinear Parameter Estimation*, Academic, Orlando, Fla., 1974.

Blackman, R. B., and J. W. Tukey, *The Measurement of Power Spectra*, Dover, New York, 1958.

Cramer, H., *Mathematical Methods of Statistics*, Princeton University Press, Princeton, N. J., 1946.

Denenberg, J. N., The estimation of spectral moments, *Tech. Rep. 23*, 82 pp., Lab. for Atmos. Probing, Dep. of Geophys. Sci., Univ. of Chicago, 1971.

Levine, M. J., Power spectrum parameter estimation, *IEEE Trans. Inf. Theory*, *IT-11*, 100–107, 1965.

Miller, K. S., *Complex Gaussian Processes: An Introduction to Theory and Application*, Addison-Wesley, Reading, Mass., 1974.

Miller, K. S., and M. M. Rochwarger, A covariance approach to spectral moment estimation, *IEEE Trans. Inf. Theory*, *IT-18*, 588–596, 1972.

Rabiner, L. R., and N. Rader (Eds.), *Digital Signal Processing—Selection of Papers Edited by the Digital Signal Processing Committee of the IEEE*, Institute of Electrical and Electronics Engineers, New York, 1976.

Rummler, W. D., Introduction of a new estimator for velocity spectral parameters, *Tech. Memo. MM-68-4121-5*, Bell Telephone Lab., Whippany, N. J., 1968.

Sato, T., and R. F. Woodman, Spectral parameter estimation of CAT radar echoes in the presence of fading clutter, *Radio Sci.*, *17*(4), 817–826, 1982.

Woodman, R. F., Error analysis of multiple delay correlation function velocity estimates, paper presented at General Assembly, URSI, Lima, Peru, 1975.

Woodman, R. F., High-altitude-resolution stratospheric measurements with the Arecibo 430-MHz radar, *Radio Sci.*, *15*(2), 417–422, 1980.

Woodman, R. F., and A. Guillén, Radar observations of winds and turbulence in the stratosphere and mesosphere, *J. Atmos. Sci.*, *31*(2), 493–505, 1974.

Woodman, R. F., and T. Hagfors, Methods for the measurement of vertical ionospheric motions near the magnetic equator by incoherent scattering, *J. Geophys. Res.*, *74*, 1205–1212, 1969.

Zrnic', D. S., Estimation of spectral moments for weather echoes, *IEEE Trans. Geosci. Electron.*, *GE-17*(4), 113–128, 1979.

R. F. Woodman, Instituto Geofísico del Perú, Apartado 3747, Lima, Peru.