8.7A SPECTRAL MOMENT ESTIMATION IN MST RADARS

R. F. Woodman

Instituto Geofisico del Peru
Aportado 3747
Lima, Peru

INTRODUCTION

Due to 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, we will consider 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, $F(\omega)$. Because of the Gaussian distribution of velocities in the turbulent scatter volume, we know the shape of these functions: they are also Gaussians. Thus, the processes we will be discussing are Gaussian stationary processes with 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.

A Gaussian power spectrum has the form

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

It is fully defined by the value of three parameters: P, Ω, W. They correspond to the total power, the frequency shift and the spectral width, respectively. They 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, $\langle \mu^2 \rangle^{\frac{1}{2}}$, under certain conditions).

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

$$P = \int s(\omega)\, d\omega \qquad (2)$$

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

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

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 can not 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 we have a problem, since there are two criteria for goodness one would like to satisfy: the estimator should be good from an 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 an statistical point, as we will see when we talk about the Maximum-Likelihood (M-L) 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 limiting, not so much because of the time scale of the processes, which are relatively slow, but for the large number of parallel channels one has to process. The demands are very large if one is after the whole MST region with high altitude resolution.

We can limit the scope of our paper, if we limit ourself to representative techniques which have been actually implemented in MST radars. We shall do this, but include also some discussion about M-L 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 references; the reader will find a very extensive list of references in Zrnic's review.

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 M-L estimator concept and discuss the limits of performance they define.

MOMENT ESTIMATORS VIA POWER SPECTRUM

The most straightforward estimators of the three parameters of interest is 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 points of frequencies.

The definitions suggest the following estimators, P', i' 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 RADAR, 1976).

We would like to point out, that unless the sampling frequency, $1/T_s$, is larger than the mean frequency plus a few spectral widths, $2\Pi(\Omega + 2W)$, Ω' and W^2 as given by (6) to (7) would be biased because of aliasing. This bias can be reduced if we assume periodicity 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^j such that $\Omega' = \omega_j + \omega_\varepsilon$, ω_ε is evaluated from

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

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) to (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 narrower than the 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, but 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 consist essentially of a single spectral line at the origin with a strength which depends on the ground shielding of the radar. At tropo- 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 strength width. This results from the slight propagation fading of the clutter. 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 non-fading 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 signal is frequency shifted by a magnitude larger than its width. This occurs frequently and presents difficulties only when one is looking too close to vertical, or the medium velocity is too slow (of the order of 1 m/sec or less). Another technique takes advantage of the symmetry of the ground clutter component. It consists of evaluating the antisymmetric component of the spectrum, replace the negative powers by zero, and evaluate 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 on the processing speed.

As far as the estimating algorithm, most modern procedures use a Fast Fourier Transform. This is an efficient way of doing it and should always be pursued, unless one has a hardwired autocorrelator. One should always use algorithms especially 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 two orders of magnitude higher than the maximum frequency content of the signals. This produces high redundance in the sampling and calls for some signal filtering; not so much to increase the system sensitivity -- as one sometime reads or hears -- as for reducing the information input into the spectrum system and the amount of signal processing. As it is well known, an FFT evaluation takes, N ln N additions and computations. A reduction, of let us say, a factor of 256 in the number of points, speeds up processing by a factor of 2000.

The simplest and easiest filter to implement digitally is a boxcar integrator (coherent integration). This simply integrates N 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 undersampling and consequent aliasing can be allowed, if (8) and (9) is used for the evaluation of Ω' and σ; 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 DFT or the time width of the weighting function (Hanning window, etc.). The latter should not be much longer than the correlation width, say 2 or 4 times the half correlation time, since this will give us 4 or 8 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^2, σ^2, σ^2 of the estimated values with respect to their expectations. $\sigma_{Pi}^2 = \xi(P_i - \langle P_i \rangle)^2$. This in general depends on the algorithm used for the evaluation of $S'(\bar{\omega})$. We will quote here the results obtained by MILLER (1974).

He gives us a simple expression for the variances for the case of a Gaussian-shaped spectrum with no additive noise. The derivation was made assuming a continuous time series weighted with a Gaussian window of width $T/2\pi$. The variances, using our notation, are given by

$$\sigma_{\Omega}^{2} = \frac{W^2}{2N[1 + (WT)^2]^{\frac{1}{2}}} \tag{10}$$

$$\sigma_{\omega} = \frac{3W^2}{16N[1 + (WT)^2]^{\frac{1}{2}}} \tag{11}$$

Usually $WT > 1$, since it is a good practice. Also we have that for a given observation time T_o, the number of DTFs is given approximately by $T_o/2T$. In terms of T_o we can then write

$$F_{\Omega} \equiv \frac{\sigma_{\Omega} \sqrt{T_o W}}{W} \approx 1 \tag{12}$$

$$F_{W} \equiv \frac{\sigma_{\omega} \sqrt{T_o W}}{W} \approx \frac{3}{8} \tag{13}$$

which are figures of merit, that serve to compare different techniques. They normalize the variances with respect to the width of the spectra and include the dependence on the square root of the number of degrees of freedom (number of independent samples) which should be common to any estimator.

One can improve on the estimators 5, 8 and 9 with additional effort. Following a rule that one should not use data that carries no information, one should use only those points in the spectrum, W_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 DFTs for every altitude. This is a time-consuming operation, but fortunately MST echoes are slow, specially 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, producing time series close to ten times faster than a 50-MHz radar and, therefore, ten 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 an FFT processor.

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 ot as much information as one has and ask only what one ignores. 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's. 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

$$\varepsilon^2 = \sum_{i=1}^{N} A_i \left[S'(\omega_i) - S(\omega_i; P, \Omega, W)\right]^2 \qquad (14)$$

The problem is nonlinear in the unknowns, P, Ω, W and involve 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. In fact, they used the technique to estimate up to 8 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 sidelobes. 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 square 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. It is feasible to perform this additional processing in real time by doubling the processing capacity. Although, for many applications, it would not be necessary to perform the nonlinear estimation in real time.

THE AUTOCOVARIANCE OR AUTOCORRELATION APPROACH

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

$$P' = \rho'(o) = \frac{1}{N} \sum_{i=1}^{N} x_i x_i^* \qquad (15)$$

$$\rho'(\tau) = \frac{1}{N-\tau} \sum_{i=1}^{N-\tau} x_i x_{i+\tau}^* \qquad (16)$$

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

$$\Omega' = \frac{\phi(\tau_1)}{\tau_1} \tag{17}$$

$$W^2 = 2\,\frac{1 - |\rho'(\tau_1)|/(\rho(o) - \text{Noise}}{\tau_1^2} \tag{18}$$

The technique takes advantage of the relationship that exists between the nth derivation of the correlation function evaluated at the origin and the nth 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 GUILLEN (1974) for stratospheric and mesospheric applications. The technique is in much use today by the weather radar community, but, 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 lnN in the case of spectral moment estimation (where N is the number of spectral points). The variance of this approach, seemly surprising, is comparable to the one 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 12) yields the same value as the one evaluated by integrating the area of the frequency spectrum (equation 5). This is only a particular case, corresponding to the zeroeth 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 Ns and small resultant values of $\tau^2\sigma_\Omega^2$(<<1 radians):

$$\sigma_\Omega^2 = \frac{\rho(o) - \rho^2(\tau)}{2\tau^2 N\,\rho^2(\tau)} \tag{19}$$

It is interesting to compare the figure of merit of this approach as with that of 12. For large S/N ratios and Gaussian-shaped autocorrelation functions, 18 takes its best values at small τs.

$$\sigma_\Omega^2 = \frac{W^2}{2N} \tag{20}$$

For a given observation time T_o, the number of (sufficiently) independent estimates is $N \simeq T_o W$. Hence,

$$F_\Omega \simeq 1/\sqrt{2} \tag{21}$$

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 is close to optimum only when the signal-to-noise ratio

is high. This relative 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 equations (5), (6) and (7), but not of 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 is 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 $\rho(o)$ and $\rho(\tau)$. These estimates can be obtained by the same methods described before for the spectral moment approach.

Going from (3) and (4) to (17) and (18) involves approximating the derivative of $\rho(\tau)$ by finite differences between $\rho(o)$ and $\rho(\tau)$. This presents a bias which could become significant for relatively large τs (MILLER, 1972). Fortunately, in the case of symmetric spectrum, equation (17) is an equality and the bias disappears. This is important since optimum values of τ, for noisy signals, are not close to the origin.

We are reproducing here, two graphs (Figures 1 and 2) from MILLER (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 $\tau(\equiv h$, 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_i' and W_i' 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 pre-filtering 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 it is to be expected, optimum results are obtained using a matched filter, matched to the shape of the signal spectrum. But, a boxcar integator (coherent integration) produces similar results and it is much easier to implement. It should be kept in mind, in any case, that filtering could be a source of Ω and W biasing. This bias can be computed theoretically and should be corrected.

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 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 (17) and (18) for different values τ_i of τ. We can always obtain a new estimate Ω_a, W_a through

$$\Omega_a = \sum_{i=1}^{M} C_i \Omega_i \tag{22}$$

$$W_a = \sum_{j=1}^{M} C_j W_j \tag{23}$$

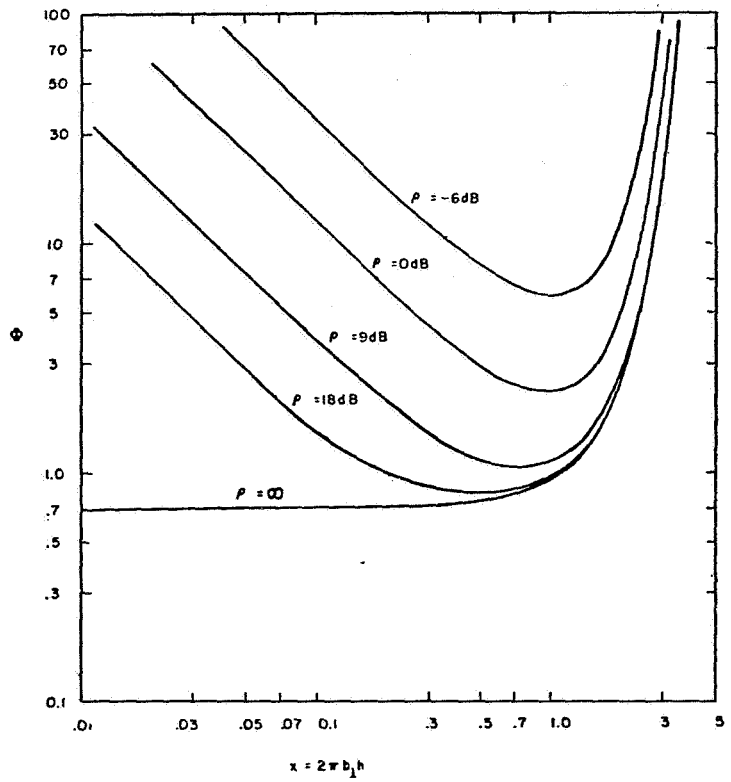

Figure 1. Normalized standard deviation of mean-frequency estimator versus pulse-pair spacing.

where C_i and C_j are weights properly selected to minimize the variances of Ω_a and W_a, and normalized such that $\Sigma C_i = \Sigma C_j = 1$. 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 points M and correlation function shapes. Figure 3 depicts the optimum set of weights for two S/N ratios, for a Gaussian-shaped autocorrelation function sampled at 32 points with a spacing of 0.1 (of the typical width). The set 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 variance of Ω_a as a function of the coordinate of the last sample (in typical width units). There are two groups corresponding to different S/N ratios. In each group there are 6 curves corresponding to 1, 2, 4 ... 32 samples. 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 variance with 32 points at optimum delay and a single point close to the origin. There is a 60% difference (30% for the standard deviation) in going from 1 to 2 points, and an additional 50% (25% for the sd) in going from 2 to 32. This last improvement, is certainly not worth the effort. The increase from 1 to 2

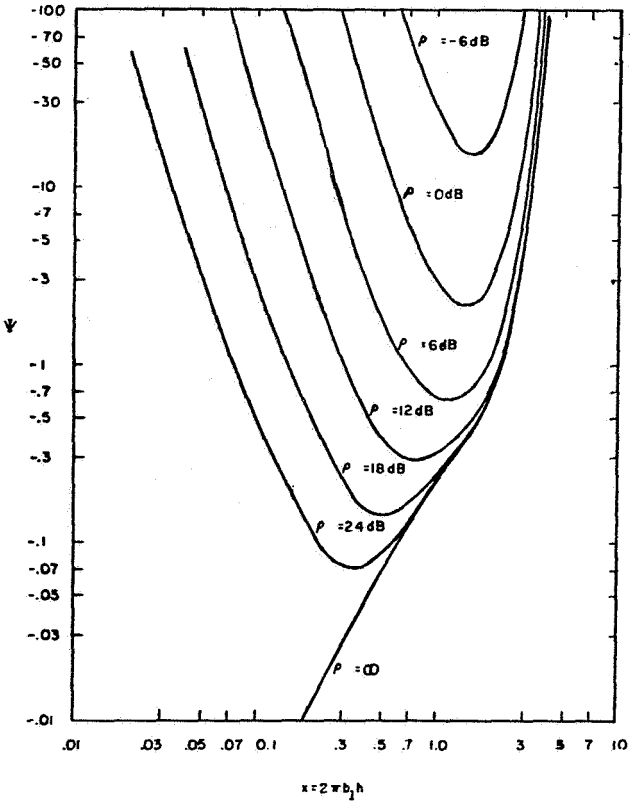

Figure 2. Normalized standard deviation of the estimate of frequency spread versus pulse-pair spacing.

could be justified if the redundancy is used to check the existence of unexpected interference.

On the other hand, we see that, for high S/N ratios, the variance is roughly inversely proportional to the number of sample points. This is not surprising since the estimates $\{\Omega_i\}$ have statistically independent errors with respect to the real Ω.

We can conclude then that Ω, as defined in (17), is a good estimator from a statistical as well as from a practical point of view when the S/N is better than one, but it is far from optimum when the signal-to-noise ratio is low.

Notice that the above discussion assumes that the sampling in the time series is such that the spacing for 32 points is optimum, i.e., that approximately 32 or N points can be fit in approximately a correlation time. Otherwise, a gain proportional to the number of samples cannot be achieved. Correlation samples which fall at points where the correlation is low do not contribute to improve the accuracy of the estimate.

It should be mentioned that the deterioration of the single delay covariance approach should not be held as an argument in favour of the sampled 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 moment approach, including the case of noise signals, as it was quoted before (RUMMLER, 1968).

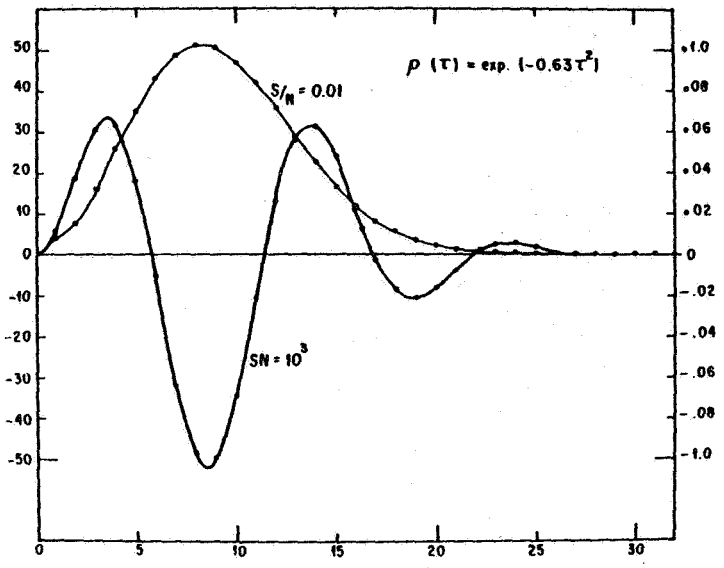

Figure 3. Optimum weighting coefficients c_i, to minimize $(\Omega_a - \Omega)^2$, where $\Omega_a = \sum_{i=1}^{M} c_i \Omega_i$. Two cases, one for low and the other for high S/N ratio, are considered.

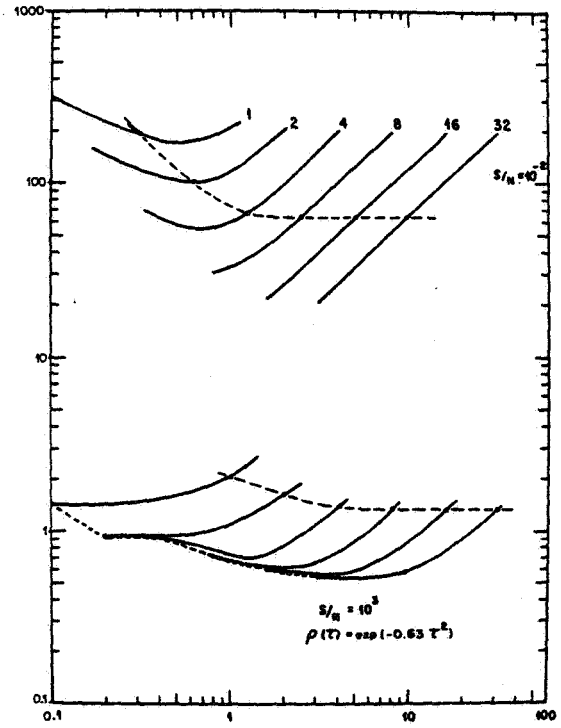

Figure 4. Variances of Ω_a, when evaluated in accordance to the weight coefficients shown on Figure 3. Labels correspond to the number of Ω_i's, and the absisa shows the delay of the last point, i = M.

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.

MAXIMUM LIKELIHOOD ESTIMATORS AND BOUNDS

Given a set or sequence (random process) of N observables x_i with an N joint probability function $F(\underset{\sim}{X};\{A\})$ such that $F(\underset{\sim}{X};\{A\}) = g(\underset{\sim}{X}<\underset{\sim}{X}<\underset{\sim}{X}+d^n\underset{\sim}{X}/\{A\})$ where $\{A\}$ is a set of parameters. It is possible to find practically an innumerable number of estimators of $\{A\}$ and estimates of $\{A_i\}$ as a function of the observables. For these estimates to be of practical use they must meet the condition that $\langle A_i' \rangle = A_i$ and that the variances $\sigma_a^2 = \langle (A_i' - A_i)^2 \rangle$ be small. For a given sample $\underset{\sim}{x}$ of the process, we can form a function $F(\underset{\sim}{x};\{A\})$ and let $\{A\}$ vary. There will be a value of $\{A\}$ for which $F(\underset{\sim}{x};\{A\})$ is a maximum in $\{A\}$ space. This value is called the maximum likelihood 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 etimators, on the other hand, the theory gives us formal expression for the ML variances, which can be used to compare the "efficiency" of a given estimator. It is possible, in the case of large N processes with a Gaussian-shaped spectrum plus white noise, and using justifiable approximations, to obtain explicit expressions for these bounds. Zrnic, for instance, using a ML approach, finds the following lower bounds.

$$\sigma_\Omega^2 \geq \frac{W^2 (WT_s/2\pi)^2}{M(1 - (WT_s/2\pi)^2)^{\frac{1}{2}}} \tag{24}$$

when the noise level is zero, and

$$\sigma_\Omega^2 \geq \frac{W^2}{M} 4\sqrt{\pi}\, (WT_s/2\pi)\, (N/S)^2 \tag{25}$$

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

We should state, though, that we find equation (24) disturbing, since for a given observation time $T_o = T_s M$, 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} gives redundant information, and should not improve the variance of any estimator. There is no explicit indication on 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 Ns (MILLER, 1972). It turns out to be the same as the covariance approach heuristically described by RUMMLER (1968) and WOODMAN and HAGFORS (1969) and discussed previously.

It is also possible to use an M-L 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 an M-L estimate of the parameters P, Ω, W and N, 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 contains all the desired information about the

process, are then used in an M-L 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 Zrnic's review 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

$$\sigma_\rho^2 \geq \frac{P^2}{M}\left(2.25 - \frac{30}{(2\pi)^2}(WT_s)^2 + \frac{180}{(2\pi)^4}(WT_s)^4\right) \tag{26}$$

$$\sigma_\Omega^2 \geq \frac{3}{\pi^2}\frac{W^2}{M}(WT_s)^2 \tag{27}$$

$$\sigma_w^2 \geq \frac{45}{4\pi^4}\frac{W^2}{M}(WT_s)^4 \tag{28}$$

These bounds are valid for large signal-to-noise ratios.

We should notice that (27) gives about the same lower bounds as (24), which means that, at least for the frequency shift variance, this approach can be as good as the M-L approach which starts with the observational time series $\underset{\sim}{x}_i$. In terms of our figures of merit we can write

$$F_\Omega = \frac{\sqrt{3}}{\pi}(WT_s)^{3/2} \tag{29}$$

$$F_W = \sqrt{45}\,\frac{1}{2\pi^2}(WT_s)^2 \tag{30}$$

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 signal delay frequency shift estimators is comparable to both M-L 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 (high sampling) rates redundant should eventually produce oversampling, which should not decrease the variance of our estimates. Figure 3, 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:

$$\sigma_\rho^2 = \frac{3}{2\sqrt{\pi}}(WT_s)\frac{N^2}{M} \tag{31}$$

$$\sigma_\Omega^2 = \frac{2W^2}{\sqrt{\pi}M}(WT_s)\left(\frac{N}{S}\right)^2 \tag{32}$$

$$\sigma_\omega^2 = \frac{2}{\sqrt{\pi}}\frac{W^2}{M}(WT_s)\left(\frac{N}{S}\right)^2 \tag{33}$$

The corresponding figure of merit for the first moment is:

$$F_\Omega = \frac{2}{\sqrt{\pi}}(WT_s)\left(\frac{N}{S}\right), \tag{34}$$

which behaves in the same way as far as its dependence on T_s and $(\frac{N}{S})$, as the multiple delay autocovariance approach we have discussed previously (Figure 2).

We have mentioned before, that parameter estimation by a least square fitting of the theoretical shape of the spectrum is an M-L technique. It is indeed an M-L estimator which starts with the frequency spectrum estimates $s(\omega_i) \equiv s_i$ as the original set of random variables. Let $F(\underset{\sim}{S};\{P_i\})$ be the multivariance distribution function, where $\underset{\sim}{S}$ 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 DFTs of weighted sections of the original time series, $F(\underset{\sim}{S}, \{P_i\})$ 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

$$\frac{\partial}{\partial P_k} \sum_{i,j} (s_i - S_i)\,(Q)^{-1}_{ij}(s_j - S_j) = 0 \qquad \text{for all ks} \tag{35}$$

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^2_{sii}. The problem is then reduced to solve the set

$$\frac{\partial}{\partial P_k} \sum_{i} (s_i - S_i(\{P_m\}))^2 \cdot \frac{1}{\sigma^2_{ii}} = 0 \qquad \text{for all ks} \tag{36}$$

But this is exactly the starting point of a least squared 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, W. The parameter estimation procedure used for the Arecibo ST data, (SATO and WOODMAN, 1982), for instance, fits up to eleven parameters.

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 is 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 specially under conditions of low S/N ratio and existence of sources of interference like fading ground, ocean or self clutter. Normally only the

estimation of the spectrum or correlation needs to be evaluated in real time.

REFERENCES

Bard, Y. (1974), Nonlinear Parameter Estimation, Academic Press, New York.

Blackman, R. B. and J. W. Tukey (1958), The Measurement of Power Spectra, Dover Publications, Inc., New York.

Cramer, H. (1946), Mathematical Methods of Statistics, Princeton University Press, Princeton.

Levine, M. J. (1965), Power spectrum parameter estimation, IEEE Trans. Inform. Theory, IT-11, 100-107.

Miller, K. S. (1974), Complex Gaussian Processes: An Introduction to Theory and Application, Addison-Wesley Publishing Co., Inc., Reading, Mass.

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

Rabiner, L. R. and N. Radar (editors) (1976), Digital signal processing, Selection of papers edited by the Digital Signal Processing Committee of the IEEE.

Rummler, W. D. (1968), Introduction of a new estimator for velocity spectral parameters, Bell Telephonse Labs., Whippany, NJ, Tech. Memo MM-68.

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

Woodman, R. F. (1975), Error analysis of multiple delay correlation function velocity estimates, Abstracts of URSI, General Assembly, Lima, Peru.

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

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

Woodman, R. F. and T. Hagfors (1969), Methods for the measurements of vertical ionospheric motions near the magnetic equator by incohering scattering, J. Geophys. Res., 75, 1205-1212.

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