CORRELOGRAMS FOR PACIFIC OCEAN WAVES

PHILIP RUDNICK
UNIVERSITY OF CALIFORNIA MARINE PHYSICAL LABORATORY

1. Introduction

This paper presents some examples of correlograms for surface waves on the ocean. Our interest in this subject parallels that of Seiwell at Woods Hole [1]. Possible uses and limitations of correlograms in the statistical treatment of wave data are suggested in the final section. The ideas may, of course, be pertinent for other processes with similar spectral character.

2. Computer facilities

Some means for rapid computation is a practical prerequisite to any extensive work with correlograms from empirical data. Even very modest amounts of data lead to prohibitive labor if only a desk calculator is used. At the Marine Physical Laboratory we have developed electronic analog computers for obtaining correlograms and also for obtaining the statistical distribution for instantaneous values of any signal [2].

Signals are furnished to the correlator through a simple resistance-capacity high pass filter whose limit corresponds to a period equal to the longest correlation interval which is to be used. Hence the correlograms obtained are comparatively free of the effect of slow fluctuation, or trend, in the data. The longest retained periods are usually not over 5 to 10 percent of the total sample length.

3. Nature of data

The wave records from which correlograms have been taken were made available through the courtesy of wave research groups, at either Scripps Institution of Oceanography or the University of California at Berkeley. Three recording stations are represented, at La Jolla and Oceanside on the southern California coast, and at Guam. In all cases the datum furnished is fluctuation of pressure near the ocean bottom at a depth ranging from 40 feet to 140 feet. The depth fixes a short period limit for the recorded waves, of 6 to 10 seconds. On the other hand a long period limit of the order of 20 seconds also usually exists, either for natural or instrumental reasons. Thus the total spectral range involved is generally not over two octaves.

The material collected is merely that which was first available. While it presumably typifies fairly common conditions, it does not pretend to any systematic inclusion of varied sea surface states.

This work represents one of the results of research carried out under contract with the Bureau of Ships, Navy Department.

4. Examples of correlograms

Any of the some fifty correlograms so far computed may be roughly described as a curve which oscillates under a diminishing envelope. Thus a period for the oscillation (which will be called a dominant period for the wave spectrum) and some sort of decay time for the envelope may usually be regarded as the two first order parameters of the correlogram.

Three groups of wave correlograms have been selected as including about the most rapid and the most gradual envelope decay among all so far determined. The records for these groups are described in the following.

Series I. Six 21 minute records taken at six hour intervals from the Scripps pier, depth 140 feet, sea surface conditions not selected.

Series II. Seven 25 minute records at six hour intervals recorded by the Berkeley wave research group at Oceanside, depth 44 feet, records selected for prominent swell.

Series III. Three 24 minute records at six hour intervals, by the Berkeley group, at Guam, depth 60 feet, records selected for prominent swell. These are taken from a report by Miller [4], being there designated as records B, C, and D. The other records A, E, and F consecutive with these were not included here because of the greater complexity of their correlograms, which is discussed in the next section.

Table I gives some numerical characteristics of the correlograms and figure 1

TABLE I

WAVE CORRELOGRAMS, QUANTITATIVE SUMMARY

SERIES	I	II	III
Average maximum correlation for interval of:			
1 period	.46	.40	.57
2 periods	.21	.29	.49
3 periods	.14	.29	.44
4 periods	.11 (av.) (4–6 periods)	.21	.38
5 periods		.15	.30
6 periods		.15	.30
7 periods	.11 (av.) (7–9 periods)		.26
8 periods			
9 periods			
Greatest irregularity in spacing of successive maxima within individual correlogram:			
For most regular curve in group	±12%	±10%	±05%
For least regular curve in group	±30%	±25%	±10%
Abscissa of first maximum (dominant period), for curves of group, in chronological order:			
	13 sec.	14 sec.	19 sec.
	14 sec.	15 sec.	18 sec.
	15 sec.	16 sec.	18 sec.
	16 sec.	17 sec.	
	13 sec.	18 sec.	
	13.5 sec.	17.5 sec.	
		16 sec.	

shows two of the curves from each of the first two groups, each pair being selected for equality of dominant periods. The two curves from series I agree in the main through the first two to four periods, and not thereafter. This suggests that about the first three periods are the significant part of each curve and that the correlations computed for larger intervals are mainly sampling errors which would tend to diminish with increased length of original record. The curve of series II would

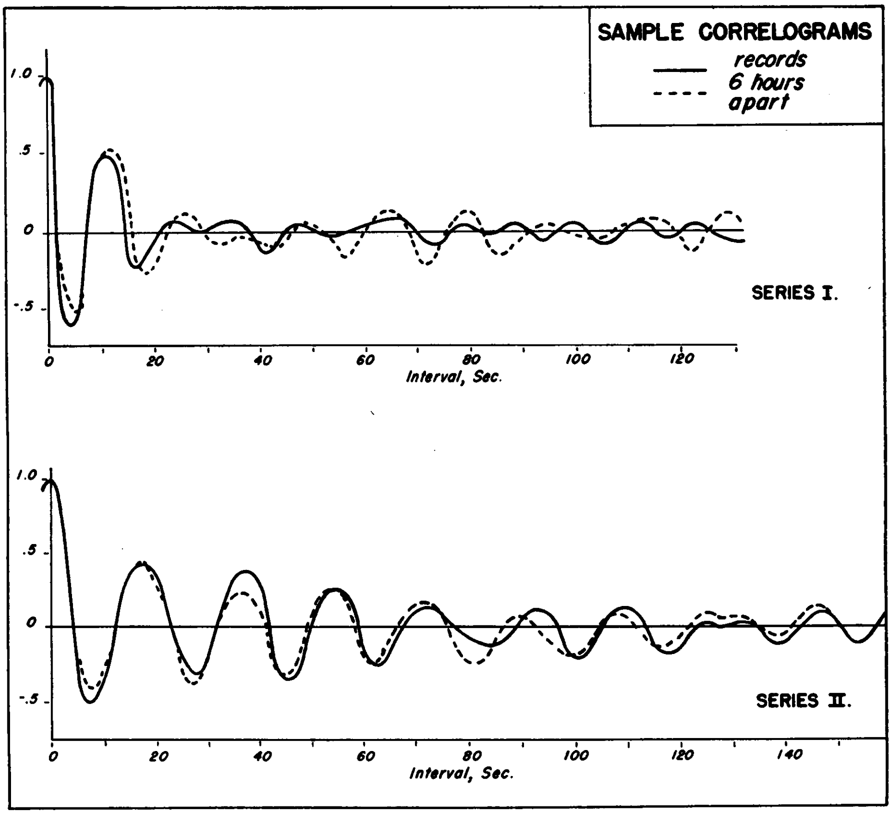

FIGURE 1

Wave correlograms. Examples from southern California coast. Series I is from unselected data, series II from data selected for prominent swell.

appear, on the same basis, significant for somewhat larger correlation intervals. The very important question of sampling errors is discussed in a later section.

Figure 2 shows an additional correlogram which is included as one for which the generating winds happened to be capable of a comparatively simple and unambiguous description. The record was made at Scripps Institution shortly after the passage of a cold front which was accompanied by a sharp rise of wind velocity and a corresponding increase in wave height, almost four fold in a period of about 12 hours. In these circumstances one would presume the postfrontal waves to have been generated almost entirely by the local wind. This has been confirmed by Mr. R. S. Arthur, who very kindly examined the pertinent meteorological charts and concluded that the local wind waves should indeed have been strongly predominant over any predictable swell from other generating areas, and that the generating condition corresponding to figure 2 was approximately a 15 knot wind limited by a duration of 12 hours.

5. Relation to spectra

Theoretically, a correlogram determines a spectrum, and conversely. The square of the Fourier amplitude spectrum (which will be called a power spectrum) and the correlogram are Fourier transforms of one another, aside from normalizing constants. Further, if a correlogram has the form $E(\Delta) \cos \omega_0 \Delta$, where Δ is the correlation interval and $E(\Delta)$ is an envelope function which has only moderate variations over the time $1/\omega_0$, then the corresponding power spectrum has the form

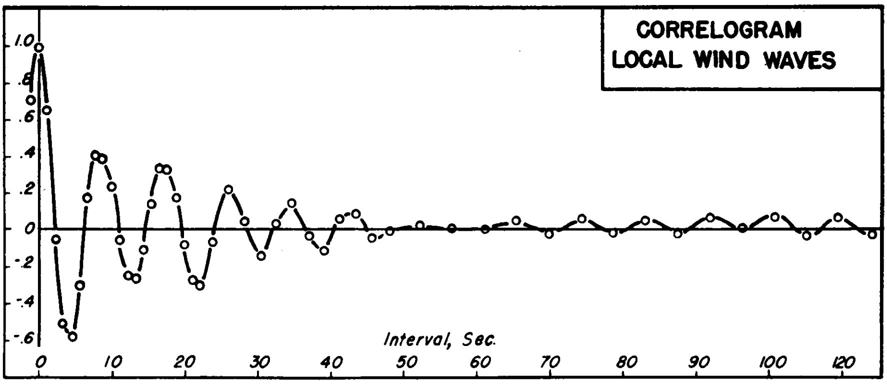

FIGURE 2

Correlogram of waves generated by local 15 knot wind in 12 hours

$F(\omega - \omega_0)$, where $F(\omega)$ is the Fourier transform of $E(\Delta)$. Thus a period of oscillation in a correlogram corresponds to the position of a central maximum in a spectral band, and the form of the correlation envelope fixes the band shape, the widths of the two being inversely related.

If the envelope of the series I correlogram (figure 1) is considered a negative exponential with decay of .46 in each period, the correlation function is representable as

$$\rho(\Delta) = e^{-|\Delta|/1.29T} \cos \frac{2\pi\Delta}{T}. \tag{1}$$

The corresponding spectrum is

$$\frac{1}{1+\frac{(f-f_0)^2}{(.12f_0)^2}}$$

where $f_0 = 1/T$. Here $.12f_0$ is the half width of the spectral band measured to either half power point. Thus the figure 1, series I, spectrum is a moderately narrow band including mainly periods in the range 13 sec. $\pm$ 12%; the bands for series II and III are even narrower, corresponding to their broader correlation envelopes.

A direct and fully detailed determination of one wave spectrum was made with a General Radio type 736A wave analyzer. A wave record of twenty-two minute duration, in variable width form on 16 mm film, was played back photoelectrically at greatly increased speed. The repetition rate was thirty per second, which is also the spacing between the various harmonic frequencies in the Fourier line spectrum. The wave analyzer was capable of resolving fully each of these harmonics and reading the amplitudes individually. The interesting part of the spectrum so measured is shown in figure 3, which also contains the correlogram for the same record. The detailed correspondence which should exist between these two has not

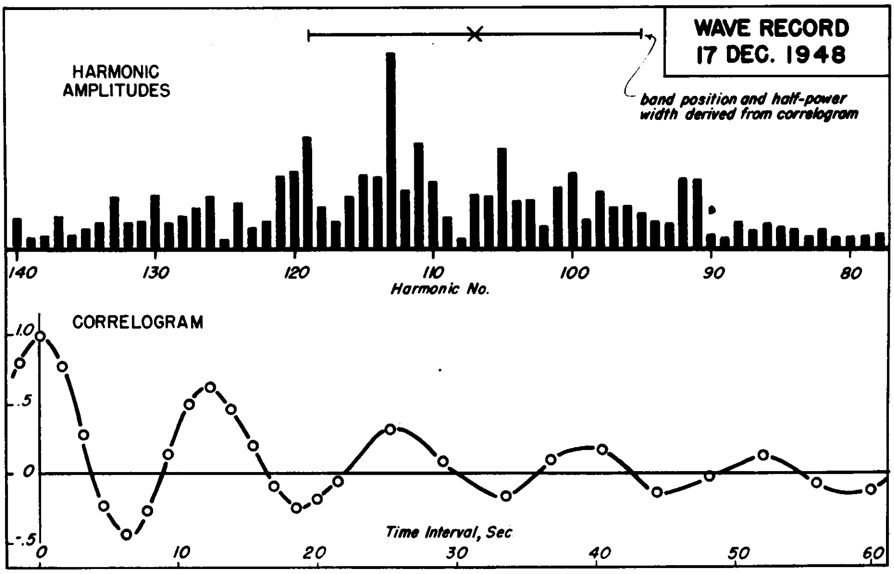

FIGURE 3

Comparison of exact Fourier spectrum of 22 minute wave record with its correlogram

been checked by calculation. However, the rough characterization of the spectrum based on dominant period and correlation envelope width is indicated in the figure. The agreement seems satisfactory. One appreciates that the exact Fourier spectrum for any specific sample requires smoothing before it can be regarded as corresponding to an infinitely long sample. The left hand end of the correlogram, involving relatively small correlation intervals and large correlations, determines the broad smoothed form of the spectrum; the residual correlations over large intervals determine the accidental detailed variations from one harmonic to the next. The discard of the right hand end of a correlogram may prove to be a useful way of obtaining a smoothed spectrum.* For example, the dominant periods listed in table I show a rather smooth and presumably real progression within each series. This might not be as readily discerned from spectra of the type of figures 4 to 6.

Wave spectra have sometimes been observed to show two maxima [3], [4].

* *Added in proof:* This point has recently been elaborated by M. S. Bartlett, "Periodogram analysis and continuous spectra," *Biometrika*, Vol. 37 (1950), pp. 1–16.

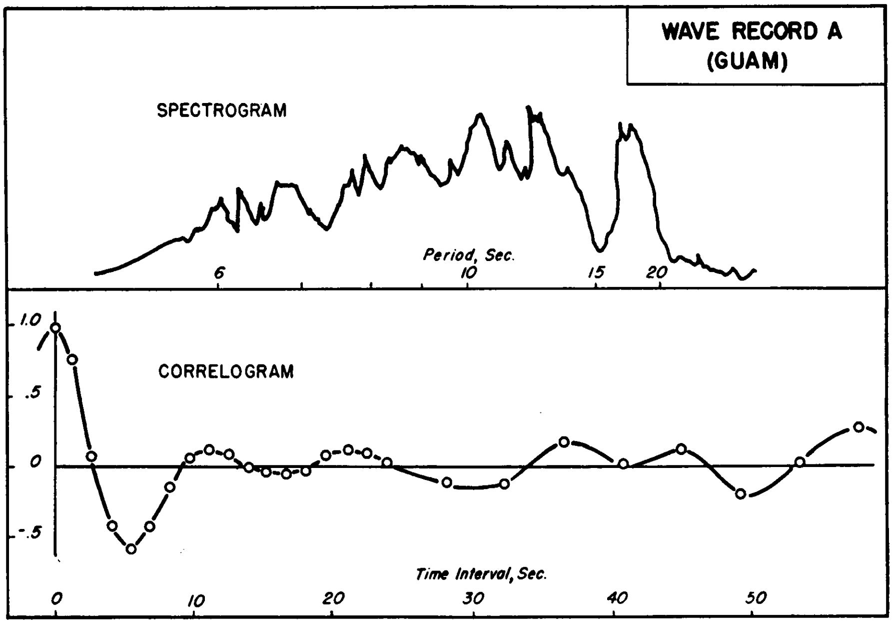

FIGURE 4

Comparison of wave spectrogram with correlogram. Example in which high frequency component is more prominent.

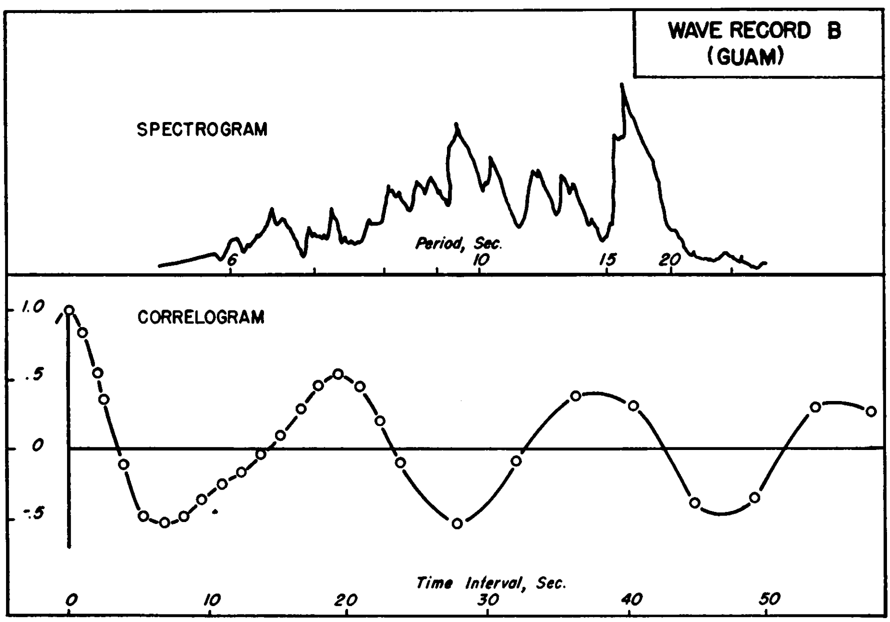

FIGURE 5

Comparison of wave spectrogram with correlogram. Example in which low frequency component is more prominent.

Correlograms have been computed from the six wave records made by the Berkeley group at Guam which are discussed by Miller [4]. Three of these and the corresponding amplitude spectra are shown in figures 4, 5, and 6. The spectra are taken from Miller's report; they were obtained with the Woods Hole wave spectrograph. The two maxima in each spectrum should give rise to superposed oscillations of two distinct periods in the correlogram; this is clearly the case in figure 6 but less evident in figures 4 and 5. The difference presumably arises from the fact that the

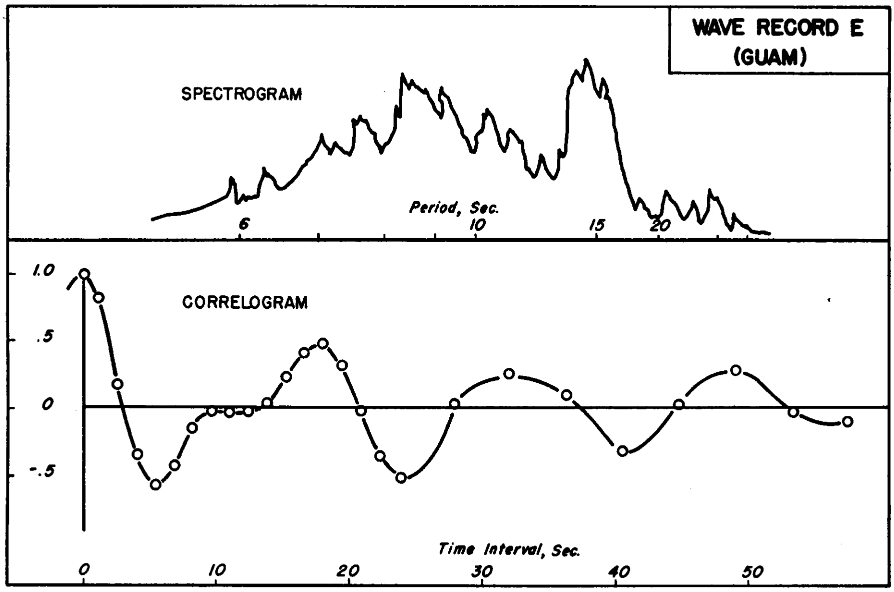

FIGURE 6

Comparison of wave spectrum with correlogram in which both spectral components are evident.

two components in the correlogram have a ratio which is the square of the amplitude ratio, so that a moderate inequality in two components in an amplitude spectrum is accentuated in the correlogram, the influence of the minor component being suppressed.

6. Sampling errors

Bartlett [5] has approximated the sampling errors in a correlation function. His results are expressed in terms of the true correlation function, and are based on hypotheses regarding the original data which are formally rather broad, so that one may hope the results to be widely applicable to empirical data. Bartlett's result, after two minor corrections, may be written in the form

$$(2)\quad \overline{\Delta(s)\Delta(t)} = \frac{1}{n}[R(s-t) + R(s+t) + 2\rho(s)\rho(t) - 2\rho(s)R(t) - 2\rho(t)R(s)].$$

In equation (2), the sampling error is

$$(3)\qquad \Delta(s) = r(s) - \rho(s)$$

where $\rho(s)$ is the true correlation function (obtainable from a time series of indefinitely great length) and $r(s)$ is a correlation function computed from a sample of the time series, of length T. The bar over the left member of (2) denotes an average over many samples, each of length T, and n is a sort of sampling number defined by

$$\frac{T}{n} = \int_{-\infty}^{\infty} \rho^2(t)\,dt = \tau\,. \tag{4}$$

The integral denoted above by τ is a time which may be regarded as measuring the extent over which correlations exist. Finally, $R(s)$ is the correlation function of the correlation function $\rho(s)$,

$$R(s) = \frac{1}{\tau}\int_{-\infty}^{\infty} \rho(t)\,\rho(s+t)\,dt\,. \tag{5}$$

Some special cases of equation (2) are of particular interest. If $s = t$, the left member becomes the mean squared error of $r(s)$;

$$\overline{\Delta^2(s)} = \frac{1}{n}[1 + R(2s) + 2\rho^2(s) - 4\rho(s)R(s)]\,. \tag{6}$$

The quantity in square brackets in the right member of (6) assumes the values 0 and 1 respectively for $s = 0$ and $s = \infty$; for intermediate values of s there appears little likelihood that the quantity will materially exceed 1. Thus the errors in general grow with s from zero to their ultimate size. Returning to equation (2), if s and t are both taken so large that $\rho(s)$, $\rho(t)$, $R(s)$, and $R(t)$ all tend to zero, we have

$$\overline{\Delta(s)\Delta(t)} = \frac{1}{n}R(s-t) \tag{7}$$

and

$$\overline{\Delta^2(s)} = \frac{1}{n} \tag{8}$$

describing what may be called the residual errors, or correlations calculated for intervals so great that no true correlations exist. Equation (7) shows that if the spectrum of the original time series contains a prominent period, the same period will appear in $\rho(s)$, $R(s)$, and hence in the residual errors $\Delta(s)$. Thus for example in figure 1 or 2, although the spacing of maxima of the small correlations toward the right hand end of each curve tends to agree with the intervals between the larger maxima at the left, this is not evidence for the reality of the small correlations, which may be residual errors. A further consequence of equation (7) is that, the greater the true correlations over a small number of dominant periods, the more highly correlated and the smoother the error curve. In comparing our computed correlograms, it is quite generally noticeable that those with highest early correlations are smoothest and most regular throughout their length, giving an appearance of significance which may be to some extent misleading, according to equation (7).

While this theory does not enable one to proceed deductively from a computed correlation function $r(s)$ to a true correlation function $\rho(s)$, it does impose a useful control on any $\rho(s)$ which one may hypothesize on the basis of a known $r(s)$. This may be illustrated by the series I correlograms of table I and figure 1. If one assumes equation (1) to give the true correlation function, one obtains $\tau = 9.6$ sec., $n = 133$, and the root mean square residual error $1/\sqrt{n} = .087$. This agrees well with .096, which is the root mean square value of all computed correlations for series I for all correlation intervals greater than three periods. Thus for the 21 minute sample length used, there is little ground for ascribing significance to the correlograms of series I beyond the third period, or about 3 per cent of the sample length.

On the basis of greater regularity, one tends to ascribe more extended significance to the correlograms of series II and III. But this must be done with caution. If one determines τ for these series by accepting the first five periods of the computed correlation function as free from error, the corresponding r.m.s. residual errors are .086 and .128, or if the curves of residual error are roughly sinusoidal, their maxima would average .12 and .18 for series II and III respectively. The last correlation maxima recorded in table I, .15 and .26, thus appear in very considerable part uncertain, and it would seem of little value to extend calculations beyond the 6 or 7 periods shown in table I, or 7% to 10% of the length of the original record. In particular, the agreement between the two series II curves shown in figure 1 at their extreme right ends should be considered in part fortuitous; one should recall that the curves shown were selected for similarity from six possible pairs.

A further investigation of sampling errors has been based on two wave records each of about 10 hours duration. Correlograms were computed for each half hour portion and for each 2 or 2.5 hour portion, as well as for each entire record. Figure 7 shows one of the over all correlograms and three of the curves based on half hour samples of the same record. There are also shown the positions of all maxima and minima for 15 of the half hour correlograms. One may see that the early maxima (at 8 to 9 sec., or 16 to 18 sec.) are sufficiently consistent in position from one half hour to the next to show small progressive changes over the 10 hour interval. On the other hand, for larger correlation intervals (20 to 30 sec. and over) the variations from one half hour to the next rapidly become larger and more random, and appear to be largely sampling errors combined with some instrumental computer errors. The mean squared deviation of the half hour correlations (over intervals greater than 20 sec.) from their 2 or 2.5 hour means was calculated and reduced by the square of .02, which is an estimate of random computer errors based on day to day reproducibility of computed correlations. The remainder gives an r.m.s. value of .048 which is ascribed to sampling fluctuation. Based on $\tau = 4.1$ sec., the Bartlett formula predicts .042 for this quantity, again with satisfactory agreement. Comparable results were found for the other 10 hour record examined.

7. Statistical distribution

Bartlett points out that if the instantaneous values in the original time series show a multivariate normal distribution, the sampling error results follow with fewer other assumptions as to the nature of the series. Almost any statistical consideration of a time series may be influenced by knowledge of its distribution function. Hence it seems appropriate to include here some results obtained with the distribution computer mentioned earlier. Figure 8 shows the cumulative distribution

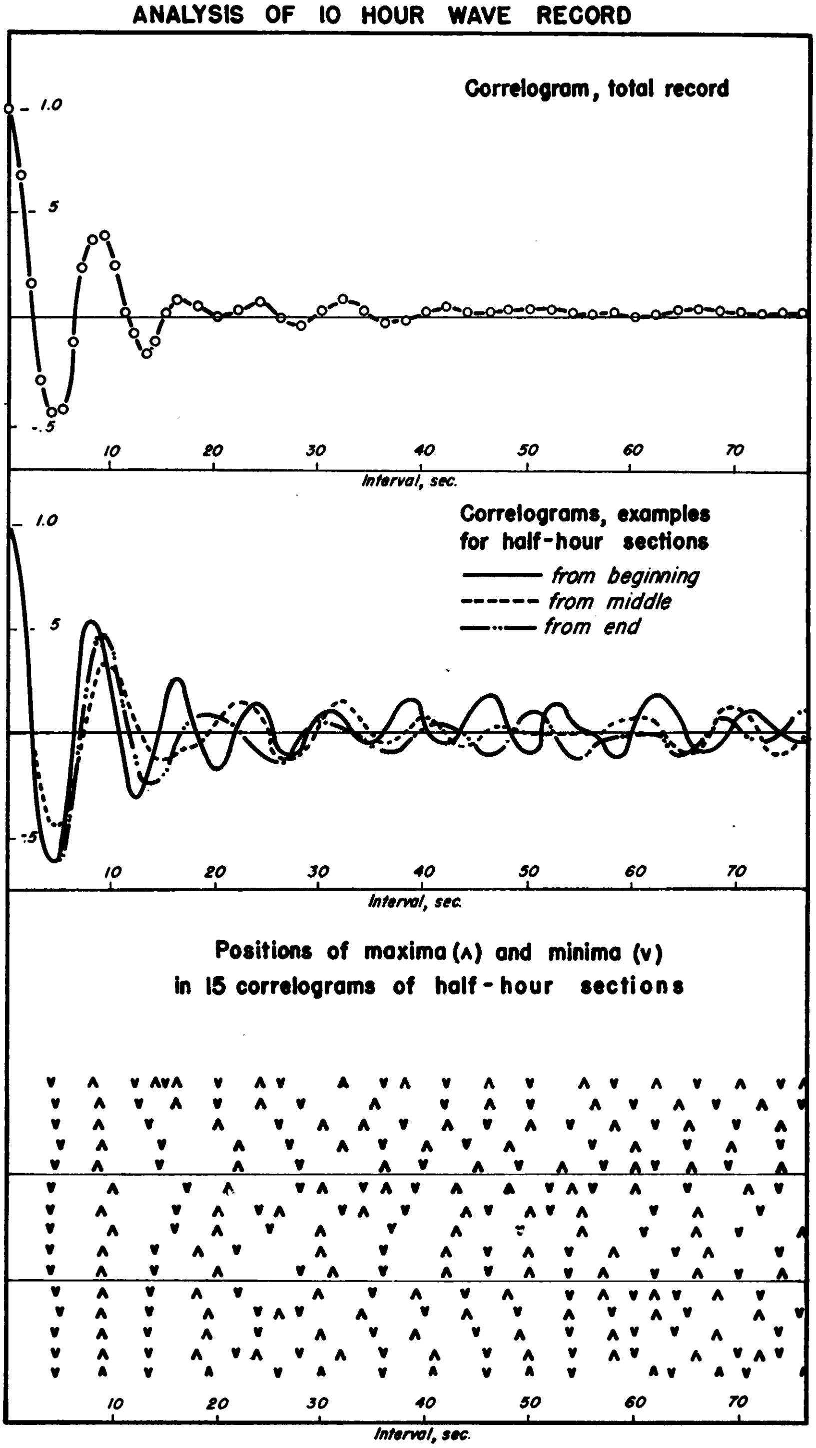

FIGURE 7

Comparison of correlograms based on half hour portions of a 10 hour wave record of displacements taken for all the series I records. The form of the curve is very closely normal. This result should be considered subject to confirmation, since a good deal of the available data has not yet been examined.

8. Summary

The wave correlograms which have been obtained have sampling errors in the range .05 to .13 when computed from records of 20 to 30 minutes duration. This is judged either by comparing successive correlation functions from the same source of data, or by Bartlett's formula. In this case the greatest interval over which significant correlations can be demonstrated ranged roughly from 40 sec. to 100 sec. The statistical distribution of instantaneous values in the wave records is found to be normal.

The correlograms frequently take the form of a nearly periodic oscillation under

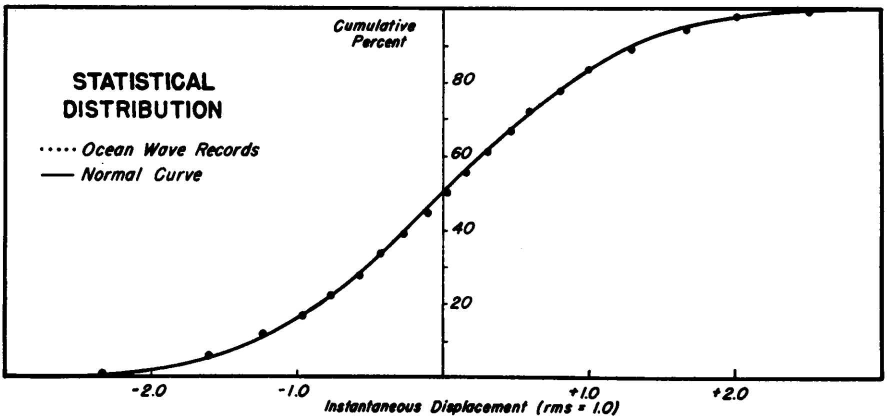

FIGURE 8

Cumulative distribution curve for instantaneous displacements in wave pressure records

a descending envelope, corresponding to a spectrum containing a single maximum, or band, of half power width of the order of 10% to 20% of its center frequency. If data exclusively of this class is under consideration, the dominant period and envelope form in the correlogram may be used in some simple way to infer the position and shape of the spectral band. Band position and width might even be determined from the abscissa and ordinate of a single point of the correlogram, namely the first maximum or first minimum. This approach through the correlogram may be a useful solution to the problem emphasized by Seiwell [1], of obtaining in some objective way a smooth spectrum, free of the nonsignificant irregularities of detail which occur in any directly measured spectrum.

If, however, the spectrum has a more complex form, the execution of a Fourier transform rapidly loses simplicity. Examples are shown to indicate that a second maximum which is a fully obvious feature of the directly measured spectrum may have very inconspicuous influence in the correlogram. Under these circumstances, if the final interpretation of the wave record is to be made in terms of frequencies, it appears that the spectrum had best be determined directly, and that correlograms are likely to be useful only if the spectrum is found to be simple, or is rendered simple by filtration.

REFERENCES

[1] H. R. Seiwell and G. P. Wadsworth, "A new development in ocean wave research," *Science*, Vol. 109 (1949), pp. 271–274. H. R. Seiwell, "The principles of time series analyses applied to ocean wave data," *Proc. Nat. Acad. Sci.*, Vol. 35 (1949), pp. 518–528.

[2] Philip Rudnick, "A system for recording and analyzing random processes," *Symposium on Applications of Autocorrelation Analysis to Physical Problems, Woods Hole, Massachusetts,* 13–14 *June*, 1949, Office of Naval Research, Washington, D.C., 1950.

[3] N. F. Barber and F. Ursell, "The generation and propagation of ocean waves and swell," *Phil. Trans. Roy. Soc.*, Ser. A, Vol. 240 (1948), pp. 527–560.

[4] R. L. Miller, "Wave and weather correlation at Apra Harbor, Guam, M. I., from 18 March to 31 May, 1949," *Wave Report from Scripps Institution of Oceanography*, No. 90 (1949).

[5] M. S. Bartlett, "On the theoretical specification and sampling properties of autocorrelated time-series," *Jour. Roy. Stat. Soc., Suppl.*, Vol. 8 (1946), pp. 27–41.