CHAPTER FIFTY TWO

Prediction of Wave Group Statistics

Steve Elgar,[1] R.T. Guza,[2] and R.J. Seymour[3]

Two methods of numerically simulating random seas, given a target power spectrum, are discussed. Wave group statistics, such as the mean length of runs of high waves, produced by the different simulation schemes are compared. For a large number of spectral components, no significant differences are found in the wave group statistics produced by the two simulation techniques. Using the simulation techniques, it is shown that ocean gravity wave group statistics are not inconsistent with an underlying wave field composed of linearly superposed random waves. The majority of the field data examined were collected in 9-10 m depth, significant wave heights ranged from about 20 to 200 cm, and the spectral shapes ranged from fairly narrow to broad. For the 9-10 m depth data, observed mean run length, variance of run length, and the probabilities of runs of a given number of high waves were statistically consistent with the linear simulations. In contrast to the apparent linear behavior in 9-10 m depth, waves in 2-3 m depth showed substantial departures from the linear simulations.

Introduction

Groups of high waves are commonly observed in the ocean. A run or group of waves is defined as a sequence of waves, the heights of which exceed a particular level (Goda, 1970). It has been suggested that such runs of large waves effect coastal structures such as breakwaters (Burcarth, 1979) and pipelines (Dean, 1980), and influence the response of ships to the wave field (Pinkster and Huijsmans, 1982). Additionally, groups of waves can excite other fluid motions, which may in turn produce noticeable effects (Bowers, 1979).

There are several linear theories which predict wave group statistics, such as the mean group length, given only the energy spectrum. One such theory considers the wave field to be composed of a succession of discrete, independent waves, an assumption appropriate for broad band spectra. Results from the theory of runs are then employed to determine certain group statistics (Goda, 1970; Nagai, 1973). The mean length of such runs of waves greater than some critical height, H_c is (Goda, 1970)

$$E[j] = 1/(1-p) \tag{1}$$

where j is the number of discrete waves in a run, E[] is the expected value operator, and p is the probability that the height of a wave is greater than H_c. The standard deviation of run length is

$$\sigma[j] = p^{1/2}/(1-p) \tag{2}$$

For Rayleigh distributed wave heights, and for $H_c = 4m_o^{1/2}$, where m_o is the variance of the time series, $p = 0.1348$. Thus, for these conditions

$$E[j] = 1.16$$

$$\sigma[j] = 0.42$$

On the other hand, for narrow band energy spectra, an expression for the mean length of runs can be derived from Rice's (1944, 1945) results for the envelope of a random process. For

[1] Research Assistant, Scripps Institution of Oceanography, A-022, La Jolla, CA 92093

[2] Associate Professor, Scripps Institution of Oceanography, Shore Processes Lab, A-009, La Jolla, CA 92093

[3] Head, Ocean Engineering Research Group, Institute of Marine Resources, SIO, A-022, La Jolla, CA 92093, M. ASCE

this case, the mean length of runs whose envelope lies above $H_c/2$, for the H_c given above, is given by (Vanmarke, 1972; Ewing, 1973)

$$E[k] = 1/2(m_2/2\pi\mu_2)^{1/2} \tag{3}$$

where m_2 and μ_2 are the second moments of the spectrum about the origin and about the centroid, respectively, and k is the number of waves in the group, not necessarily discrete.

Extensive numerical simulations (Elgar, Guza, and Seymour, 1984) indicate agreement ($\pm$ 10 percent) with Goda's prediction of a constant mean run length, equation (1), for very broad spectra, $Q_p < 2$ Here Q_p is defined as (Goda, 1970)

$$Q_p = (2/m_0^2)\int_0^\infty f S^2(f)\,df \tag{4}$$

where f is frequency and S(f) is the power spectral density. Unfortunately, Elgar, et al. (1984) could find no simple shape parameter that quantitatively indicated the region of validity of equation (3). Indeed, for spectral shapes similar to those found in the ocean, neither of the theories described above adequately predicts mean run length. Consequently, in order to predict wave group statistics given an arbitrary spectral shape, a simulation procedure is utilized.

Simulations

The fundamental assumption of linear waves is that the sea surface can be represented as a linear combination of waves with random phases,

$$\eta(t) = \sum_{n=1}^{N} C_n \cos(\omega_n t - \phi_n) \tag{5a}$$

where $N >> 1$ and

$$C_n = (2S(f_n)\Delta f)^{1/2} \tag{5b}$$

are the Fourier amplitudes, $\omega_n = 2\pi f_n, f_n = n\Delta f$, and ϕ_n are random phase angles, uniformly distributed in $[0, 2\pi]$. An alternate expression for the sea surface is

$$\eta(t) = \sum_{n=1}^{N} a_n \cos(\omega_n t) + b_n \sin(\omega_n t) \tag{6}$$

where a_n and b_n are independent, Gaussian distributed random variables with zero mean and variance $S(f_n)\Delta f$. Simulations using equation (5), which will be referred to as a random phase scheme, have spectra that always exactly match the target spectrum, $S(f)$, while simulations with equation (6), a random Fourier coefficient scheme, have spectra with a statistical variation about $S(f)$. Both methods were implemented, and as discussed below, yield nearly identical results.

Comparison of simulation schemes

For the random phase scheme, Fourier coefficients were coupled with random phases produced by a numerical random number generator (equation (5)). An inverse fast Fourier transform of the unsmoothed spectrum results in a simulated time series with the identical spectral shape as the target spectrum, but with random phases. To obtain random Fourier coefficients (equation (6)), Gaussian distributed, zero-mean, unit-variance random variables were generated, and then multiplied by $(S(f_n)\Delta f)^{1/2}$, producing new Fourier amplitudes with the desired properties (Andrew and Borgman, 1982). Again, an inverse fast Fourier transform yields a simulated time series.

Rice (1944 and 1945), invoking the central limit theorem, points out that both representations (5) and (6) will yield the same statistics in the limit as $N \to \infty$. Nevertheless, both forms were used in the simulations because there is some question as to whether or not they produce the same results (Tuah and Hudspeth, 1982; Tucker, Challenor, and Carter, 1984).

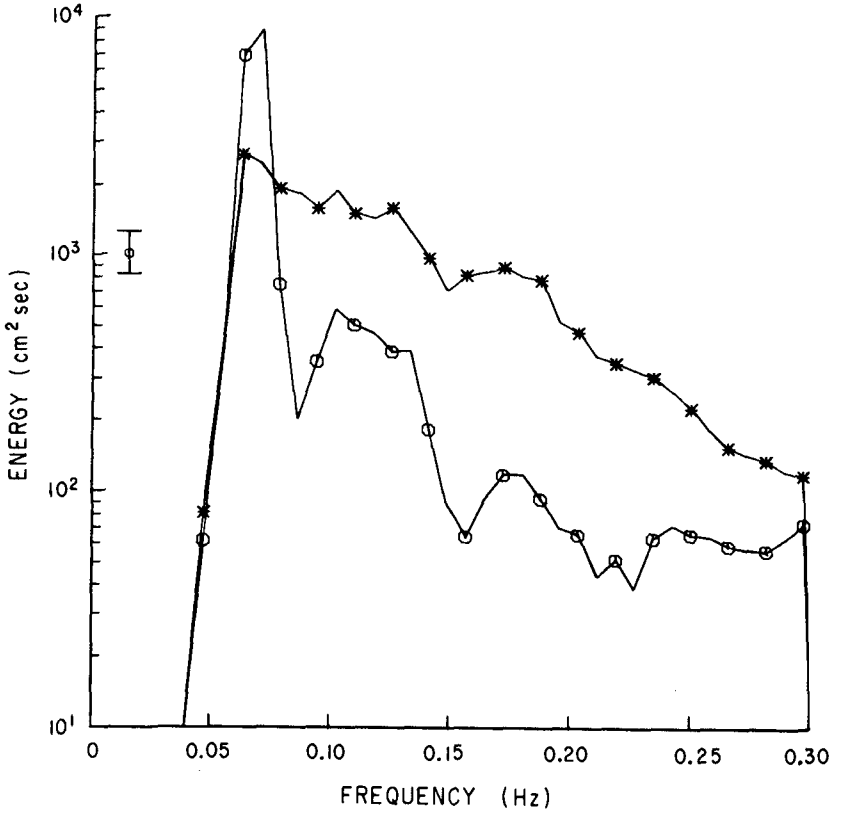

Figure 1. Power spectral density of sea surface elevation in water 10 m deep. Circle, narrow band; asterisk, broad band. The spectra have 128 degrees of freedom, and the 90 percent confidence limits are indicated by the bars. (Reprinted with permission of AGU.)

Each set of random phases or coefficients, via the simulations described above, produces a time series whose properties are statistics which fluctuate about mean values. To compare the two simulation methods for a particular target spectrum, 100 simulated time series were produced for each simulation scheme, each with its own set of random phases or Fourier coefficients. This procedure was repeated for 29 target spectra, thus a total of 5800 time series were produced.

The target spectra used to compare the two methods of simulating random waves were obtained from field measurements at Santa Barbara, California during the Nearshore Sediment Transport Study, conducted in January and February, 1980 (Gable, 1981). The time series used for this study were obtained from bottom mounted pressure sensors in water approximately 10 m deep. These spectra represent a wide range of ocean conditions, including very narrow (by ocean standards) and quite broad band spectral shapes. Figure 1, and significant wave heights between 20 and 200 cm. With typical peak periods from 8 to 20 seconds, the wave steepness (product of significant amplitude and wave number of the spectral peak) is in the range 0.006 to 0.1. Some of the field data were characterized by swell from distant storms, others by locally generated seas, and a few had multiple-peaked spectra, representing a combination of sea and swell.

Each time series used in this study was 8192 s (2.27 hr) long and was band-pass filtered between 0.04 and 0.3 Hz. The bottom pressures were converted to sea surface elevation using linear theory. Individual wave heights were determined using a zero-upcrossing definition, and were considered to belong to a group if the crest to trough distance exceeded $4m_o^{1/2}$, the significant wave height. The mean period of the waves was about 10 s, so there are approximately 800 waves per record, and about 80,000 waves per target spectrum for each of the two simulation schemes. The mean length of runs of waves greater than the significant wave height in the simulations varies from about 1 to 2.5, and the number of groups in each time series is between 30 and 100. It was shown in Elgar, et al. (1984) that simulations with 100 realizations as described above are extensive enough to estimate the mean length of runs and the frequency distribution of the number of waves per group to within a few percent of their true values.

For each realization, the mean run length and the frequency distribution of the number of waves per group were calculated. These quantities were then averaged over the 100 realizations per target spectrum. Higher order moments, such as the variance of run lengths, were calculated from the averaged frequency distributions. Average values were calculated for each simulation scheme, and compared to determine if there are any statistical differences between the two simulation schemes. Figure 2 shows that the mean run lengths from the random phase and random coefficient schemes are visually very similar. To test if the collection of mean run lengths from the random phase scheme was statistically consistent with the collection produced by the random coefficient scheme, Student's t test for paired data was calculated. Essentially, this test examines whether or not the two treatments (random phase and random coefficients) of the same data (target spectrum) produce the same result (mean run length). The t statistic obtained will be exceeded about 25 percent of the time due to random fluctuations. Thus, there is no support for the hypothesis that the mean group lengths produced by the random phase scheme are statistically different than those produced by the random coefficient simulations.

Similarly, the variances of run lengths obtained from the random phase and coefficient methods were compared. Figure 3 shows the two simulation procedures have negligibly different run length variances. The ratios of the square of run length coefficients of variation (standard deviation normalized by the mean, random phase and random coefficient schemes) for each of the 29 target spectra were compared to tabulated values of Fisher's F distribution. None of the values exceeded the tabulated values at the 99 percent significance level.

Finally, the frequency distributions of the number of waves per group produced by each simulation technique for each target spectrum were compared. A chi-square test was used to test if the entire collection of frequency distributions produced by the random phase scheme differed significantly from those produced by the random coefficient scheme. The chi-square value obtained (with 77 degrees of freedom) is such that the hypothesis that the two collections of frequency distributions come from the same population can be accepted with more than 99 percent confidence. Indeed, as displayed in Figure 4, when corresponding frequency distributions

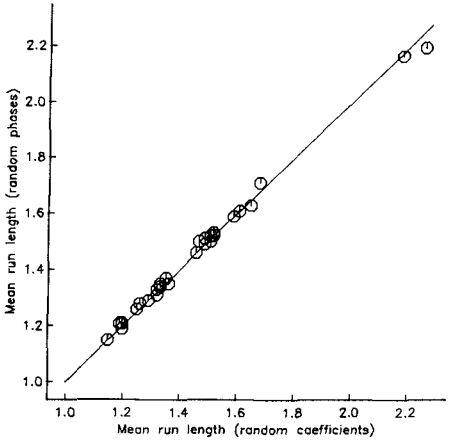

Figure 2. Mean length of runs greater than the significant wave height from the random phase scheme (equation (5)) versus mean length of runs greater than the significant wave height from the random coefficient scheme (equation (6)). The solid line indicates agreement between the two simulation methods.

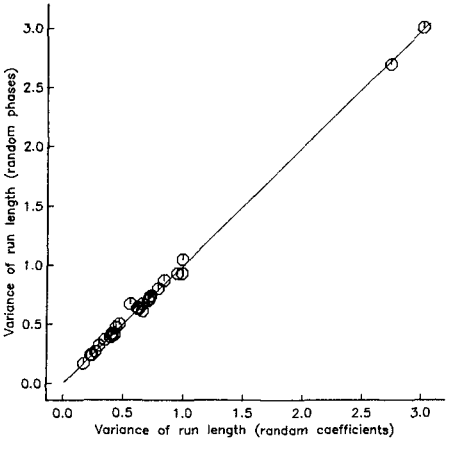

Figure 3. Variance of lengths of runs greater than the significant wave height from the random phase scheme (equation (5)) versus variance of the lengths of runs greater than the significant wave height from the random coefficient scheme (equation (6)). The solid line indicates agreement between the two simulation methods.

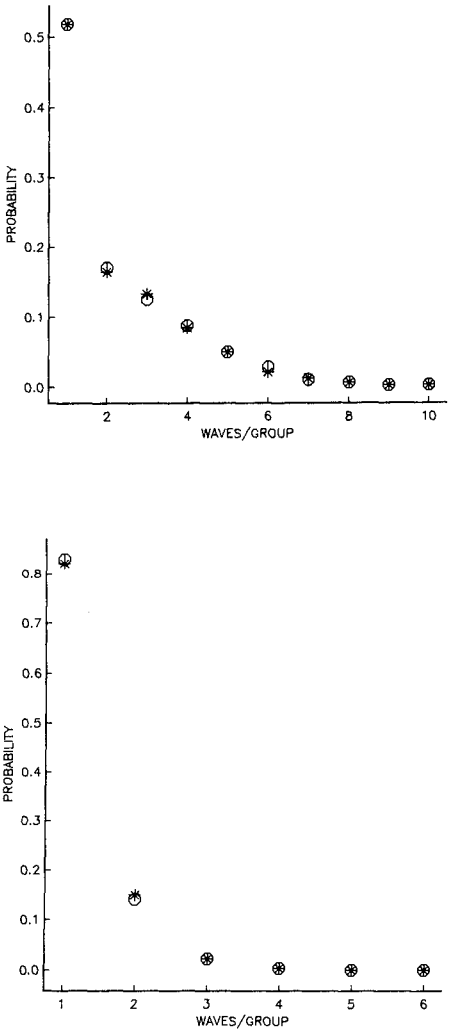

Figure 4. Frequently distribution of the number of waves per group corresponding to the spectra in Figure 1; circle, random coefficient scheme; asterisk, random phase scheme. (top) narrow band spectrum, (bottom) broad band spectrum.

from each simulation method are compared, they are seen to be almost identical. A more detailed discussion of the variability and statistics of the frequency distributions can be found in Elgar, et al. (1984).

The parameters investigated above indicate that the random phase scheme produces wave group statistics which do not differ from the random coefficient scheme statistics any more than two collections of random coefficient (or random phase) generated statistics would differ from each other. Further discussion of the similarities of and differences between the two simulation procedures can be found in Elgar, Guza, and Seymour (in Press).

Simulation-data comparison

In order to determine whether or not linear simulations (i.e., random phase or random Fourier coefficient schemes) are adequate for predicting ocean wave group statistics, several different comparisons of measured and simulated data were made. The simulation procedures are the same as described above, but in this case statistics from ensemble averages of the 100 simulation realizations are compared to the actual field data values for each target spectrum. One such comparison is of the mean length of runs. Figure 5 shows that the mean run lengths from the ocean data are visually well correlated with the mean run lengths from the simulations. As shown in Elgar, et al. (1984), the E[j] for the 100 realizations for each target spectrum are Gaussian distributed, and the field values of E[j] deviated from the simulation mean no more than would be expected for a Gaussian distribution, with 77 percent of the ocean E[j] falling within one standard deviation of the simulation mean.

To test if the entire collection of ocean E[j] were statistically consistent with the simulated E[j], Student's t test for paired data was calculated (Elgar, et al., 1984). The value obtained will be exceeded about 50 percent of the time due to random fluctuations. Hence, the hypothesis that the ocean mean group lengths are statistically consistent with linear wave theory cannot be rejected, and the simulation procedure successfully predicts ocean mean group lengths given a target spectrum.

Similarly, the variances of run length obtained from the ocean data were compared with the corresponding variances produced by the simulations, Figure 6. Fisher's F distribution was used to compare the ratios of the square of run length coefficients of variation. As shown in Elgar, et al. (1984), the hypothesis that the run length variances come from the same population cannot be rejected. Thus, the simulation procedure correctly predicts field values of run length variance given the energy spectrum.

Finally, a comparison of the frequency distributions of the number of waves per group was made, as shown in Figure 7. The details of this comparison can be found in Elgar, et al. (1984). Again, the conclusion is that the simulation schemes are capable of predicting field probability densities given only the power spectrum.

All the above statistical tests were applied to both the random phase simulations and the random Fourier coefficent simulations, with negligible differences as expected. The values presented above are from the random phase simulations.

Simulations of shoaling waves

The results presented so far indicate that the assumption of a linear, Gaussian process, as expressed by equation (5) or equation (6), produces statistics consistent with observations of ocean wave group statistics in 10 m depth. However, as waves shoal they are expected to become more nonlinear. Consequently, a linear representation, such as (5) or (6) should not necessarily produce wave group statistics consistent with observations of shoaled waves. That this is the case is shown in Figure 8. Values of E[j], both observed and simulated (from the measured spectrum at the appropriate depth) are shown as a function of depth, from 10 m, through the breaking region (about 2 m), and into 1 m of water. The simulated run length varies during shoaling because of substantial changes in the observed spectrum. Simulations and observations are similar in 9-10 m depth, but as the waves shoal the observed E[j] becomes much greater than

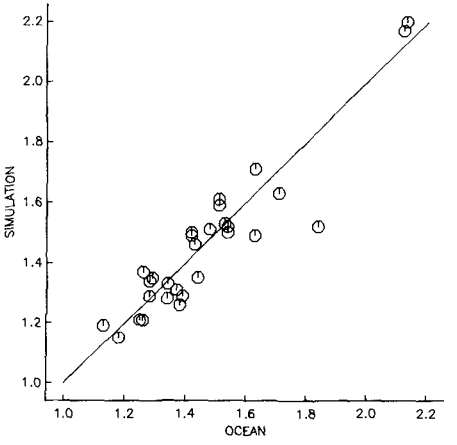

Figure 5. Mean length of runs greater than significant wave height from the numerical simulations versus mean length of runs greater than the significant wave height from the ocean field data The 45 ° solid line indicates agreement between simulations and ocean field data. (Reprinted with permission of AGU.)

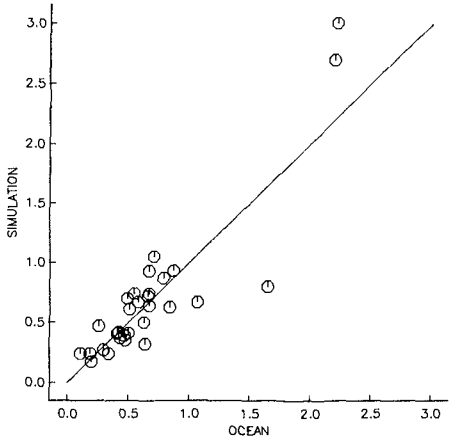

Figure 6. Variance of the lengths of runs greater than the significant wave height from the numerical simulations versus variance of the lengths of runs greater than the significant wave height from the ocean field data. The 45 ° solid line indicates agreement between simulations and field data.

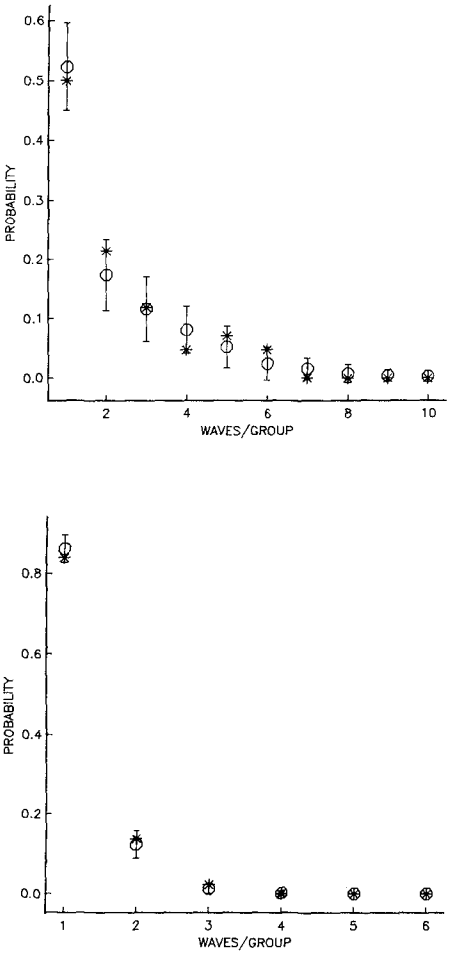

Figure 7. Frequency distribution of the number of waves per group corresponding to the spectra in Figure 1; circle, simulations; asterisk, ocean field data. (top) narrow band, February 2 (42 groups were observed in the field data); (bottom) broad band, February 15 (74 groups were observed in the field data). Bars indicate ± 1 standard deviation of simulated values. (Reprinted with permission of AGU.)

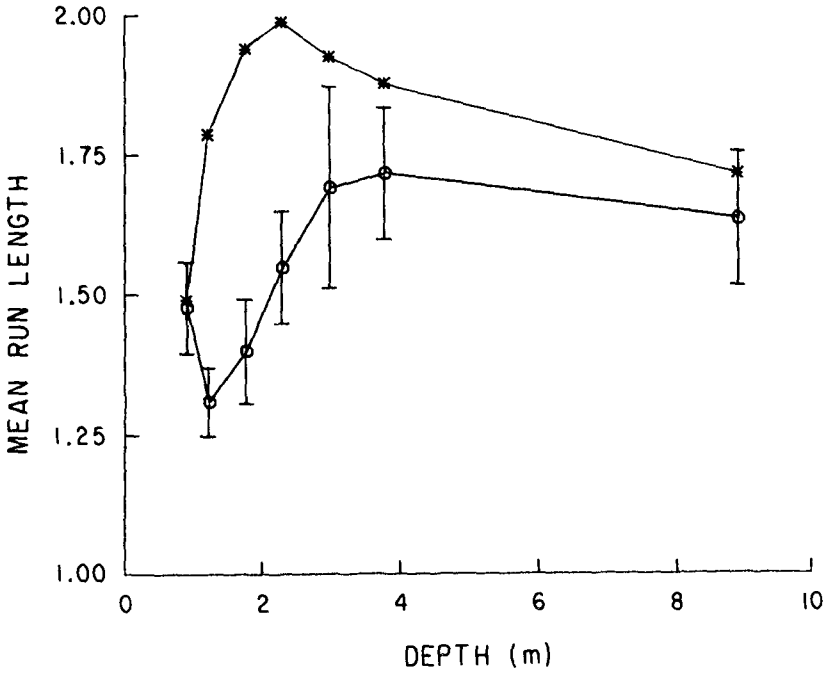

Figure 8. Mean length of runs greater than the significant wave height versus depth of water. Asterisk, ocean field data; circle, simulations, bars indicate ± 1 standard deviation of simulated values. (Reprinted with permission of AGU.)

linear simulations predict. The ocean E[j] remains higher than the corresponding value from the simulations until the waves break. This trend occurs in many of the data sets, and requires a nonlinear model to be predicted from a given deep water spectrum (Elgar, et al., in preparation). Other group statistics observed in the field data are also inconsistent with the linear simulations, as described in Elgar, et al. (1984). Consequently, it is inappropriate to use the linear simulation technique to predict wave group statistics for very shallow water waves.

Conclusions

The theoretical models which predict wave groups statistics given only the power spectral density are not appropriate for the vast majority of ocean spectra. On the other hand, wave group statistics such as the mean length of runs of high waves, the run length variance, and the frequency distribution of the number of waves per group can be predicted by a linear numerical simulation from the energy spectrum. The two methods discussed here, a random phase scheme and a random Fourier coefficient scheme, produce nearly identical statistics for the spectra and conditions considered in this study.

Although the linear simulations accurately predict wave group statistics in water 8-10 m deep, very substantial disagreement with the simulations was found for shoaled waves in 2-3 m depth. Thus, the linear simulation procedure is inappropriate to use in shallow water where nonlinearities are important.

Acknowledgements

This study was funded by the National Oceanic and Atmospheric Administration Office of Sea Grant contracts NOAA-04-8-MO1-193, R-CZ-N-4E (S.L.E. and R.J.S.) and Office of Naval Research contract N00014-75-C-0300 (R.T.G.) S.L.E. would like to thank Wayne Hager, chairman of the Engineering Science Department, University of Idaho, where the bulk of this manuscript was prepared, for providing a most hospitable environment during the summer of 1984.

References

Andrew, M.E., and L.E. Borgman, Procedures for studying wave grouping in wave records from California Coastal Data Collection Program, report, U.S. Army Corps of Eng., San Francisco, Calif., November, 1981.

Bowers, E.C., Harbour resonance due to set-down beneath wave groups, JFM *79*, 71-92, 1977.

Burcharth, H.F., The effect of wave grouping on on-shore structures, Coastal Engineering *2*, 189-199, 1979.

Dean, R.G., Kinematics and forces due to wave groups and associated second-order currents, Proceedings Second International Conference on the Behavior of Offshore Structures (BOSS'80), London, paper # 6, 1980.

Elgar, Steve, R.T. Guza, and R.J. Seymour, Groups of waves in shallow water, Journal of Geophysical Research, *89*, 3623-3624, 1984.

Elgar, Steve, R.T. Guza, and R.J. Seymour, Wave group statistics from numerical simulations of a random sea, Applied Ocean Research, in Press.

Goda, Y., Numerical experiments on wave statistics with spectral simulation, report, volume 8, pp.3-57, Port and Harbor Research Institute, Nagase, Japan, 1970.

Nagai, K., Runs of the maxima of the irregular sea, Coastal Engineering Japan, *16*, 13-18, 1973.

Pinkster, J.A. and R.H.M. Huijmans, The low frequency motions of a semi-submersible in waves, Proceedings of the Third International Conference on Behavior of Off Shore Structures (BOSS'82), Boston, 1982.

Rice, S.O., The mathematical analysis of random noise, Bell System Technical Journal, *23*, 282-332, 1944.

Rice, S.O., The mathematical analysis of random noise, Bell System Technical Journal, *24*, 46-156, 1945.

Tuah, H. and R.T. Hudspeth, Comparisons of numerical random sea simulations, Journal of Waterway, Port, Coastal, and Ocean Division, American Society of Civil Engineers, *108*, 569-584, 1982.

Tucker, M.J., P.J. Challenor, and D.J.T. Carter, Numerical simulation of a random sea: a common error and its effect upon wave group statistics, Applied Ocean Research, *6* , 118-122, 1984.