Wave group statistics from numerical simulations of a random sea

S. ELGAR, R. T. GUZA and R. J. SEYMOUR

Scripps Institution of Oceanography, Mail Code A-022, University of California, La Jolla, California 92093, USA

Two commonly used methods of simulating random time series, 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. The target spectra used are obtained from ocean measurements, and cover a wide range of ocean conditions. For a sufficiently large number of spectral components, no significant differences are found in the wave group statistics produced by the two simulation techniques.

Key Words: Wave simulation, wave groups, wave statistics.

INTRODUCTION

Because the dynamics of surface gravity waves are quite complicated, it has become common practice to simulate random seas in order to gain information about various statistics that cannot be obtained analytically. These simulations are sometimes produced in the laboratory, with a programmable wave paddle for example. More often, however, random seas are simulated numerically on a digital computer. A recent paper by Tucker, Challenor and Carter[1] discusses various methods of digitally simulating random time series. Two of the most common methods will be discussed here. The first, called a random phase scheme, represents the time series as:

$$\zeta(t) = \sum_{n=1}^{N} C_n \cos(2\pi f_n t + \phi_n) \qquad (1)$$

where

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

are the Fourier amplitudes, $S(f)$ is the energy density spectrum, Δf is the frequency resolution, $f_n = n\,\Delta f$, and ϕ_n are random phase angles, uniformly distributed in $[0, 2\pi]$. The second method, called a random coefficient scheme, is:

$$\zeta(t) = \sum_{n=1}^{N} a_n \cos(2\pi f_n t) + b_n \sin(2\pi f_n t) \qquad (2)$$

where a_n, b_n are independent, Gaussian distributed random variables with zero mean and variance $S(f_n)\,\Delta f$. Tucker *et al.*[1] correctly point out that the proper representation of a Gaussian sea is equation (2), and only in the limit $N \to \infty$ does equation (1) truly represent a Gaussian sea. Tucker *et al.*[1] correctly state that it is not clear which wave statistics produced by the random phase scheme are in error. On the other hand, they claim that 'statistics of wave groups are certainly affected'. Furthermore, they conclude that the random phase scheme 'incorrectly reproduces the distribution of lengths of wave groups'.

During a recent study to ascertain whether or not certain wave group statistics observed in ocean field data were consistent with linear dynamics, time series were numerically simulated using both the random phase and the random coefficient schemes.[2] That paper very briefly mentions that no substantial differences were found, with respect to wave group statistics, between the two simulation schemes. However, in light of Tucker *et al.*[1] and the comments regarding wave group statistics therein, this question is now examined in more detail. The simulations discussed immediately below are those of Elgar *et al.*,[2] while the effects of varying N, the number of spectra components (equation (1) and equation (2)), are considered in the discussion section.

SIMULATIONS

For the random phase scheme, equation (1), the Fourier coefficients (C_n) from a target spectrum were coupled with random phases produced by a numerical random number generator. An inverse Fast Fourier transform of the unsmoothed spectrum produces a time series. To obtain random Fourier coefficients (a_n, b_n of equation (2)), Gaussian distributed, zero-mean variables with unit variance were generated, and then multiplied by $(S(f_n)\,\Delta f)$ producing new Fourier amplitudes with the desired properties.[3] Again, an inverse Fast Fourier transform yields a simulated time series.

For both the random phase and the random coefficient schemes, 100 simulated time series (i.e. realizations) were produced, 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. These spectra represent a wide range of conditions, including narrow (by ocean standards) and quite broadband spectral shapes. Examples of smoothed spectra are shown in Fig. 1. Some of these field data

Accepted September 1984. Discussion closes June 1985.

0141-1187/85/020093-04 $2.00

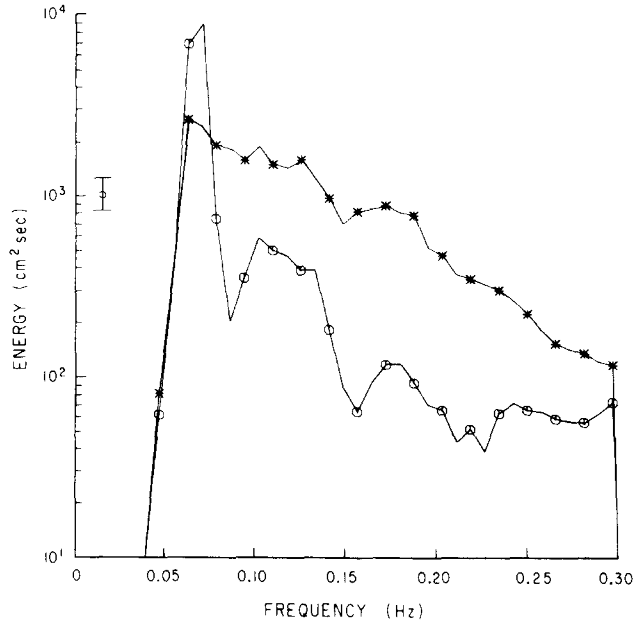

Figure 1. Power spectral density of sea surface elevation in water 10 m deep. Circle, narrow band; asterisk, broad-band. The spectra have 128 degrees of freedom, and the 90% confidence limits are indicated by the bars

were characterized by swell from distant storms, others by locally generated seas, and a few had multiple-peaked spectra, representing different combinations of sea and swell.

Individual wave heights were determined by using a zero-upcrossing definition, and were considered to belong to a group of high waves if the crest to trough distance exceeded four standard deviations of the time series (the significant wave height). Each time series was 8192 s long and was band-passed filtered between 0.04 and 0.3 Hz. Thus, the frequency resolution of the target spectra is 1.22×10^{-4} Hz, and between 0.04 and 0.3 Hz there are 2130 random phases (or 4260 random Fourier coefficients). The mean period was about 10 s so there are about 800 waves per simulated time series, and about 80 000 waves per target spectrum for each of the two simulation schemes. The mean length of runs greater than the significant wave height in the simulations varies from about 1 to almost 2.5, and the number of groups in each time series is between 30 and 100.

It was shown by Elgar *et al.*[2] that simulations with 100 realizations as described above are extensive enough to estimate the mean length of runs and the frequency distributions of the number of waves per group within a few per cent of their true values.

RESULTS

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. Other group statistics, such as the variance of run lengths, were calculated from the averaged frequency distributions. Values from each simulation scheme were compared to determine if there were any statistically significant differences. Figure 2 shows that the mean run lengths from the random phase and random coefficient schemes are visually well correlated. To test if the collection of mean run lengths from the random phase scheme were statistically consistent with those 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 was $t = 1.3$, which will be exceeded about 25% 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% significance level.

Finally, the estimated probability function 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 estimated probability functions 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 estimated probability functions come from the same population can be accepted with more than 99% confidence. Indeed, when corresponding probability functions from each simulation method are compared, they are seen to be almost identical (Fig. 4).

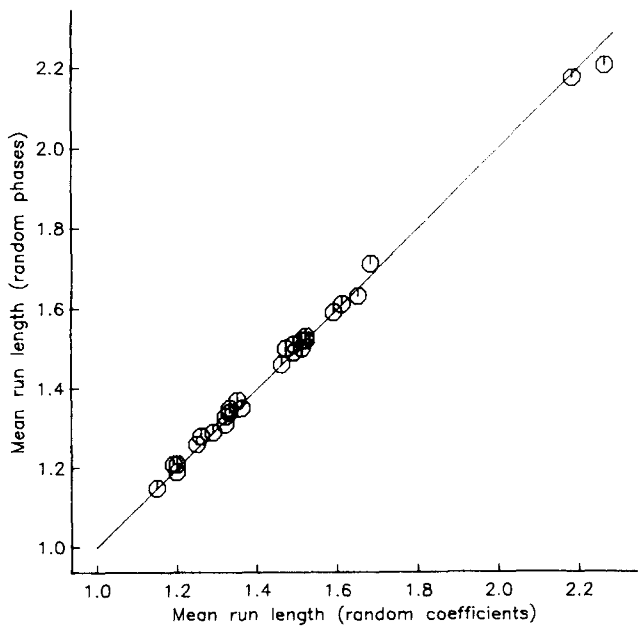

Figure 2. Mean length of runs greater than the significant wave height from the random phase scheme (equation (1)) versus mean length of runs greater than the significant wave height from the random coefficient scheme (equation (2)). The solid line indicates agreement between the two simulation methods

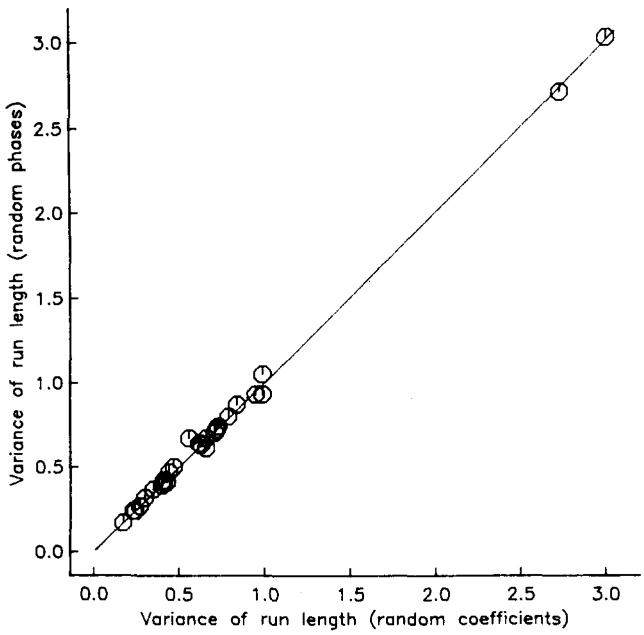

Figure 3. Variance of the lengths of runs greater than the significant wave height from the random phase scheme (equation (1)) versus variance of the lengths of runs greater than the significant wave height from the random coefficient scheme (equation (2)). The solid line indicates agreement between the two simulation methods

A more detailed discussion of the variability and statistics of these probability functions can be found in Elgar *et al.*[2]

The parameters investigated above indicate that the random phase scheme can produce wave group statistics which do not differ from the random coefficient scheme statistics any more than two collections of random coefficient generated statistics would differ from each other.

DISCUSSION

Based on the results discussed above, the hypothesis that the low order, simple wave group statistics produced by the random phase scheme are necessarily different than those produced by the random coefficient scheme must be rejected. Indeed, for the group statistics considered here the differences between the two simulation procedures are essentially negligible. This may not be surprising considering the large number of random phases (2130) and random coefficients (4260) used here. The proof that the two simulation schemes produce identical statistics as the number of spectral amplitudes approaches infinity is based on the Central Limit Theorem.[4] In many applications of statistics, the Central Limit Theorem is invoked when the number of degrees of freedom is greater than 30 or so. Although simulations of a random sea based upon only 30 frequencies may not be adequate for most studies, 2130 frequencies is apparently large enough to invoke the Central Limit Theorem, at least for the particular spectra and wave group statistics considered here.

In general, the frequency resolution required for the two simulation schemes to produce similar statistics is spectral shape dependent. To produce a Gaussian sea from a very narrow spectrum using the random phase scheme will require densely spaced (in frequency) coefficients near the spectral peak. On the other hand, simulating a broad spectrum will require spectral components spread over the entire energetic part of the spectrum. The effect of spectral shape and the number of spectral components on the simulation of the mean length of runs produced by the two simulation schemes is qualitatively illustrated in Fig. 5. The 8192 s records corresponding to the spectra shown in Fig. 1 were subdivided into ensembles of several shorter records, each with a decreased N. Both of the simulation procedures were applied to each short record, producing 100 simulated time series from each simulation scheme for each

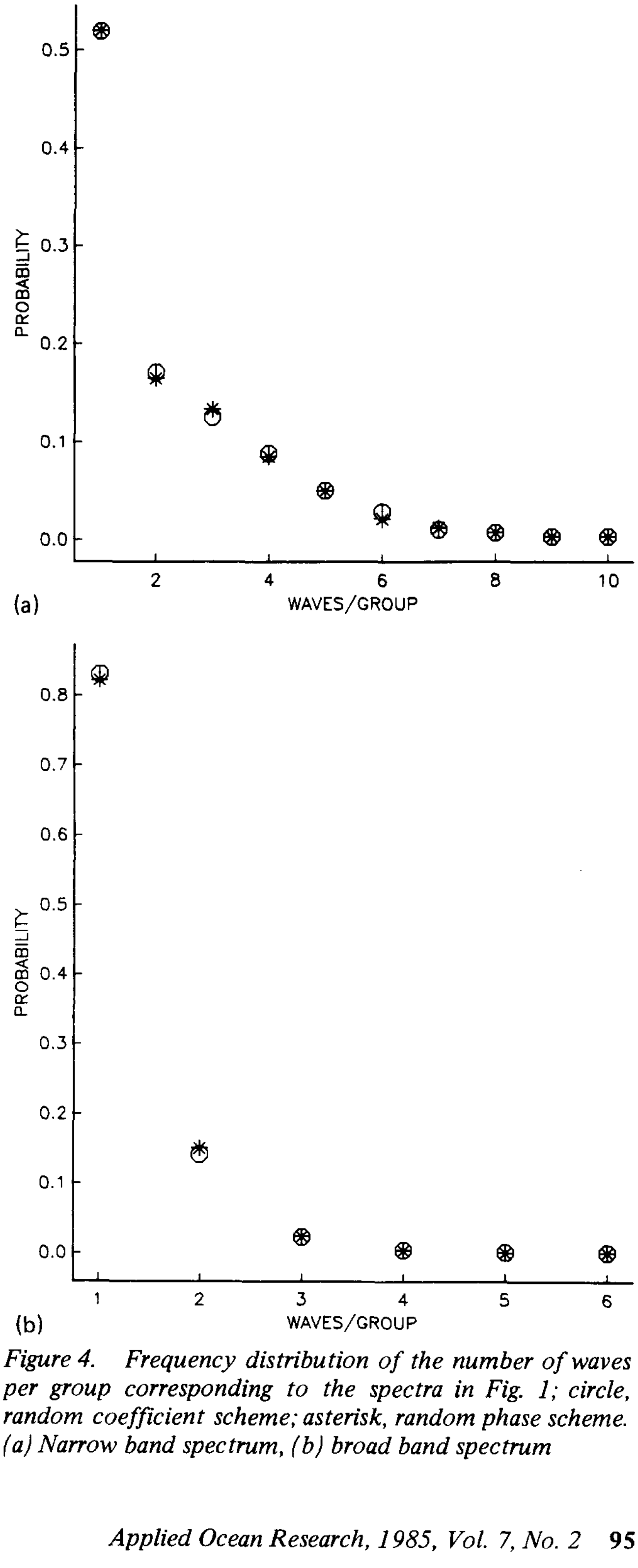

Figure 4. Frequency distribution of the number of waves per group corresponding to the spectra in Fig. 1; circle, random coefficient scheme; asterisk, random phase scheme. (a) Narrow band spectrum, (b) broad band spectrum

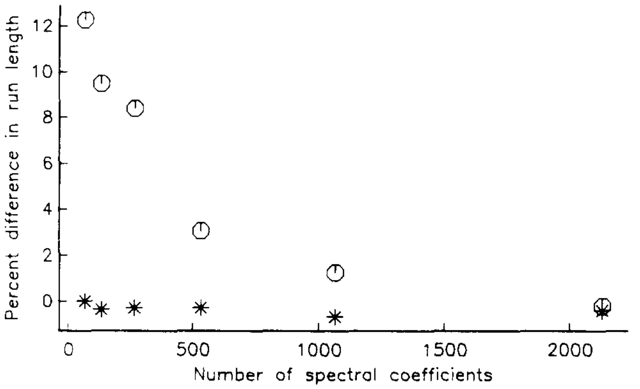

Figure 5. Percentage difference in the lengths of runs greater than the significant wave height produced by the two simulation schemes, (random phase run length − random coefficient run length)/random phase run length × 100%, versus number of spectral components. Circle, narrow band spectrum; asterisk, broad band spectrum

of the short records. The simulated mean run length was calculated by averaging the mean run lengths produced from all the short records in the ensemble. Thus, the number of spectral coefficients in a single simulated spectrum is reduced, but the total number of degrees of freedom remains constant. For the narrow band energy spectrum, the difference in the mean length of runs produced by the two simulation schemes increases as the number of spectral components decreases, although never more than about 12%. On the other hand, mean run lengths from the two simulation schemes are almost identical for the broad band spectrum, even for $N = 66$, the smallest number of spectral components considered. The differences between the broad and narrow spectra (Fig. 5) suggest that it is more appropriate to consider the effective number of spectral coefficients to be those within the energetic part of the spectrum, and not the total number of coefficients. For example, about 80% of the energy of the narrow band spectrum is contained within a frequency band only 0.06 Hz wide, while in the broad spectrum the same relative amount of energy is distributed in a band 0.13 Hz wide. Thus, the effective number of spectral coefficients for the broad band spectrum may be larger than that for the narrow band spectrum.

CONCLUSIONS

Two common methods of simulating random time series have been investigated. The random Fourier coefficient scheme (equation (2)) reproduces a Gaussian sea, while the random phase scheme (equation (1)) theoretically results in a Gaussian sea only in the limit of infinitely many spectral components. However, for many wave group statistics there is no significant difference between the two simulation schemes when a sufficiently large number of spectral components is used. For the particular spectral shapes used for the simulations in this study, which are representative of a broad range of ocean conditions, the wave group statistics produced by the two simulation procedures are essentially identical for 1000 (or more) Fourier components per spectrum.

Clearly, certain spectral statistics are not properly modelled by the random phase simulations, the statistical fluctuations of spectral levels being an obvious example. Hence, the random phase scheme is not suitable for all applications. On the other hand, the implication (Tucker *et al.*[1]) that use of the random phase method has necessarily corrupted the numerical simulations of Rye and Levrik[5] and others is contradicted by the results presented in this study.

ACKNOWLEDGEMENTS

This study was funded by the National Oceanic and Atmospheric Administration Office of Sea Grant Contracts NOAA-04-8-M01-193, R-CZ-N-4E (S.L.E. and R.J.S.) and Office of Naval Research Contract N00014-75-C-300 (R.T.G.). J. A. Battjes made helpful suggestions.

REFERENCES

1 Tucker, M. J., Challenor, P. J. and Carter, D. J. T. Numerical simulation of a random sea: a common error and its effect upon wave group statistics, *Applied Ocean Research* 1984, **6**, 93
2 Elgar, S., Guza, R. T. and Seymour, R. J. Groups of waves in shallow water, *J. Geophys. Res.* 1984, **89**, 3623
3 Andrew, M. E. and Borgman, L. E. Procedures for studying wave grouping in wave records from California Coastal Data Collection Program, report, US Army Corps of Engineers, San Francisco, California, November 1981
4 Rice, S. O. The mathematical analysis of random noise, *Bell System Technical Journal* 1944, **23**, 282; 1945, **24**, 46
5 Rye, H. and Lervik, E. Wave grouping studied by means of correlation techniques, Preprint *International Symposium on Hydrodynamics in Ocean Engineering, Trondheim, Norway*, August 1981, **1**, 25;