JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 89, NO. C6, PAGES 10,547–10,552, NOVEMBER 20, 1984

The Distribution of Measured and Simulated Wave Heights as a Function of Spectral Shape

GEORGE Z. FORRISTALL

Shell Development Company

First-order theory predicts that zero-crossing wave heights should have a Rayleigh distribution, but measured waves are slightly smaller than the theoretical prediction. Suggested explanations for this discrepancy have included the nonlinearity of the wave surface, limitations on height as a result of breaking, and the effect of spectral width. In a recent paper, Tayfun (1981*b*) showed that the shape of the spectrum influences the height distribution because the wave envelope has different amplitudes at the time of the crest and trough. We have compared the distribution developed by Tayfun to simulated waves with different spectral shapes as well as to observations and found excellent agreement. The theoretical, simulated, and measured wave height distributions agree to within 1% in height. This result reinforces the conclusion that linear Gaussian simulations can reproduce many features of ocean waves.

INTRODUCTION

The envelope of band-limited, normally distributed noise has a Rayleigh distribution. If the band of frequencies present in the noise is rather narrow, it is natural to associate crest-to-trough wave heights with twice the amplitude of the envelope, so that zero-crossing wave heights also have a Rayleigh distribution. Since this argument was applied to ocean waves by *Longuet-Higgins* [1952], the distribution has proved to be very useful in engineering studies of wave loading. Nevertheless, *Thompson* [1974], *Haring et al.* [1976], and *Forristall* [1978] all found that the Rayleigh distribution overpredicted the heights of the highest waves in long field recordings. Figure 1 shows the data from Forristall along with the Rayleigh distribution and a Weibull distribution that was empirically fit to the data.

There have been numerous suggestions for the cause of the overprediction. In *Forristall* [1978] I speculated that the likely source of the discrepancy was the failure of the assumption of linearity of the wave profile and that the theoretical understanding of the distribution would not improve in the near future. It now appears that I was wrong on both counts. Both *Tayfun* [1980] and *Longuet-Higgins* [1980] have pointed out that second-order locked harmonics do not change the zero-crossing wave height. *Longuet-Higgins* [1980] also showed that higher-order nonlinearities would tend to increase the heights for a given wave spectrum.

Tayfun [1981*a*] suggested that the overprediction could be due to wave heights being steepness limited by the physical process of breaking. However, the data analysis by *Chen et al.* [1979] showed that measured wave heights and periods were rather far away from the breaking criteria, even in hurricane conditions.

Several discussions of the importance of spectral width have appeared. The distribution for crest heights as a function of spectral width that was calculated by *Rice* [1945] and discussed by *Cartwright and Longuet-Higgins* [1956] does not apply to wave heights. The Rice distribution is for local maxima, including maxima that may be negative, and not for crest-to-trough heights. *Nolte and Hsu* [1979] pointed out that waves with frequencies that are even multiples of the frequencies of the primary components of the spectrum will not contribute to the wave height. This is true regardless of whether these components are free or nonlinearly locked. Nolte and Hsu thus proposed filtering out the portion of the spectrum that did not contribute to height before calculating the Rayleigh distribution. This correction has the effect of shifting the solid curve in Figure 1 to the left by the same ratio at all levels of probability. A test calculation gave good agreement for the probabilities of the highest waves but underestimated the heights of moderate waves somewhat.

Longuet-Higgins [1980] pointed out that he originally normalized the heights by the rms wave amplitude ($\bar{a}$) rather than by the square root of the spectral variance ($m_0^{1/2}$), as was done by *Forristall* [1978]. He then calculated the ratio $\bar{a}/m_0^{1/2}$ as a function of spectral bandwidth. This procedure again shifts the solid curve in Figure 1 proportionally to the left, although not as far as the filtering scheme of *Nolte and Hsu* [1979].

Some support for the influence of spectral width on the wave height distribution was given by *Larsen* [1981], who analyzed records from a deep pressure transducer. All of the records had rather narrow spectral widths, but they did show a trend in the same direction as predicted by *Longuet-Higgins* [1980].

Finally, *Tayfun* [1981*b*] studied the consequences of the fact that the crest and trough of the wave do not occur at the same time. If the spectrum is not narrow, the envelope will change during that half-wave period. Furthermore, if the wave is high, so that the crest is near an extreme on the envelope, it is likely that the associated trough will have a smaller amplitude, and the wave height will be less than twice the value of the envelope at the crest.

Rice [1945] derived the joint distribution for the amplitudes of two points on the envelope separated by time τ, and *Tayfun* [1981*b*] integrated it to give a distribution for zero-crossing wave heights. This approach seems to give a good represenation of the effect of spectral width on the wave height distribution, but its predictions need to be tested.

Simulated wave series are very helpful in studying the form of the height distribution, since long records can be produced with specified spectral forms. Furthermore, the simulations can be made to conform to the assumptions in the theory. Gaussian simulations contain none of the contaminating effects of nonlinearities or wave breaking that might confuse the issue in measurements of real waves.

We begin by reviewing the derivation presented by *Tayfun*

Paper number 4C0961.
0148-0227/84/004C-0961$05.00

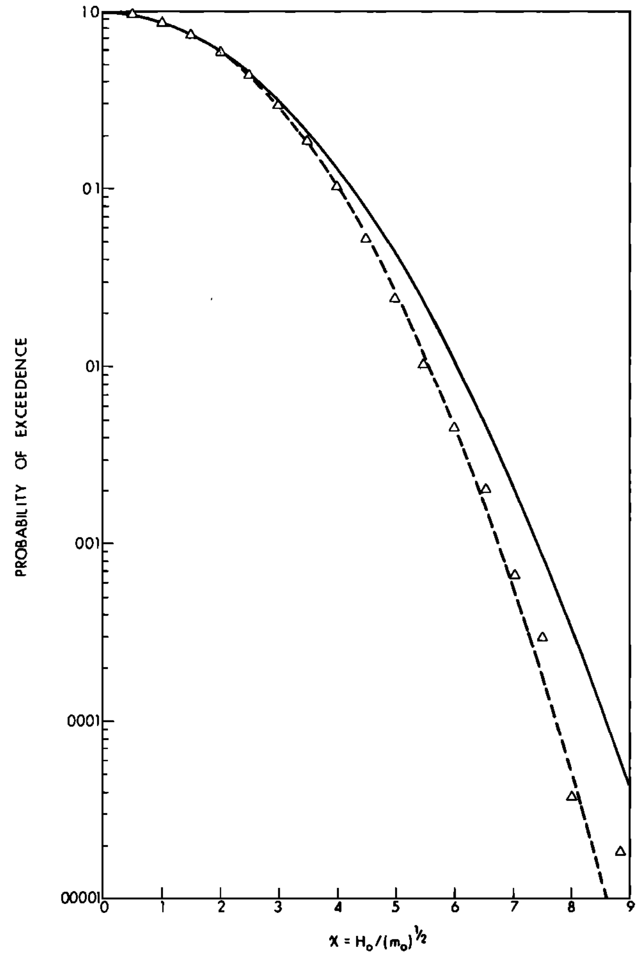

Fig. 1. Probability of exceeding a given normalized wave height: (solid line) Rayleigh distribution; (dashed line) empirical Weibull distribution fit to data; (triangles) measurement.

[1981*b*]. His distribution is then compared to that found from analysis of simulated wave records with various spectral moments. The agreement is very good. Finally, we calculate the spectral moments of the measurements used by *Forristall* [1978] and show that Tayfun's distribution also fits the distribution of those wave heights well. It thus is reasonable to conclude that the spectral shape controls the wave height distribution and that nonlinearities, wave breaking, and other effects not included in the simulations have negligible effect.

TAYFUN'S DISTRIBUTION

The variation with time of the sea surface at a point can be described as a Fourier series:

$$\eta(t) = \sum a_i \cos(\omega_i t + \phi_i) \tag{1}$$

where the a_i are amplitudes, ω_i are frequencies, and ϕ_i are phases. For a linear model the phases are uniformly distributed, and $\eta(t)$ has a Gaussian distribution.

The wave record can also be described by its power spectrum $S(\omega)$, which has moments defined by

$$m_r = \int_0^\infty \omega^r S(\omega)\, d\omega \tag{2}$$

By Parseval's theorem, m_0 is the variance of the wave record.

A mean frequency can be defined as

$$\omega_0 = m_1/m_0 \tag{3}$$

and the spectral width is given by

$$\nu^2 = m_2/m_0\omega_0^2 - 1 \tag{4}$$

It is always possible to find $A(t)$ and $\theta(t)$ such that (1) can be rewritten as

$$\eta(t) = A(t) \cos(\omega_0 t + \theta(t)) \tag{5}$$

The function $A(t)$ can then be shown to have a Rayleigh distribution [*Rice*, 1945, Equation (3.7-10)]. When the spectrum has a relatively narrow width, $\eta(t)$ has the appearance of a carrier wave at frequency ω_0 modulated by the envelope function $A(t)$. The peaks of $\eta(t)$, then, have the same Rayleigh distribution as the envelope.

If the envelope varies slowly, the distance between a crest and a preceding trough will be approximately twice the height of the crest, and the wave height distribution will also be Rayleigh. However, there will be cases where the spectrum is sufficiently narrow for an amplitude-modulated carrier wave to be a good description of the process but where the variation in amplitude between crest and trough is important.

Rice [1945, equation (3.7-10)] also derived the joint distribution for the value of A at time t and its value at some later time $t + \tau$. If we define the normalized variables

$$x = A(t)/m_0^{1/2}$$

and

$$y = A(t+\tau)/m_0^{1/2} \tag{6}$$

then their joint probability density function is

$$g(x, y; \tau) = \frac{xy}{1 - r^2} I_0\left\{\frac{xyr}{1 - r^2}\right\} \exp\left\{-\frac{x^2 + y^2}{2 - 2r^2}\right\} \tag{7}$$

where

$$r^2 = \rho^2 + \lambda^2 \tag{8}$$

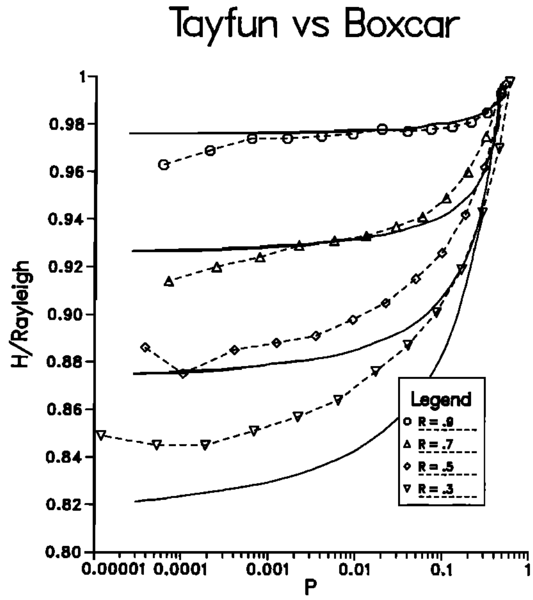

Fig. 2. Wave height distribution from a boxcar spectrum: (solid lines) *Tayfun*'s [1981*b*] distribution for various values of r; (dashed lines with points) results from numerically simulated waves.

$$\rho = \frac{1}{m_0} \int_0^\infty S(\omega) \cos(\omega - \omega_0)\tau \, d\omega \quad (9)$$

$$\lambda = \frac{1}{m_0} \int_0^\infty S(\omega) \sin(\omega - \omega_0)\tau \, d\omega \quad (10)$$

and I_0 is a modified Bessel function. The mean time between crests and troughs can be estimated from the mean frequency to be π/ω_0. Using this constant value of $\tau = \pi/\omega_0$ as an approximation, the density function of the normalized wave heights defined by

$$z = x + y \quad (11)$$

can then be found from evaluation of the convolution integral

$$f(z) = \int_0^z g(z - u, u; \pi/\omega_0) \, du \quad (12)$$

Equation (12) can be evaluated numerically without difficulty. Spectra of different shapes will have different values of the parameter r as evaluated by using (8)–(10). *Tayfun* [1981*b*] found an approximation for r in terms of the spectral width defined in (4). However, this approximation is not very good for moderately large values of the spectral width, and the evaluation of the trigonometric moment in (9) and (10) is no more difficult than the evaluation of the moments in (2). The form of (9) and (10) indicates that spectral components far from the peak will have little influence on the wave height distribution. On the other hand the spectral width is rather heavily influenced by high-frequency components, since it depends on the second moment.

The probability that that normalized wave height will exceed a given value z is given by

$$1 - F(z) = 1 - \int_0^z f(u) \, du$$

$$1 - F(z) = \int_z^\infty f(u) \, du \quad (13)$$

$$1 - F(z) = \int_0^{1/z} f(1/v) \, dv/v^2$$

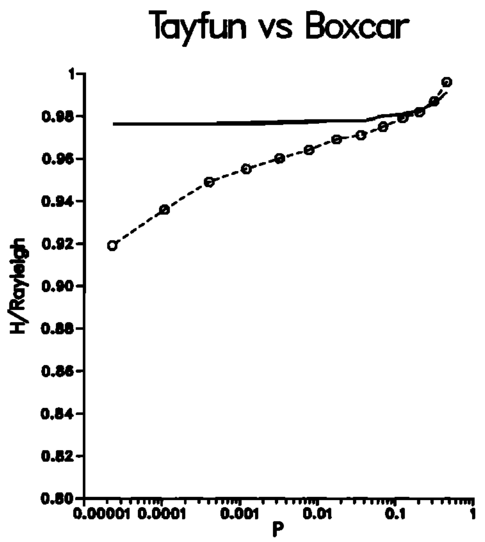

Fig. 3. Wave height distribution for a narrow spectrum and short simulation lengths: (solid line) Tayfun distribution for $r = 0.9$; (dashed line with points) simulation with a transform length of 2048.

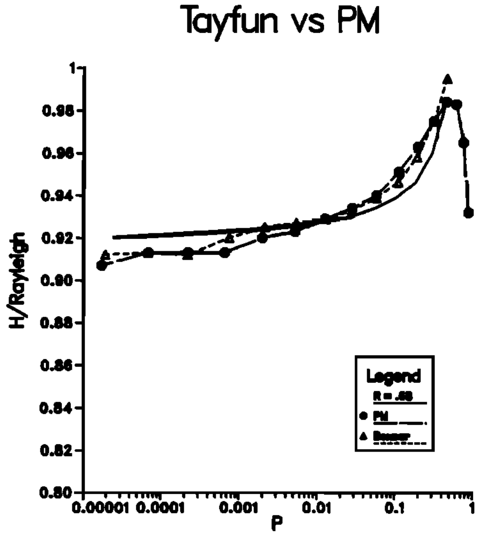

Fig. 4. Wave height distributions for a Pierson-Moskowitz spectrum: (solid line) Tayfun distribution with $r = 0.68$; (broken line) simulation of Pierson-Moskowitz spectrum; (dashed line) simulation of boxcar spectrum with $r = 0.68$.

When z is large, the numerical evaluation of (13) is more accurate if the last form is used.

SIMULATED WAVE HEIGHTS

The derivation of (12) included two approximations: first, the assumption that the wave height distribution can be calculated from its envelope even when the spectrum is not very narrow and second, the approximation of the wave period distribution by a delta function at $2\pi/\omega_0$. The effect of these approximations can be isolated from other factors that might influence the distribution of measured wave heights by analyzing simulated wave series.

Waves that conform to the model given by (1) can be economically simulated by using a fast Fourier transform to sum the series [*Borgman*, 1969]. The amplitudes of the wavelets are chosen to fit a given spectrum, and the phases are chosen from a uniformly distributed random variable. *Osborne* [1982] has used a similar Monte Carlo simulation to study wave statistics. A Fourier transform with frequency resolution $\Delta\omega = 2\pi/N\Delta t$ includes all the information in a time series of length N and sample spacing Δt. If the phases in the frequency domain are uniformly distributed, then the time series should have a Gaussian distribution, and in fact it does.

Most of our simulations were made by using a boxcar spectrum, which can be thought of as the result of applying a perfect bandpass filter to white noise. If the passband is given by

$$-\varepsilon \le \pi(\omega - \omega_0)/\omega_0 \le \varepsilon \quad (14)$$

then evaluation of (8)–(10) gives

$$r = \rho = \sin \varepsilon/\varepsilon \quad (15)$$

Simulations for various values of r were made with $\omega_0 = 0.408$ s^{-1}, $\Delta t = 0.5$, and Fourier transforms of length 8192. The choices of ω_0 and Δt insured that the wave peaks would be adequately sampled. The simulations were repeated 1500 times to produce approximately 400,000 waves for each spectrum. The sample statistics are thus reliable to low probability levels. Zero-crossing wave heights were defined as the dis-

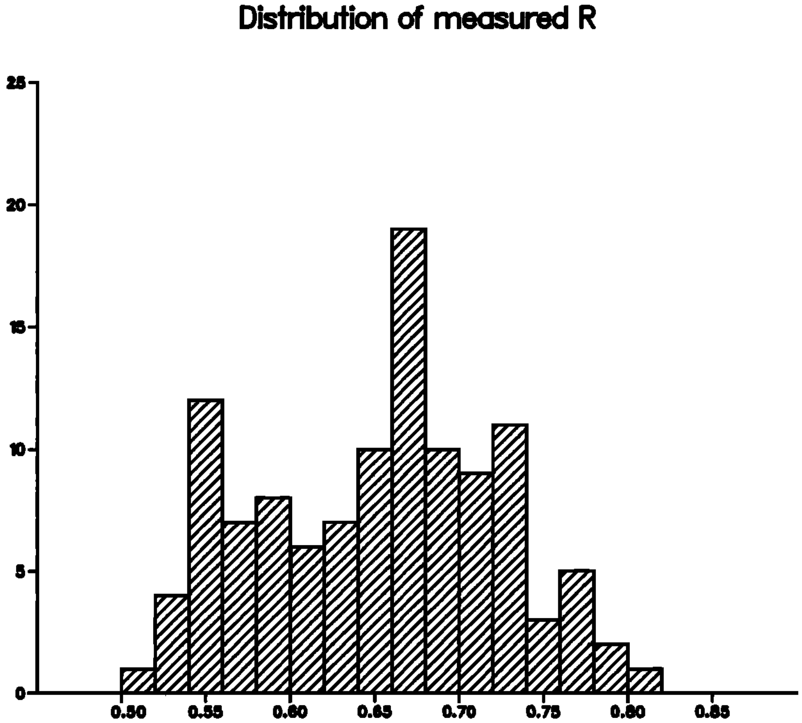

Fig. 5. Histogram of spectral parameter r for the wave records used to develop Forristall's empirical distribution.

tances between the highest and lowest samples found between two successive zero down-crossings.

Figure 2 shows comparisons between the sample distributions of the simulations and *Tayfun*'s [1981*b*] distribution for various values of r. The form of the plot is designed to emphasize the differences between distributions that are really quite similar. The abscissa is the probability that a given normalized height will be exceeded. The ordinate is the ratio between that height and the height having the same probability according to the Rayleigh distribution. These are the ratios of the distributions along horizontal lines in plots with the form of Figure 1. For $r = 1$, (12) converges to the Rayleigh distribution, which would plot as a straight line with an ordinate of 1 in Figure 2. The theoretical distributions are shown by solid lines, and the sample distributions from the simulations are shown by dashed lines with data points.

Figure 2 only shows about the higher half of the wave distributions, since the lower waves have probabilities higher than the Rayleigh distribution as the parameter r decreases. For the waves shown the height decreases steadily as the spectrum becomes broader and r decreases. The agreement between the theoretical and simulated distributions is remarkably good. Some discrepancy is evident only for the broadest spectra, where the theoretical distribution is a bit too low. Even there, the disagreement is only about 2% in wave height.

The shape of the distribution is also accurately predicted. The curves droop farther below the Rayleigh distribution as the probability decreases. The methods of *Nolte and Hsu* [1979] and *Longuet-Higgins* [1980] produce distributions that would plot as horizontal lines on Figure 2.

Simulations of very narrow spectra must be made with long Fourier transforms to ensure enough variability. Figure 3 shows the result of simulating a boxcar spectrum with $r = 0.9$ and a transform length of 2048 instead of 8192. The higher wave heights predicted by the theory were not produced. For the short transform there were only 34 Fourier lines in the boxcar, and this was evidently few enough to reduce the probability of the highest waves. The sample distributions for the widths shown in Figure 2 did not change significantly for transform lengths longer than 8192.

Simulations of spectra with more natural shapes can also be made. Figure 4 shows wave height distributions for a Pierson-Moskowitz spectrum with $\omega_0 = 0.446$. Numerical integration of the spectrum in (9) and (10) gives $r = 0.68$, and *Tayfun*'s [1981*b*] distribution for this value is plotted as the solid line in the figure. The result of simulating a boxcar spectrum with

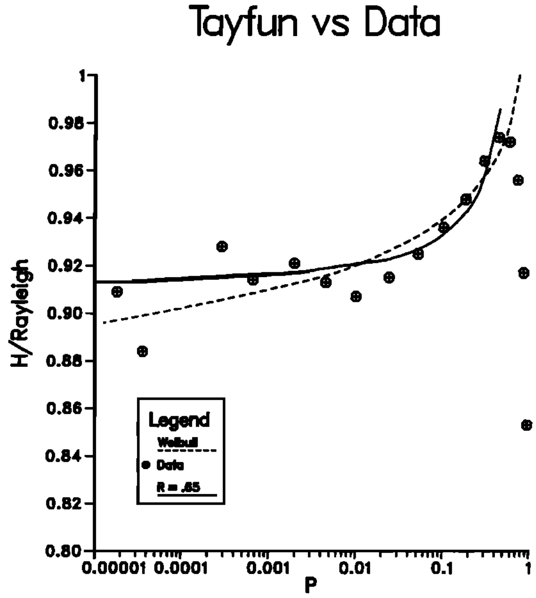

Fig. 6. Measured, fitted, and theoretical distributions: (data points) measurements used by *Forristall* [1978]; (dashed line) empirically fit Weibull distribution; (solid line) Tayfun distribution for $r = 0.65$.

the same value of r is also shown. The agreement between the three curves is excellent. However, the simulated Pierson-Moskowitz spectrum has too few low waves. The agreement, using rather broad spectra, indicates that the approximations made in deriving (12) were justified.

Longuet-Higgins [1980] calculations gave $\bar{a}/m_0^{1/2} = 0.931$ for a Pierson-Moskowitz spectrum, so his distribution would plot as a straight line with an ordinate of 0.931 on Figure 4. This is a good approximation for the probabilities of the highest waves but misses the shape of the distribution in the intermediate range.

Since *Tayfun*'s [1981*b*] distribution depends on the trigonometric moments (9) and (10), it is not sensitive to the details of the high-frequency tail of the distribution or a high-frequency cutoff in simulations or data. This is a pleasant contrast to distributions that depend on the spectral width defined by (4). If a Pierson-Moskowitz spectrum is cut off at $2\omega_0$, then $\nu = 0.30$ and $r = 0.68$. If the cutoff is raised to $10\omega_0$, then $\nu = 0.42$ and $r = 0.68$.

MEASURED WAVE HEIGHTS

The agreement of the theoretical distribution with simulated wave records indicates that the mathematical approximations used in its derivation do not cause large errors for representative spectra. However, other physical factors could cause the distribution of measured wave heights to differ from the theory. Comparisons of measured distributions with the theory should indicate whether any of these factors are important.

Nature will not cooperate by providing a long series of waves with a constant spectral shape. Ideally, one could still find a number of spectra with the same value of the parameter r. Figure 5 shows the distribution of the values of r for the wave records used by *Forristall* [1978]. The distribution is rather narrow, with a mean of $r = 0.65$. The mean spectral shape is just slightly broader than the Pierson-Moskowitz form. Since Figure 2 indicates that the variation of *Tayfun*'s [1981*b*] distribution with r is nearly linear, it is reasonable to use the mean values of r as representative of the entire data set.

Figure 6 shows the comparison of Forristall's measured and empirically fitted distributions with *Tayfun*'s [1981*b*] distribution for $r = 0.65$. The three agree quite closely. There is more statistical variability in the data at low probabilities than there was for the simulations, since the data only included about 55,000 waves. It seems possible that Tayfun's distribution fits the trend of the data at low probabilities better than the empirical distribution.

The data and *Tayfun*'s [1981*b*] distribution are compared in an ordinary exceedance diagram in Figure 7. The fit is excellent. Both Figures 6 and 7 show a slight deficit of very low waves in the data, similar to that found in the Pierson-Moskowitz simulations.

CONCLUSIONS

The distributions of measured and simulated zero-crossing wave heights agree excellently with a distribution derived by *Tayfun* [1981*b*] from the work of *Rice* [1945]. It seems that, given the wave spectrum, the expected value of the wave height can be predicted to within 1%. The trigonometric moments of the spectrum given by (9) and (10) give a more appropriate measure of the spectral shape than the spectral width.

The theoretical distribution describes measured and simulated wave heights equally well. Simulated and measured waves with the same value of r have the same wave height distribution. It thus seems likely that wave breaking and nonlinear phase locking of wave components have little effect on the distribution. *Osborne* [1982] also reached this conclusion from his simulations. The agreement of measurements and simulations is also a good argument for the usefulness of linear simulations of waves.

Tayfun's [1981*b*] wave height distribution should be used whenever great accuracy is demanded, and the spectral shape is known from measurements or wave hindcasts. The approximate method resulting from *Longuet-Higgins* [1980] gives reasonable values for the low-probability tail of the distribution. The empirical distribution proposed by *Forristall* [1978] is still a useful approximation, if details of the spectral shape are not known, since the wave spectra that were used in deriving it were representative of storm conditions. This empirical distribution also offers some advantage in ease of manipulation. Finally, if heights good to within 10% are adequate, the Rayleigh distribution is still useful.

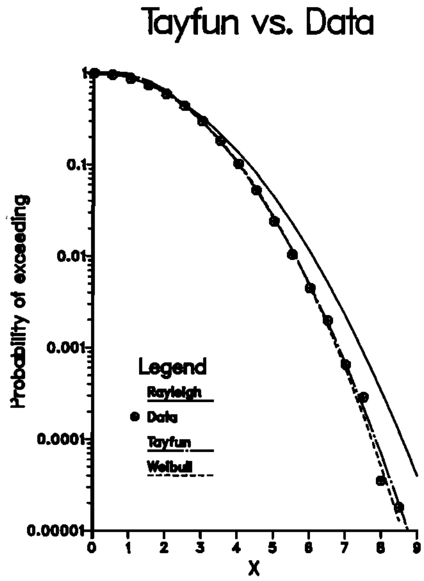

Fig. 7. Wave height distributions: (solid line) Rayleigh distribution; (data points) measurements used by *Forristall* [1978]; (dash-dot line) Tayfun distribution for $r = 0.65$, (dashed line) empirical Weibull distribution.

REFERENCES

Borgman, L. E., Ocean wave simulation for engineering design, *J. Waterways Harbors Div., Am. Soc. Civil Eng.*, *95*, 557–583, 1969.

Cartwright, D. E., and M. S. Longuet-Higgins, The statistical distribution of the maxima of a random process, *Proc. R. Soc. London, Ser. A*, *237* (1209), 212, 1976.

Chen, E., L. E. Borgman, E. Yfantis, Height and period distribution of hurricane waves, paper presented at Civil Engineering in the Oceans IV, Am. Soc. Civil Eng., San Francisco, September 10–12, 1979.

Forristall, G. Z., On the statistical distribution of wave heights in a storm, *J. Geophys. Res.*, *83*, 2353–2358, 1978.

Haring, R. E., A. R. Osborne, and L. P. Spencer, Extreme wave parameters based on continental shelf storm wave records, paper presented at 15th Coastal Engineering Conference, Am. Soc. Civil Eng., Honolulu, July 11–17, 1976.

Larsen, L. H., The influence of bandwidth on the distribution of heights of sea waves, *J. Geophys. Res.*, *86*, 4299–4301, 1981.

Longuet-Higgins, M. S., On the statistical distribution of the heights of sea waves, *J. Mar. Res.*, *11*, 245–266, 1952.

Longuet-Higgins, M. S., On the distribution of the heights of sea waves: Some effects of nonlinearity and finite and bandwidth, *J. Geophys. Res.*, *85*, 1519–1523, 1980.

Nolte, K. G., and F. H. Hsu, Statistics of larger waves in a sea state, *J. Waterways Harbors Div., Am. Soc. Civil Eng.*, *105*, 389–404, 1979.

Osborne, A. R., The simulation and measurement of random ocean wave statistics, in *Topics in Ocean Physics*, edited by A. R. Osborne and P. M. Rizzoli, North-Holland, Amsterdam, 1982.

Rice, S. O., The mathematical analysis of random noise, *Bell Syst. Tech. J.*, *24*, 46–156, 1945.

Tayfun, M. A., Narrow-band nonlinear sea waves, *J. Geophys. Res.*, *85*, 1548–1552, 1980.

Tayfun, M. A., Breaking-limited wave heights, *J. Waterways Harbors Div., Am. Soc. Civil Eng.*, *107*, 59–69, 1981*a*.

Tayfun, M. A., Distribution of crest-to-trough wave heights, *J. Waterways Harbors, Am. Soc. Civil Eng.*, *107*, 149–158, 1981*b*.

Thompson, E. F., Results from CERC wave measurement program, paper presented at International Symposium on Ocean Wave Measurement and Analysis, Am. Soc. Civil Eng., New Orleans, September 9–11, 1974.

G. Z. Forristall, Shell Development Company, Box 481, Houston, TX 77001.

(Received April 25, 1984;
revised May 24, 1984;
accepted May 24, 1984.)