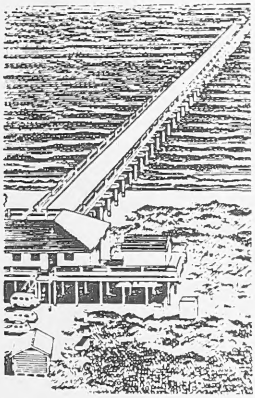
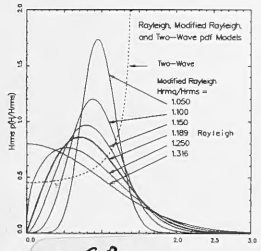
US Army Corps of Engineers





COASTAL ENGINEERING RESEARCH CENTER

TECHNICAL REPORT CERC-91-7

WAVE HEIGHT DISTRIBUTIONS IN MULTIPLE-PEAKED SEAS

by

Charles E. Long

Coastal Engineering Research Center

DEPARTMENT OF THE ARMY
Waterways Experiment Station, Corps of Engineers
3909 Halls Ferry Road, Vicksburg, Mississippi 39180-6199

July 1991
Final Report

Prepared for DEPARTMENT OF THE ARMY
US Army Corps of Engineers
Washington, DC 20314-1000

Under Civil Works Research Work Unit 32484

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 1991	Final report

4. TITLE AND SUBTITLE

Wave Height Distributions in Multiple-Peaked Seas

5. FUNDING NUMBERS

Civil Works Research
Work Unit 32484

6. AUTHOR(S)

Charles E. Long

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

USAE Waterways Experiment Station
Coastal Engineering Research Center
3909 Halls Ferry Road
Vicksburg, MS 39180-6199

8. PERFORMING ORGANIZATION REPORT NUMBER

Technical Report
CERC-91-7

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

US Army Corps of Engineers
Washington, DC 20314-1000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

Available from National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Knowledge of distributions of heights of sea waves under realistic oceanic conditions is critical to the successful design of coastal shore protection projects. In this report, a simple examination is made of wave height distributions in sea states characterized by energy spectra with multiple peaks, i.e., having energy centered at two or more distinct frequencies. Sea states with broad frequency spectra are also included since such seas are composed of waves of diverse frequencies. These cases violate the assumptions of unimodal, narrow spectra that are formally required for the conventionally used Rayleigh distribution of wave heights to apply.

Reported here are tests of the Rayleigh and Modified Rayleigh wave height distribution models. These models are compared with wave heights

(Continued)

14. SUBJECT TERMS

Height distributions
Multimodal spectra
Wave heights
Wind waves

15. NUMBER OF PAGES

63

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED		

13. (Concluded).

from idealized, synthetic time series having spectra with variable widths, modes, and modal separations. The Modified Rayleigh model is the two-parameter, deepwater asymptotic form of the Beta-Rayleigh distribution.

The first part of this report involves determining some of the limits implied by the phrase "unimodal, narrow spectrum" in the synthesis of time series having Rayleigh distributions of wave heights. It is found that for unimodal, band-limited, white spectra, there are constraints on both the overall bandwidth and the number of component waves for a close approximation to a Rayleigh process to occur.

The primary part of this investigation is examination of wave height distributions in wave signals derived from multimodal spectra. Using criteria established in the unimodal tests, bimodal spectra are constructed from two unimodal spectra, each of which yields a Rayleigh wave height distribution by itself. The two primary variables in these bimodal tests are a measure of modal separation in the frequency domain and the relative amount of energy in one mode compared with the other.

Spectra with broad ranges of modal separation parameter and with a ratio of low-frequency modal energy to high-frequency modal energy roughly greater than 1.0 have wave height distributions that are well-represented by the Modified Rayleigh model. The Rayleigh model tends to overpredict wave height averages for these parametric ranges in much the same way as was found for time series derived from broad-banded spectra and as have been found frequently in natural observations. For large modal separations and small high-frequency modal energy relative to low-frequency modal energy, synthetic data deviate strongly from the Rayleigh model and deviate significantly from the Modified Rayleigh model. Further study is needed to determine the relative importance of such conditions. Further research is also required to determine the parameters of the Modified Rayleigh model in terms of spectral parameters. Finally, these conclusions need to be tested with natural data to ensure that the artifice of synthesis has not biased the results.

PREFACE

The study in this report was authorized by Headquarters, US Army Corps of Engineers (HQUSACE), Coastal Engineering Area of Civil Works Research and Development. Work was performed under Civil Works Research Work Unit 32484, "Directionality of Waves in Shallow Water," Coastal Flooding Program, at the Coastal Engineering Research Center (CERC) of the US Army Engineer Waterways Experiment Station (WES). The HQUSACE Technical Monitors were Messrs. John H. Lockhart, Jr.; John G. Housley; James E. Crews; and Robert H. Campbell. Dr. C. Linwood Vincent was the CERC Program Manager.

This investigation was conducted by Dr. Charles E. Long at the WES/CERC Field Research Facility (FRF) in Duck, NC. The purpose of this study was to help clarify the behavior of wave height statistics for sea states characterized by broad or multimodal spectra.

The report was prepared under the direct supervision of Mr. William A. Birkemeier, Chief, FRF, and under the general supervision of Mr. Thomas W. Richardson, Chief, Engineering Development Division; Mr. Charles C. Calhoun, Jr., Assistant Chief, CERC; and Dr. James R. Houston, Chief, CERC. This report was edited by Ms. Lee T. Byrne, Information Technology Laboratory, WES.

Commander and Director of WES during publication of this report was COL Larry B. Fulton, EN. Dr. Robert W. Whalin was Technical Director.

CONTENTS

WAVE HEIGHT DISTRIBUTIONS IN MULTIPLE-PEAKED SEAS

PART I: INTRODUCTION

1. Successful design of shore protection projects depends critically on knowledge of water level variation. Sea surface elevations change because of a suite of processes that include tides, storm surge, and wind waves. Wind waves do significant amounts of work on coastal boundaries, in general. Of particular importance is that water level changes associated with heights of wind waves contribute to extremes of beach and structural runup and hence to high beach erosion rates or potential for overtopping and flooding behind coastal defenses. For engineering design purposes, it is useful to have a statistical description of the heights of sea waves so that probabilities can be assigned to particular water level extremes. With these probabilities, a likelihood of project survival can be estimated. Ideally, probable distributions of wave heights in a given sea state can be described with a mathematical model. An important problem in coastal engineering is the determination of a model that can represent realistic sea states with reasonable fidelity.

2. One of the earliest and most widely used models is the Rayleigh probability density function (pdf), first applied in ocean work by Longuet-Higgins (1952). It is a one-parameter model, relying only on specification of a root-mean-square (RMS) wave height to allow estimation of the probability that a randomly chosen wave height will fall within a small range about a specified wave height. As derived by Longuet-Higgins (1952), the Rayleigh pdf is valid formally only when an ocean surface can be described by the linear sum of a very large number of randomly phased and directed sinusoidal components which conform to a unimodal frequency spectrum wherein all frequencies are very nearly (but not quite) equal to some central, characteristic frequency. Unfortunately, not all sea states have unimodal, narrow-banded spectra. In a rather broad study of naturally occuring sea states, Thompson (1980) noted that well over half of his observed spectra were distinctly multimodal. This condition means that one can expect wave height distributions to deviate to some extent from a Rayleigh distribution. If such a deviation is important, alternative pdf models must be sought.

3. A comparison of the Rayleigh pdf model with a large set of observations is described in the Shore Protection Manual (SPM) (1984). That comparison indicates that if both the number of waves and the RMS wave height from a sea surface elevation record are modified to minimize differences between model and data, a discrepancy of 10 to 15 percent remains in the low-probability but high-wave tails of the distributions. These tails are important for estimating extreme wave conditions, so it is important that they be modeled correctly. The above result suggests that either the model or the data (or both) are not representative of real ocean conditions to within the stated percentages.

4. On the other hand, Longuet-Higgins (1980) cites favorable comparisions of the Rayleigh pdf with data reported by Earle (1975) and Forristall (1978). To achieve a favorable comparison, Longuet-Higgins renormalized Forristall's data with a spectrally based characteristic wave height modified for effects of finite bandwidth. His argument was justified in that the RMS wave height is the controlling parameter in a Rayleigh distribution and the relationship between RMS wave height and spectrum-based wave height can be altered when a narrow-band process is perturbed by low-level energy at frequencies away from the peak frequency, as occurs often in nature.

5. Thornton and Guza (1983) extended tests of the Rayleigh pdf to include highly nonlinear, actively breaking wave conditions in very shallow water. They used data from a number of sensors in water depths as shallow as 1 m in the breaker zone at Torrey Pines Beach, California. The Rayleigh pdf provided very good estimates of wave height statistics in comparison with their observations. There was, however, a very slight overprediction of the wave population in the high-wave tail of the pdf, somewhat like the result given in the SPM (1984).

6. None of the above tests specifically address the problem of height distributions under conditions where an energy spectrum is distinctly multimodal. A simple example of such a condition occurs where two narrow-band processes coexist but at well-separated frequencies. In nature, such a condition can arise where a local, wind-driven sea is generated in the presence of low-frequency swell. In general, one would not expect wave heights from this scenario necessarily to be Rayleigh distributed because the requirement of a unimodal, narrow-band process has been violated. In light of

Thompson's (1980) findings, it appears that multimodal processes are common in nature so that one is obligated to investigate wave height distributions in them.

Purpose of Study

7. The intent of the study reported here is to conduct a preliminary examination of wave height distributions that arise in multimodal processes. Only bimodal processes are considered here in order to keep the investigation simple. To maintain maximum control of experiment conditions, time series from which to analyze extrema are artificially produced linear sums of sinusoidal components. This procedure makes the study somewhat idealized, but avoids some of the vagaries associated with field data. Amplitudes and frequencies of component wave signals are established through a spectral definition. Phases of component signals are selected at random.

8. Statistical distributions of artificially generated wave heights are compared with two distribution models. The first is the Rayleigh model discussed above. The second model is the deepwater asymptotic form of the so-called Beta-Rayleigh distribution introduced by Hughes and Borgman (1987). This second model is a two-parameter function that collapses to the Rayleigh model for a specific ratio of its parameters. In general, it is less restrictive than the Rayleigh model (because it has two parameters instead of just one) and is called the Modified Rayleigh model. It is included in this analysis because an analysis by Long (in preparation) indicates that it represents observations better than either the Rayleigh model or the full Beta-Rayleigh model for waves having a variety of directional distributions in intermediate and shallow water. The hypothesis posed here is that the Modified Rayleigh pdf may represent adequately distributions that deviate significantly from Rayleigh distributions due to either broad-bandedness or multimodality.

Scope of Investigation

9. Mathematical descriptions of the models used in this study are given in Part II. Part III describes the way in which test data are generated. Part IV discusses an analysis of unimodal processes. This analysis is necessary because it is not clear how many components distributed over what bandwidths are necessary to approximate reasonably a Rayleigh process. Such

treatment also helps in understanding what happens in spectral filtering of time series to eliminate unwanted parts of a signal. The analysis discussed in Part V uses the results of Part IV to generate some simple bimodal processes where each mode is individually a Rayleigh process. Results can then be classified purely by modal separation and ratio of modal energies. Comparison of analyzed synthetic time series with model curves then reveals which processes are approximately Rayleigh and under which, if any, circumstances the Modified Rayleigh model excels.

PART II: MODEL DEFINITIONS

Basic Statistical Notation

10. The models used here are all defined* basically as probability density functions p of specified crest-to-trough wave height H in the form $p(H)$, having dimensions of inverse length. When multiplied by an incremental range of height dH, the resulting expression $p(H)\,dH$ is dimensionless and gives the probability that a randomly chosen height $\hat{H}$ lies in the range between H and $H + dH$. This relationship is written as

$$p(H)\,dH = \text{Prob}[H < \hat{H} < H + dH] \tag{1}$$

where Prob means *the probability that*.

11. If Equation 1 is integrated over all increments of height between zero and some specified height H, one obtains the *cumulative distribution function* (cdf) denoted as $P(H)$, which is the probability that a randomly chosen height $\hat{H}$ is less than a specified height H. This expression takes the form

$$P(H) = \int_0^H p(x)\,dx$$

$$= \text{Prob}[\hat{H} < H] \tag{2}$$

where x is simply the dummy integration variable. In Equation 2, it can be seen that $P(\infty) = 1$ because all wave heights are less than infinitely high. One can take the complement of the cumulative probability to obtain the *exceedence probability function* $Q(H)$, which is the probability that a randomly chosen height $\hat{H}$ is greater than a specified height H. Formally, $Q(H)$ is found from the probability density function by

$$Q(H) = \int_H^\infty p(x)\,dx$$

$$= \text{Prob}[\hat{H} > H] \tag{3}$$

---

* For convenience, symbols and abbreviations are listed in the Notation (Appendix A).

If P(H) is known, Q(H) can also be computed from the result of the following derivation

$$Q(H) = \int_0^{\infty} p(x)\, dx - \int_0^{H} p(x)\, dx$$

$$= P(\infty) - P(H)$$

$$= 1 - P(H) \qquad (4)$$

12. Conventional statistical definitions allow other properties of a wave height distribution to be found from basic probability density functions. For instance, the *mode*, or most probable wave height, is the maximum of p(H) . The *mean*, or average wave height is the integral of the product H p(H) over all H . The point here is that all statistical properties of interest for a given process can be found once the probability density function is known. It is this function, therefore, that is most fundamental to define.

Two-Wave Model

13. In the analysis described below, synthetic series of points are computed by summing a number of sinusoidal components. Aside from the trivial case of a single sine wave (where mean, mode, RMS, and all other measures of wave height are constant and equal to each other), the simplest sum is of two waves of equal amplitudes, very slightly different frequencies, and arbitrary initial phases. These two wave trains will beat together, forming the familiar grouping pattern where the waves are nearly in phase and then become out of phase over relatively long periods.

14. Statistical properties of this two-wave process were derived by Longuet-Higgins (1952). In particular, he found the pdf to be

$$p_2(H) = \begin{cases} \dfrac{2}{\pi(2H_{rms}^2 - H^2)^{1/2}} & H < \sqrt{2}\, H_{rms} \\ 0 & H > \sqrt{2}\, H_{rms} \end{cases} \qquad (5)$$

where the subscript 2 indicates this two-wave model and H_{rms} is the RMS wave height, determined from a set of observed wave heights through the equation

$$H_{rms} = \left(\frac{1}{N} \sum_{n=1}^{N} H_n^2 \right)^{1/2} \tag{6}$$

In Equation 6, N is the number of observed waves and H_n is the height of the n^{th} wave. It is useful to treat mathematical models in dimensionless form, where practicable. The only parameter in Equation 5 is H_{rms}. Hence, the pdf can be made dimensionless in the form

$$H_{rms}\, p_2\left(\frac{H}{H_{rms}}\right) = \begin{cases} \dfrac{2}{\pi\left[2 - \left(\dfrac{H}{H_{rms}}\right)^2\right]^{1/2}} & \dfrac{H}{H_{rms}} < \sqrt{2} \\ 0 & \dfrac{H}{H_{rms}} > \sqrt{2} \end{cases} \tag{7}$$

The exceedence probability corresponding to this pdf is found by integrating Equation 7 in accordance with Equation 3. The result, in dimensionless form, is

$$Q_2\left(\frac{H}{H_{rms}}\right) = \begin{cases} 1 - \dfrac{2}{\pi} \sin^{-1} \dfrac{1}{\sqrt{2}} \dfrac{H}{H_{rms}} & \dfrac{H}{H_{rms}} < \sqrt{2} \\ 0 & \dfrac{H}{H_{rms}} < \sqrt{2} \end{cases} \tag{8}$$

15. The two-wave model is useful for testing synthetic data generation. When used with two wave trains of arbitrarily different frequencies, synthetic wave heights should approximate this statistical distribution, becoming asymptotically identical as the two frequencies become nearer to (but not the same as) each other. For these conditions to hold, synthetic time series must be long enough to include sufficient cycles of the beat period to obtain a meaningful sample of wave heights.

Rayleigh Model

16. As described by Longuet-Higgins (1952), the Rayleigh pdf is defined by the equation

$$p_R(H) = \frac{2H}{H_{rms}^2} e^{-(H/H_{rms})^2} \tag{9}$$

where the subscript R identifies the pdf as Rayleigh. Equation 9 can be written in dimensionless form by multiplying both sides by H_{rms} to form

$$H_{rms}\ p_R\left(\frac{H}{H_{rms}}\right) = 2\ \frac{H}{H_{rms}}\ e^{-(H/H_{rms})^2} \tag{10}$$

17. The cumulative probability function for the Rayleigh distribution is found in dimensionless form by integrating Equation 10 with respect to H/H_{rms} over the limits from zero to H/H_{rms}. The result is

$$P_R\left(\frac{H}{H_{rms}}\right) = e^{-(H/H_{rms})^2} \tag{11}$$

The Rayleigh exceedence probability function is found by substituting Equation 11 in Equation 4 to yield

$$Q_R\left(\frac{H}{H_{rms}}\right) = 1 - e^{-(H/H_{rms})^2} \tag{12}$$

Modified Rayleigh Model

18. The Modified Rayleigh pdf is the deepwater asymptotic form of the Beta-Rayleigh model introduced by Hughes and Borgman (1987). They defined the Modified Rayleigh pdf as

$$H_{rms}\ p_{MR}\left(\frac{H}{H_{rms}}\right) = \frac{2\alpha^{\alpha}}{\Gamma(\alpha)}\left(\frac{H}{H_{rms}}\right)^{2\alpha-1} e^{-\alpha(H/H_{rms})^2} \tag{13}$$

where the subscript MR stands for Modified Rayleigh, Γ is the gamma function (see Abramowitz and Stegun 1970), and α is a parameter defined as

$$\alpha = \frac{1}{\left(\frac{H_{rmq}}{H_{rms}}\right)^4 - 1} \tag{14}$$

In Equation 14, α depends on H_{rmq}, which stands for *root-mean-quad wave height* and is computed from a set of N observed wave heights H_n through the expression

$$H_{rmq} = \left(\frac{1}{N}\sum_{n=1}^{N} H_n^4\right)^{1/4} \tag{15}$$

Note that when $\alpha = 1$ (or $H_{rmq}/H_{rms} = 2^{1/4}$ from Equation 14), Equation 13 becomes identical to Equation 10 and the Modified Rayleigh pdf becomes a Rayleigh pdf. It should be noted that Hughes and Borgman (1987) defined H_{rmq} using a square root instead of the fourth root used here on the right side of Equation 15. It makes more sense to use the fourth root since it gives a height parameter with dimensions of length instead of length squared. This is only a modification of notation and not of the basic model. One would simply use the square root of the H_{rmq} in Equation 14 above.

19. To find the cumulative distribution function of the Modified Rayleigh model, one must integrate the pdf numerically. For computations used in this study, integration was done using Simpson's rule in double precision with integration steps of 0.001 in H/H_{rms} .

20. Examination of Equations 13 and 14 reveals that the shape of the dimensionless pdf depends only on one parameter, the ratio H_{rmq}/H_{rms} . Figure 1 shows a variety of Modified Rayleigh distributions found for select values of H_{rmq}/H_{rms} . The bold curve in Figure 1 is the Rayleigh pdf to which the Modified Rayleigh pdf degenerates when $H_{rmq}/H_{rms} = 2^{1/4}$. For completeness, the two-wave pdf is also shown in Figure 1.

Additional Parameters

21. Where spectra are used for synthetic time series generation, some additional parameters can be defined. These parameters are based solely on spectral shape. Perhaps the most frequently used spectral scale of wave height is denoted H_{mo} and is defined by

$$H_{mo} = 4\, m_o^{1/2} \tag{16}$$

In Equation 16, m_o is the spectral zeroth moment. The more general n^{th} spectral moment m_n is defined by

$$m_n = \int_{f_1}^{f_2} f^n\, S(f)\, df \tag{17}$$

where f is frequency; f_1 and f_2 are, respectively, low and high frequency bounds of the spectral definition; and S is signal variance spectral density.

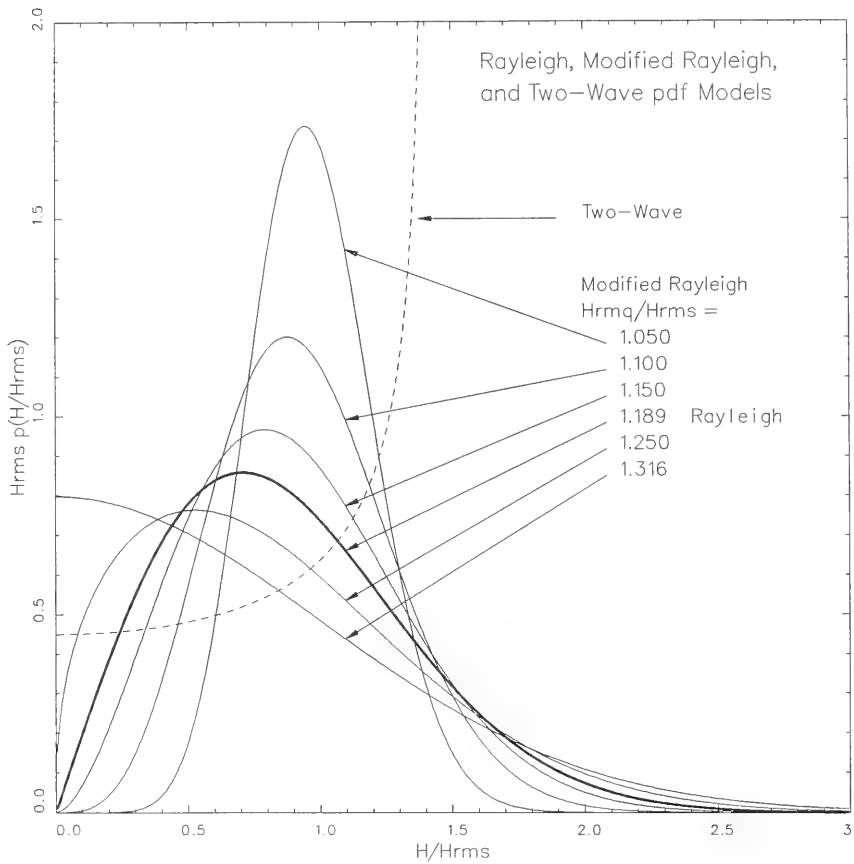


Figure 1. Examples of model curves

22. Ideally, $f_1 = 0$ and $f_2 = \infty$, but in practice, the two frequencies are bound to less diverse values. Where purely synthetic time series are used, a low frequency of $f_1 = 0$ is realizable, but it results only in a constant offset of the signal that would normally be removed in subsequent wave height analysis. For computer-generated discrete time series, the lowest practical frequency f_1 is that which allows enough cycles that wave heights extracted from the time series include the effects of these low-frequency waves.

23. The constraint on the upper frequency f_2 is related to the time step Δt used to simulate a time series. If a frequency is too high, there will be too few time steps per wave. The probability will then be small that the crests and troughs of these waves will be represented in the discrete sample. For practical purposes, about 20 time steps per wave is a reasonable sampling. A crest or trough is then always within one-fortieth of a wave period of some discrete time step. The fractional error in crest or trough estimation is then the ratio of the cosine of one-fortieth of a wave cycle (9 deg) to the cosine of zero degrees. This ratio is about 0.988, corresponding to an error of about 1.2 percent in crest elevation estimation or about 2.4 percent overall. A frequency upper limit of 20 time steps then requires $f_2 \leq 1/(20\ \Delta t)$.

24. Another parameter of use is one that gives a measure of spectral width. A wave height distribution is theoretically Rayleigh if the corresponding frequency spectrum is narrow-banded. A measure of such spectral width has been given by Cartwright and Longuet-Higgins (1956) as the parameter ϵ defined by

$$\epsilon = 1 - \frac{m_2^2}{m_0 m_4} \tag{18}$$

where m_o , m_2 , and m_4 are determined through Equation 17. A small ϵ indicates a narrow spectrum, and an ϵ near 1 indicates a broad spectrum.

25. Further parameters associated with spectral definitions are given in Part III, wherein test data generation is described. Parameters associated with averages of particular subsets of wave heights (for example, average of the highest one-third) are defined in Part IV, where model test criteria are described.

PART III: TEST DATA GENERATION

26. In this study, wave height distributions are extracted from synthetic time series that have specific spectral definitions in the frequency domain. A very efficient computational scheme for synthesizing a time series from spectra is by way of inverse Fourier transforms. In this chapter, this technique and the nature of the spectra used in this study are described.

Inverse Fourier Transform Technique

27. Following the description by Bendat and Piersol (1971), a time series of N water surface elevations or an artificial representation thereof can be represented by the notation $x_n = x(t_n) = x[(n - 1)\Delta t]$, for $n = 1, 2, 3, \ldots, N - 1, N$ where x_n is the n^{th} discrete sample of a continuous function $x(t)$ and t_n is the n^{th} discrete increment of time Δt after an initial time of zero for the first sample ($n = 1$). A second, independent, N-point time series can be represented by the notation $y_n = y(t_n) = y[(n - 1)\Delta t]$, for $n = 1, 2, \ldots, N$. Each of these time series can be subjected to a discrete Fourier transformation (DFT). For time series x_n, the k^{th} element of the DFT is denoted by X_k, and the whole set is defined by

$$X_k = \sum_{n=1}^{N} x_n \, e^{-2\pi i(n-1)(k-1)/N} \qquad k = 1, 2, \ldots, N \tag{19}$$

For the second time series, DFT elements are denoted by Y_k and are defined by

$$Y_k = \sum_{n=1}^{N} y_n \, e^{-2\pi i(n-1)(k-1)/N} \qquad k = 1, 2, \ldots, N \tag{20}$$

28. Because each time series represents a piecewise continuous function, each can also be represented as a Fourier series of points. The first time series can thus be written

$$x_n = \sum_{k=1}^{N/2} \left\{ A_{xk} \cos\left[\frac{2\pi(k - 1)(n - 1)}{N}\right] + B_{xk} \sin\left[\frac{2\pi(k - 1)(n - 1)}{N}\right] \right\} \qquad n = 1, 2, \ldots, \frac{N}{2} \tag{21}$$

where A_{xk} and B_{xk} are the k^{th} Fourier cosine and sine coefficients, respectively, of time series x_n. Similarly, the second time series can be expressed in the form

$$y_n = \sum_{k=1}^{N/2} \left\{ A_{yk} \cos\left[\frac{2\pi(k-1)(n-1)}{N}\right] + B_{yk} \sin\left[\frac{2\pi(k-1)(n-1)}{N}\right] \right\} \qquad n = 1, 2, \ldots, \frac{N}{2} \tag{22}$$

where A_{yk} and B_{yk} are the k^{th} Fourier cosine and sine coefficients, respectively, of time series y_n. The Fourier coefficients can be found from the time series points through the formulae

$$A_{x1} = \frac{1}{N} \sum_{n=1}^{N} x_n \tag{23}$$

$$A_{xk} = \frac{2}{N} \sum_{n=1}^{N} x_n \cos\left[\frac{2\pi(k-1)(n-1)}{N}\right] \qquad k = 2, 3, \ldots, \frac{N}{2} \tag{24}$$

$$B_{x1} = 0 \tag{25}$$

$$B_{xk} = \frac{2}{N} \sum_{n=1}^{N} x_n \sin\left[\frac{2\pi(k-1)(n-1)}{N}\right] \qquad k = 2, 3, \ldots, \frac{N}{2} \tag{26}$$

for time series x_n, and the formulae

$$A_{y1} = \frac{1}{N} \sum_{n=1}^{N} y_n \tag{27}$$

$$A_{yk} = \frac{2}{N} \sum_{n=1}^{N} y_n \cos\left[\frac{2\pi(k-1)(n-1)}{N}\right] \qquad k = 2, 3, \ldots, \frac{N}{2} \tag{28}$$

$$B_{y1} = 0 \tag{29}$$

$$B_{yk} = \frac{2}{N} \sum_{n=1}^{N} y_n \sin\left[\frac{2\pi(k-1)(n-1)}{N}\right] \qquad k = 2, 3, \ldots, \frac{N}{2} \tag{30}$$

for time series y_n.

29. Noting that the real part of Equation 19 is identical to the summation term in Equation 24 and that the imaginary part of Equation 19 is identical to minus the summation term in Equation 26, it is clear that

$$A_{x1} - i\, B_{x1} = \frac{1}{N} X_1 \tag{31}$$

and

$$A_{xk} - i\, B_{xk} = \frac{2}{N} X_k \qquad k = 2, 3, \ldots, \frac{N}{2} \tag{32}$$

Similarly,

$$A_{y1} - i\, B_{y1} = \frac{1}{N} Y_1 \tag{33}$$

and

$$A_{yk} - i\, B_{yk} = \frac{2}{N} Y_k \qquad k = 2, 3, \ldots, \frac{N}{2} \tag{34}$$

for time series y_n. These relationships show how the Fourier series coefficients and Fourier transform components are related.

30. It is also useful to note that the Fourier series of Equation 21 can be written as

$$x_n = \sum_{k=1}^{N/2} C_{xk} \cos\left[\frac{2\pi(k - 1)(n - 1)}{N} + \phi_{xk}\right] \qquad n = 1, 2, \ldots, N \tag{35}$$

where the cosine and sine amplitudes A_{xk} and B_{xk}, respectively, of Equation 21 are related to the modulus C_{xk} and phase ϕ_{xk} of Equation 35 by

$$A_{xk} = C_{xk} \cos \phi_{xk} \tag{36}$$

and

$$B_{xk} = C_{xk} \sin \phi_{xk} \tag{37}$$

or

$$C_{xk}^2 = A_{xk}^2 + B_{xk}^2 \tag{38}$$

and

$$\phi_{xk} = \tan^{-1} \frac{B_{xk}}{A_{xk}} \tag{39}$$

These relationships allow definition of an individual wave component in terms of a wave amplitude C_{xk} and initial phase ϕ_{xk}. These relationships are useful when defining a wave field from a spectrum as is done below.

31. The spectrum or variance spectral density S_{xk} of the k^{th} frequency component of time series x_n is found by multiplying Equation 32 by its complex conjugate, dividing the result by 2, and dividing again by the frequency bandwidth represented by element k. This frequency bandwidth Δf is defined as

$$\Delta f = \frac{1}{N\Delta t} \tag{40}$$

The set of mathematical operations outlined above yields

$$S_{xk} = \frac{1}{2\Delta f} (A_{xk} - iB_{xk})(A_{xk} + iB_{xk}) \tag{41a}$$

$$= \frac{1}{2\Delta f} (A_{xk}^2 + B_{xk}^2) \tag{41b}$$

$$= \frac{\frac{1}{2}C_{xk}^2}{\Delta f} \tag{41c}$$

$$= \frac{2\Delta t}{N} X_k X_k^* \tag{41d}$$

where the asterisk (*) denotes complex conjugate. In the above derivation, Equation 38 has been used in the step from Equation 41b to Equation 41c. Equation 41c is essentially the definition of spectral density since the variance of a sinusiodal wave of amplitude C_{xk} is $\frac{1}{2}C_{xk}^2$, and the variance density is the variance divided by the frequency bandwidth represented by the single sinusiodal wave. Equation 41c also shows how the k^{th} spectral element relates to the Fourier series representation of the process. Equation 41d comes about by using the right side instead of the left side of Equation 32 in Equation 41a. Equation 41d shows how the spectral density of element k is related to element k of the DFT.

32. Clearly, if one knows the spectral density S_{xk}, the time step Δt, and the number of time steps N in a time series, the resolution

frequency bandwidth can be determined from Equation 40, and wave amplitude C_{xk} can be found by solving Equation 41c to form

$$C_{xk} = (2\ \Delta f\ S_{xk})^{1/2} \tag{42}$$

or, if Equation 40 is used to define Δf ,

$$C_{xk} = \left(\frac{2\ S_{xk}}{N\Delta t}\right)^{1/2} \tag{43}$$

The frequency of element k is the $(k - 1)^{th}$ increment of the resolution frequency bandwidth Δf in the form

$$f_k = (k - 1)\ \Delta f \tag{44}$$

or, if Equation 40 is used to define Δf ,

$$f_k = \frac{(k - 1)}{N\Delta t} \tag{45}$$

The value of the independent variable time at the n^{th} time step of both time series x_n and y_n is

$$t_n = (n - 1)\ \Delta t \tag{46}$$

Finally, the initial phase is a uniform random deviate in the range from 0 to 2π radians. This definition conforms to the assumption of a Gaussian random process for natural ocean waves. Notation for random phase of the k^{th} wave component is

$$\phi_{xk} = 2\pi\ U_{xk}[0,1] \tag{47}$$

where $U_{xk}[0,1]$ represents a uniform random deviate in the range 0 to 1, and, in this example, one associated with time series x_n . A variable of this type can be obtained routinely from a computer-based random number generator.

33. With the definitions of Equations 43, 45, 46, and 47, the contribution to a time series at the n^{th} time step from the k^{th} frequency component is

$$C_{xk} \cos(2\pi f_k t_n + \phi_{xk}) = C_{xk} \cos\left[\frac{2\pi(k - 1)(n - 1)}{N} + \phi_{xk}\right] \tag{48a}$$

$$= \left(\frac{2\ S_{xk}}{N\Delta t}\right)^{1/2} \cos\left\{\frac{2\pi(k - 1)(n - 1)}{N} + 2\pi U_{xk}[0,1]\right\} \tag{48b}$$

Equation 48a is recognized as the k^{th} contribution of the Fourier series of x_n given by Equation 35. The Fourier coefficients A_{xk} and B_{xk} of Equations 36 and 37, respectively, then become

$$A_{xk} = \left(\frac{2\ S_{xk}}{N\Delta t}\right)^{1/2} \cos\{2\pi U_{xk}[0,1]\} \tag{49}$$

$$B_{xk} = \left(\frac{2\ S_{xk}}{N\Delta t}\right)^{1/2} \sin\{2\pi U_{xk}[0,1]\} \tag{50}$$

Using these Fourier coefficients in Equations 32 and solving for X_k yields

$$X_k = \frac{N}{2}\left[\left(\frac{2\ S_{xk}}{N\Delta t}\right)^{1/2} \cos\{2\pi U_{xk}[0,1]\} - i\left(\frac{2\ S_{xk}}{N\Delta t}\right)^{1/2} \sin\{2\pi U_{xk}[0,1]\}\right] \tag{51}$$

as an expression for the k^{th} element of the DFT of time series x_n. Equation 51 can be written in more abbreviated form by combining terms and using Euler notation, which results in

$$X_k = \left(\frac{2\ S_{xk}}{N\Delta t}\right)^{1/2} e^{-i2\pi U_{xk}[0,1]} \tag{52}$$

A similar set of steps can be created for time series y_n. If the k^{th} component of the frequency spectrum of y_n is S_{yk} and its k^{th} initial phase is $2\pi U_{yk}[0,1]$, the corresponding DFT element Y_k is given by

$$Y_k = \left(\frac{2\ S_{yk}}{N\Delta t}\right)^{1/2} e^{-i2\pi U_{yk}[0,1]} \tag{53}$$

34. By taking advantage of complex notation and certain properties of DFT's, an efficient means of computing multiple time series is achieved. If the DFT for time series y_n (Equation 20) is multiplied by i and added to the DFT for time series x_n (Equation 19) and the k^{th} element of the result is denoted as Z_k, then

$$Z_k = X_k + iY_k$$

$$= \sum_{n=1}^{N} (x_n + iy_n)\, e^{-i[2\pi(k-1)(n-1)/N]} \qquad k = 1, 2, \ldots, \frac{N}{2}$$

$$= \frac{N}{2} (A_{xk} - iB_{xk}) + i\frac{N}{2} (A_{yk} - iB_{yk}) \qquad k = 1, 2, \ldots, \frac{N}{2} \tag{54}$$

where Equations 32 and 34 were used to obtain the last line. The complex conjugate of element $N - k + 2$ of Equation 54 has the form

$$Z^*_{N-k+2} = X^*_{N-k+2} - iY^*_{N-k+2}$$

$$= \sum_{n=1}^{N} (x_n - iy_n)\, e^{i[2\pi(N-k+2-1)(n-1)/N]}$$

$$= \sum_{n=1}^{N} (x_n - iy_n)\, e^{-i[2\pi(k-1)(n-1)/N]} \qquad k = 1, 2, \ldots, \frac{N}{2} \tag{55}$$

where use is made of the identity $e^{i[2\pi N(n-1)/N]} = 1$. Again using Equations 32 and 34, Equation 55 can be written

$$Z^*_{N-k+2} = \frac{N}{2} (A_{xk} - iB_{xk}) - i\frac{N}{2} (A_{yk} - iB_{yk}) \qquad k = 1, 2, \ldots, \frac{N}{2} \tag{56}$$

35. Adding Equations 54 and 56 and multiplying by $1/N$ yields an expression for just the Fourier coefficients of time series x_n, i.e.,

$$A_{xk} - iB_{xk} = \frac{1}{N} \left(Z_k + Z^*_{N-k+2} \right) \tag{57}$$

Subtracting Equation 56 from Equation 54 and multiplying by $-i/N$ yields an expression for just the Fourier coefficients of time series y_n in the form

$$A_{yk} - iB_{yk} = \frac{-i}{N} \left(Z_k - Z^*_{N-k+2} \right) \tag{58}$$

Equations 57 and 58 can be inverted to recover Equation 54 representing Z_k for $k = 1, 2, \ldots, N/2$ and the complex conjugate of Equation 56 representing Z_k for $k = N/2 + 2, N/2 + 3, \ldots, N$. The A_{xk} and B_{xk} for time series x_n are assigned values from user-defined spectral densities S_{xk} and a set of random phases $2\pi U_{xk}[0,1]$ following Equations 49 and 50. A similar set of relations defines the A_{yk} and B_{yk} for time series y_n. With these

definitions made, the complex DFT Z_k can be inverse transformed to recover a complex array for which the real part is the time series x_n and the imaginary part is the second time series y_n .

36. In summary, two time series can be synthesized from two distinct spectral density definitions using a single algorithm. Complex numbers are formed from the Fourier coefficients of the time series. For one time series, the complex Fourier coefficients are formed from

$$A_{xk} - iB_{xk} = \left(\frac{2\ S_{xk}}{N\Delta t}\right)^{1/2} e^{-i2\pi U_{xk}[0,1]} \qquad k = 1,\ 2,\ \ldots,\ \frac{N}{2} \tag{59}$$

and for the other time series, the Fourier coefficients are

$$A_{yk} - iB_{yk} = \left(\frac{2\ S_{yk}}{N\Delta t}\right)^{1/2} e^{-i2\pi U_{yk}[0,1]} \qquad k = 1,\ 2,\ \ldots,\ \frac{N}{2} \tag{60}$$

Equations 59 and 60 are then combined to define the DFT of the complex sum of the two time series in the form

$$Z_k = \frac{N}{2}\left[(A_{xk} - iB_{xk}) + i(A_{yk} - iB_{yk})\right] \qquad k = 2,\ 3,\ \ldots,\ \frac{N}{2} \tag{61}$$

and

$$Z_{N-k+2} = \frac{N}{2}\left[(A_{xk} - iB_{xk}) - i(A_{yk} - iB_{yk})\right]^* \qquad k = 2,\ 3,\ \ldots,\ \frac{N}{2} \tag{62}$$

with $Z_1 = Z_{N/2+1} = 0$ arbitrarily assigned without loss of generality in the present application for the mean values and Nyquist frequency values, respectively. Finally, the complex array Z_k $(k = 1,\ 2,\ \ldots,\ N)$ is inverse Fourier transformed to recover the time series through the sum

$$x_n + iy_n = \sum_{k=1}^{N} Z_k\ e^{i[2\pi(k-1)(n-1)/N]} \qquad n = 1,\ 2,\ \ldots,\ N \tag{63}$$

In Equation 63, the real part is the time series x_n and the imaginary part is the time series y_n . Where a collection of such time series (or their properties) is useful for statistical purposes, the above algorithm can be applied repeatedly, with the procedure providing two new time series in each application. Computer techniques known as Fast Fourier Transforms enable sums like those on the right sides of Equations 19, 20, and 63 to be computed very

rapidly. In this way, statistically stable data sets can be synthesized with reasonable efficiency.

Unimodal Data

37. In using synthetic data as described above, some assurance must be gained that the procedure used replicates, in fact, the process it intends to synthesize. In the present case, a narrow-banded process containing a large number of spectral lines should yield a Rayleigh distribution of wave heights. This distribution is the result predicted by Longuet-Higgins (1952). Although ideally a Rayleigh process contains an infinite number of spectral lines, it is a practical impossibility to synthesize such a process using discrete computational techniques. However, it should be possible to approximate the process if a sufficiently large number of lines is used. Thus, an important question in the current context is how many lines it takes for a spectrum of a given width to approximate closely a Rayleigh process.

38. A second question is related to the first. That question is how narrow a spectrum must be before a Rayleigh process is realized. The derivation given by Longuet-Higgins (1952) simply states that the process must have a narrow spectrum, but does not give a practical definition of narrowness.

39. Part of the investigation described here is devoted to examining these two questions. The idea is to find what is necessary to simulate a Rayleigh process in a unimodal spectrum and then combine several of these independently Rayleigh processes to investigate wave height distributions from processes with multimodal spectra. The approach for this part is to generate time series from band-limited white spectra having variable bandwidths and different numbers of spectral lines, with each spectral component being assigned a random initial phase. Due to the random phasing process, a given sample time series can have a highly variable maximum wave height depending on how nearly all the components are in phase at some point in the synthetic time series. To get a better estimate of a characteristic maximum height, as well as other high-wave statistics, an average of the wave height distributions of several time series, having the same generating parameters (except for phasing) and truncated at the same number of waves, is computed. This average is compared with both the Rayleigh model, Equations 10 and 12 for the pdf and exceedence curves, respectively, and the Modified Rayleigh model, Equation 13

for the pdf and its intgral using Equation 4 for the exceedence curve. The fundamental parameters of these time series averages are the defining spectral widths and numbers of components.

40. Mathematically, the spectra are defined by first specifying the number of time steps N in a time series and a time step dt between time steps. By Equation 40, the basic frequency increment of the discrete spectrum is $df = 1/Ndt$ so that the n^{th} frequency of the spectrum is $f_n = (n-1)df$. It then remains to assign variance densities to each of the $N/2 + 1$ frequency bands at or below the Nyquist frequency. For band-limited white spectra, four more variables are needed: a total variance, a reference center frequency, a bandwidth, and a number of spectral lines to which to assign energy. In all of the tests discussed here, the total variance is constrained to be equivalent to $0.25\ m^2$. This yields an $H_{mo} = 2.0\ m$, which is a convenient number of order one for computational purposes, but is otherwise unimportant as all results are normalized with the computed RMS upcrossing wave height H_{rms} .

41. The variables in the problem are then the center frequency f_c , the overall bandwidth Δf , and the number N of spectral lines within Δf which are assigned a finite variance. The bandwidth is normalized by the center frequency to take the form $\Delta f/f_c$ as a parameter. The bandwidth is then found from the product $f_c \cdot \Delta f/f_c$. The low-frequency bound of this band is the discrete raw frequency band nearest to $f_c - \Delta f/2$. The high-frequency bound is the discrete frequency band nearest to $f_c + \Delta f/2$.

42. Assignment of the number of bands within the bounding frequencies is determined by specifying the interval between raw frequency increments to which finite variance is assigned. If every line is of finite variance, the number of lines is $N = \Delta f/df$. If every other line has finite variance (the intervening lines having zero variance), the number of lines is $N = \Delta f/2df$. In this case, the effective overall bandwidth of the spectrum is still the same Δf , but it is represented with only half as many lines, as if the resolution bandwidth is twice as wide. If the variance in every other band of the second case is twice the variance in each band of the first case, the total energy represented by the spectrum remains the same. This same procedure can be followed by assigning finite variance to every third line, every fourth line, and so on until as few as two lines remain to represent the spectrum. Two lines is clearly a lower limit for the number of lines with

which to represent a spectrum because the pdf is no longer even approximately Rayleigh, as derived by Longuet-Higgins (1952), as expressed in Equation 7, and as shown in Figure 1. Figure 2 illustrates the effect of varying the number of lines with which to define a band-limited white spectrum. It also shows the equivalent bandwidths and spectral densities of a "continuous" spectrum having the reduced number of lines and yet retaining the same overall bandwidth.

43. For the unimodal spectra, both the overall bandwidth Δf and the number of lines N within this band are varied within the limits imposed by time series of finite length. In all tests, time series lengths are $N = 65,536$ points. With a time step of $dt = 0.5$ sec , this represents a record of 32,768 sec or 9 hr, 6 min, 8 sec, a record much longer than is normally obtained in nature. In all cases, the center frequency is kept at $f_c = 0.1$ Hz , corresponding to 10-sec waves. Hence, each simulation contains in excess of 3,000 waves, enough with which to compute some reasonably concise statistics.

44. For each simulation, 20 runs were made to establish a mean characteristic pdf. With 20 samples, other statistics can be computed as well. For the present tests, a standard deviation for the exceedence curve is also computed. Since the number of waves is not exactly constant in the full length of the time series, the run results are scanned to find the case with the fewest waves and all other runs are truncated in time at this same number of waves. In this way, all 20 runs for each case have the same number of waves. When wave heights are placed in order of ascending height, there are then 20 samples at exactly the same discrete exceedence probability estimate.

Empirical Probability Estimates

45. A set of N discrete wave heights H_n can be used to estimate the pdf of the governing process. If all H_n are normalized by H_{rms} (defined in Equation 6), the number J_u which fall in the range $u\Delta H/H_{rms}$ to $(u+1)\Delta H/H_{rms}$ can be counted and divided by N to compute the fraction (or estimated probability) of this occurance. When this fraction is divided by ΔH , a wave height range arbitrarily chosen by the investigator, and multiplied by H_{rms} to make the result dimensionless, the result is an estimate of the normalized pdf for that data set.

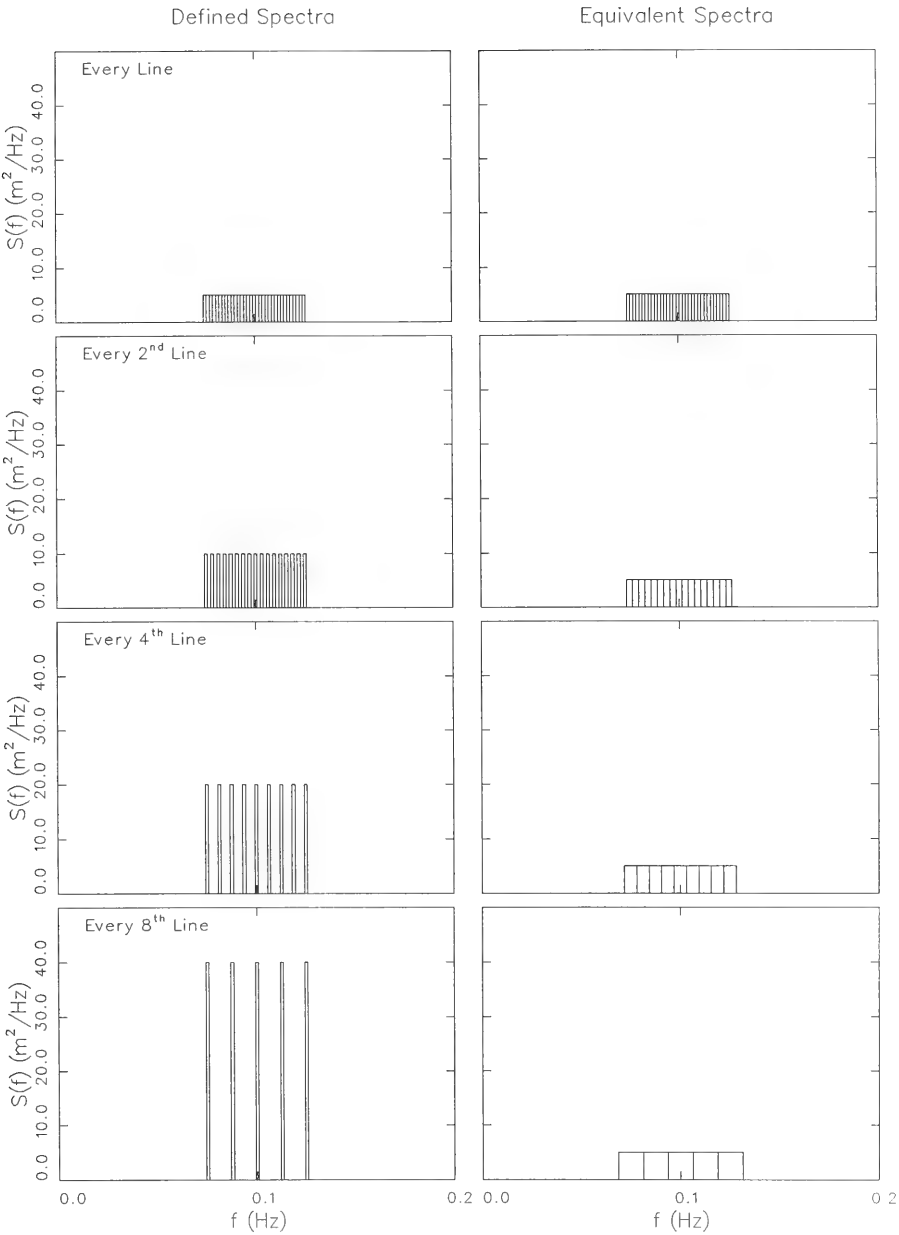


Figure 2. Examples of spectral line decimation

46. If more than one set of wave heights is available, a better estimate of the pdf of the underlying process is found from the average number $\bar{J}_u$ of contributions to each range bin. Since H_{rms} is unlikely to vary much from run to run, the effect is simply to increase the sample population of wave heights. Such an increase gives a better pdf estimate because there are more degrees of freedom for the population of each range bin.

47. The same set of normalized wave heights H_n/H_{rms} can be used to estimate the cumulative or exceedence probability functions. If the H_n/H_{rms} are ordered from smallest to largest, the probability that any of the wave heights is less than the height of wave n is the fraction $n/(N + 1)$. This computation provides an estimate of the cumulative distribution function. The data estimate of exceedence probability $Q_D(H_n/H_{rms})$ is one minus the cumulative distribution at height n, or

$$Q_D\left(\frac{H_n}{H_{rms}}\right) = 1 - \frac{n}{N + 1} \tag{64}$$

48. A third set of statistics that are of use is the average of the highest fraction r of an observed set of waves, denoted $H^{(r)}$. One of the most commonly cited members of this group is the average of the highest one-third of the waves in a given sample. This is given the symbol $H^{(1/3)}$ following the notation of Longuet-Higgins (1952). Note that the SPM (1984) and some other references use the symbol $H_{1/3}$, instead. From the set of normalized, ordered wave heights H_n/H_{rms}, the computation of the j^{th} discrete, monotonically increasing fraction r_j is given by

$$r_j = \frac{j}{N} \tag{65}$$

and the average of that fraction of the highest waves is

$$\frac{H^{(r_j)}}{H_{rms}} = \frac{1}{j} \sum_{i=1}^{j} \frac{H_{N-i+1}}{H_{rms}} \tag{66}$$

The result of Equation 66 can be interpolated if a particular desired fraction does not fall on one of the discrete fractions r_j.

49. For the Rayleigh and Modified Rayleigh models being tested here, the corresponding average heights are found following the derivation given by

Longuet-Higgins (1952). The fraction r of waves that are larger than a given value of H/H_{rms} is the exceedence probability for that H/H_{rms}. The average value of waves higher than a given H/H_{rms} is the integral from that H/H_{rms} to infinity of the normalized heights weighted by the pdf and divided by the area over the same limits under the weighting function (i.e., the pdf). The general forms for r and $H^{(r)}/H_{rms}$ are

$$r = \int_{H/H_{rms}}^{\infty} \left[H_{rms}\ p\left(\frac{H}{H_{rms}}\right)\right]\ d\left(\frac{H}{H_{rms}}\right) \tag{67}$$

$$\frac{H^{(r)}}{H_{rms}} = \frac{1}{r} \int_{H/H_{rms}}^{\infty} \left(\frac{H}{H_{rms}}\right) \left[H_{rms}\ p\left(\frac{H}{H_{rms}}\right)\right]\ d\left(\frac{H}{H_{rms}}\right) \tag{68}$$

The terms in square brackets are the dimensionless pdf models under consideration (Equation 10 for the Rayleigh model or Equation 13 for the Modified Rayleigh model). Where the above integrals have no simple analytic form, numerical methods are used for evaluation. The simple trapezoid rule with increments of $\Delta H/H_{rms} = 0.001$ and an upper limit of integration of 6.0 (in place of infinity) has been found to yield quite accurate results.

Bimodal Test Data Generation

50. The above definitions are used to establish criteria for simulating a Rayleigh process using discrete inverse Fourier transform techniques to produce time series from unimodal, band-limited, white spectra having given bandwidths and numbers of component waves. Once such criteria are established, a bimodal spectrum can be constructed of two unimodal spectra, each of which yields a Rayleigh wave height distribution by itself. Wave height distributions derived from such bimodal spectra can then be compared with test models to determine the quality of representation.

51. Since each mode in such a bimodal spectrum yields a Rayleigh wave height distribution by itself, the details of the mode structure (bandwidth and number of component wave trains) are no longer important. The two primary variables in bimodal tests are a measure of modal separation in the frequency domain and the relative amount of energy in one mode compared with the other. Modal separation is important because the two modes could be very close to each other and so just yield a slightly wider unimodal spectrum (and

corresponding wave height distribution). If the modes are well separated, low- and high-frequency waves are present at the same time. Under these conditions, one would expect a rather diverse character to a wave height distribution.

52. However, if the energy (i.e., variance) in one mode is much less than in the other, the high-energy mode would be expected to dominate, and the process should become asymptotically Rayleigh. Hence, the ratio of energy in one mode relative to that in the other is important. It seems clear that the greatest deviations would occur when the two modes are of approximately equal energy so that neither clearly dominates.

53. To characterize modal separation, a dimensionless difference between the modal center frequencies is used. If $f_{c,1}$ is the center frequency of the first, lower frequency mode and $f_{c,2}$ is the center frequency of the other mode, a separation parameter can be defined as

$$\text{modal separation} = \frac{f_{c,2} - f_{c,1}}{\frac{1}{2}(f_{c,2} + f_{c,1})} \tag{69}$$

It can be as small as zero (if the two modal center frequencies are co-located) or as large as 2 (if $f_{c,1} \ll f_{c,2}$).

54. To characterize relative energy, the ratio of the variance in the second mode to the variance in the first mode is used. Since four times the square root of the variance in a spectral mode can be identified as an H_{mo} for that mode, the square of the ratio $H_{mo,2}/H_{mo,1}$ can be used to characterize relative energy, i.e.,

$$\text{relative energy} = \left(\frac{H_{mo,2}}{H_{mo,1}}\right)^2 \tag{70}$$

Relative energy can vary from zero to infinity, but for the practical cases considered here, the range is from 0.25 to 4.0.

PART IV: UNIMODAL TESTS

Test Conditions

55. A total of 53 cases were examined for the unimodal tests. Table 1 lists the parameters used in the tests. The primary variables were the bandwidth Δf and the number of spectral lines N. In all cases, the center frequency was $f_c = 0.1$ Hz, the characteristic spectrum-based wave height was $H_{mo} = 2.0$ m, and the number of runs for obtaining statistical averages was 20. Note that spectral component wave amplitudes were found by dividing total signal variance (equal to the square of $\frac{1}{4}H_{mo}$) evenly among all N components. Note also that there is an upper limit to the number of component waves for a given spectral bandwidth because of the finite raw bandwidth imposed by the discrete Fourier technique used to create the time series. Hence, the narrowest case ($\Delta f/f_c = 0.05$) was limited to 164 bands, but the broadest case ($\Delta f/f_c = 1.6$) had 5,243 bands.

56. Test criteria were the percentage differences of $H^{(1/100)}$, $H^{(1/10)}$, $H^{(1/3)}$, and $H^{(1/1)}$ from the synthetic results with those from the test models. Computations followed the pattern shown here for the comparison of synthetic data with the Rayleigh model estimate of the average of the highest 1 percent of waves

$$\% \text{ Difference} = \frac{H_D^{(1/100)} - H_R^{(1/100)}}{H_R^{(1/100)}} \times 100\% \tag{71}$$

Comparison with the Modified Rayleigh model entails replacing $H_R^{(1/100)}$ with $H_{MR}^{(100)}$ in Equation 71. Replacement of the highest 1-percent averages with the highest 10-percent averages $H_D^{(1/10)}$, $H_R^{(1/10)}$, and $H_{MR}^{(1/10)}$ allows comparison among data and models for this statistic, and so on.

Test Results

57. Results of such comparisons are shown in Figures 3 and 4. Figure 3 shows synthetic data results compared with Rayleigh model statistics for the four wave height averages mentioned above. Percentage differences are shown as displacements on ordinate axes. Abscissa displacements are the numbers of component waves N. Symbols denote synthetic data of constant $\Delta f/f_c$. The

Table 1

Parameters* of Unimodal Test Data

Group 1: $\Delta f/f_c = 0.05$		Group 2: $\Delta f/f_c = 0.10$		Group 3: $\Delta f/f_c = 0.20$	
Case #	# of Lines	Case #	# of Lines	Case #	# of Lines
38	164	1	328	11	655
39	82	2	164	12	328
40	41	3	82	13	164
41	20	4	41	14	82
42	10	5	20	15	41
43	5	6	11	16	20
44	3	7	10	17	10
		8	5	18	5
		9	3	19	3
		10	2		

Group 4: $\Delta f/f_c = 0.40$		Group 5: $\Delta f/f_c = 0.80$		Group 6: $\Delta f/f_c = 1.60$	
Case #	# of Lines	Case #	# of Lines	Case #	# of Lines
20	1311	29	2621	45	5243
21	655	30	1311	46	2621
22	328	31	655	47	1311
23	164	32	328	48	655
24	82	33	164	49	328
25	41	34	82	50	164
26	20	35	41	51	82
27	10	36	20	52	41
28	5	37	10	53	20

* For all cases: $f_c = 0.1$ Hz , $H_{mo} = 2.0$ m , 20 runs averaged for each case

inset in the upper part of Figure 3 shows the correspondence between $\Delta f/f_c$ and the spectral width parameter ϵ of Cartwright and Longuet-Higgins (1956), defined in Equation 18. Figure 4 shows the same pattern of comparison but for the Modified Rayleigh model vice the Rayleigh model.

58. Figure 3 reveals some interesting characteristics of synthetic data. Perhaps the most obvious is that the smaller the fraction r in the average $H^{(r)}$, the more wave components are required to approximate a Rayleigh distribution. For the average heights $H^{(1/1)}$, the curves are relatively flat for all numbers of component lines. The same is true for $H^{(1/3)}$. For $H^{(1/10)}$, it appears that about 20 component waves are necessary to differ from a Rayleigh $H^{(1/10)}$ by less than 10 percent. For $H^{(1/100)}$, it

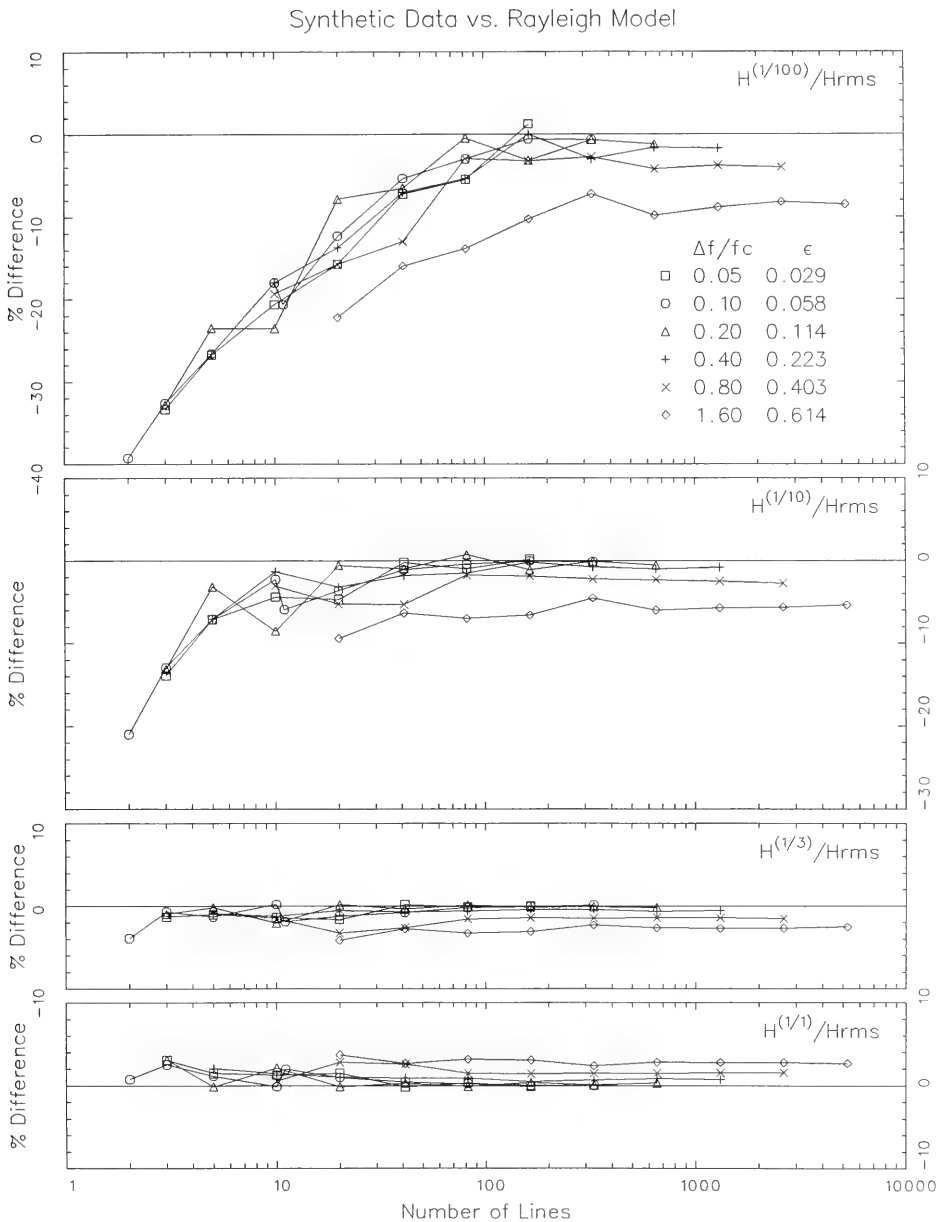


Figure 3. Synthetic wave height averages compared with Rayleigh averages

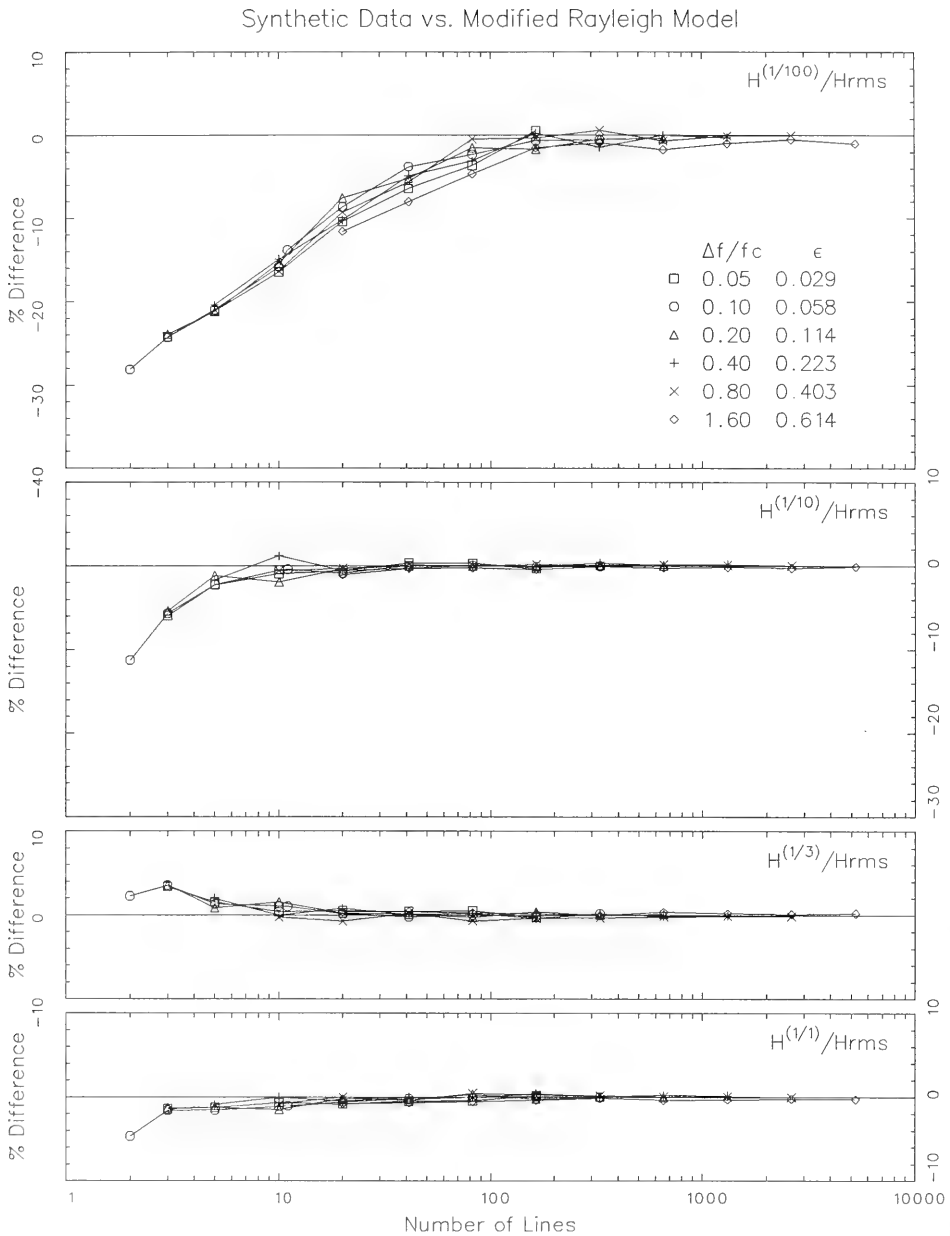


Figure 4. Synthetic wave height averages compared with Modified Rayleigh averages

takes about 100 component waves to stay within 10 percent of a Rayleigh $H^{(1/100)}$.

59. The reason for this behavior can be seen in the sequence of Figures 5, 6, 7, and 8, which show the synthetic, Rayleigh, and Modified Rayleigh probability density and exceedence functions for the cases with $f_c = 0.1$ Hz , $\Delta f/f_c = 0.1$, and $N = 2, 3, 10$, and 328 component waves, respectively. Because the synthetic curves are averages of 20 different runs, standard deviations could also be computed and plotted. Standard deviations are shown as dashed lines in the exceedence graphs. Figure 5 represents the case of two waves. For this case, the synthetic pdf in the upper part of Figure 5 looks very much like the two-wave model pdf shown in Figure 1. Also, the standard deviation is very small. This result is as it should be since the two-wave case has an exact solution (Equation 7) and the synthetic results reflect this solution very well.

60. When a third wave train is added, the problem becomes more complicated, having no analytic solution. The result for a single run depends strongly on the initial phases of the three component waves. If all three are nearly in phase at some point in the time series, the maximum wave heights will be larger than if the waves are not nearly in phase at any point in the time series. Since initial phases vary from run to run, the variation in wave height distribution is suggested by the standard deviation of the set of runs. This is shown in Figure 6 as the dashed lines on either side of the average exceedence curve in the lower part of the figure. The mean of the 20 runs is higher, however, than the mean of the two-wave results of Figure 5. This trend continues when the number of component waves increases to 10, as shown in Figure 7. Figure 7 also shows that the smaller waves begin to conform to the Rayleigh model more than the two-wave or three-wave cases. This is the result summarized in Figure 3. At the point where there are 10 component waves in Figure 3, averages over the highest 1/10, 1/3, and all the waves are very close to the Rayleigh values. However, the $H^{(1/100)}$ still differs substantially from the Rayleigh curve, evidently due to the limited number of wave components.

61. In Figure 8, the synthetic data are composed of 328 wave trains and appears to conform to the Rayleigh curve over the whole plotted domain. Presumably, if a longer time series had been generated, there would be a region of limiting heights at some point on the high-wave tail. This effect

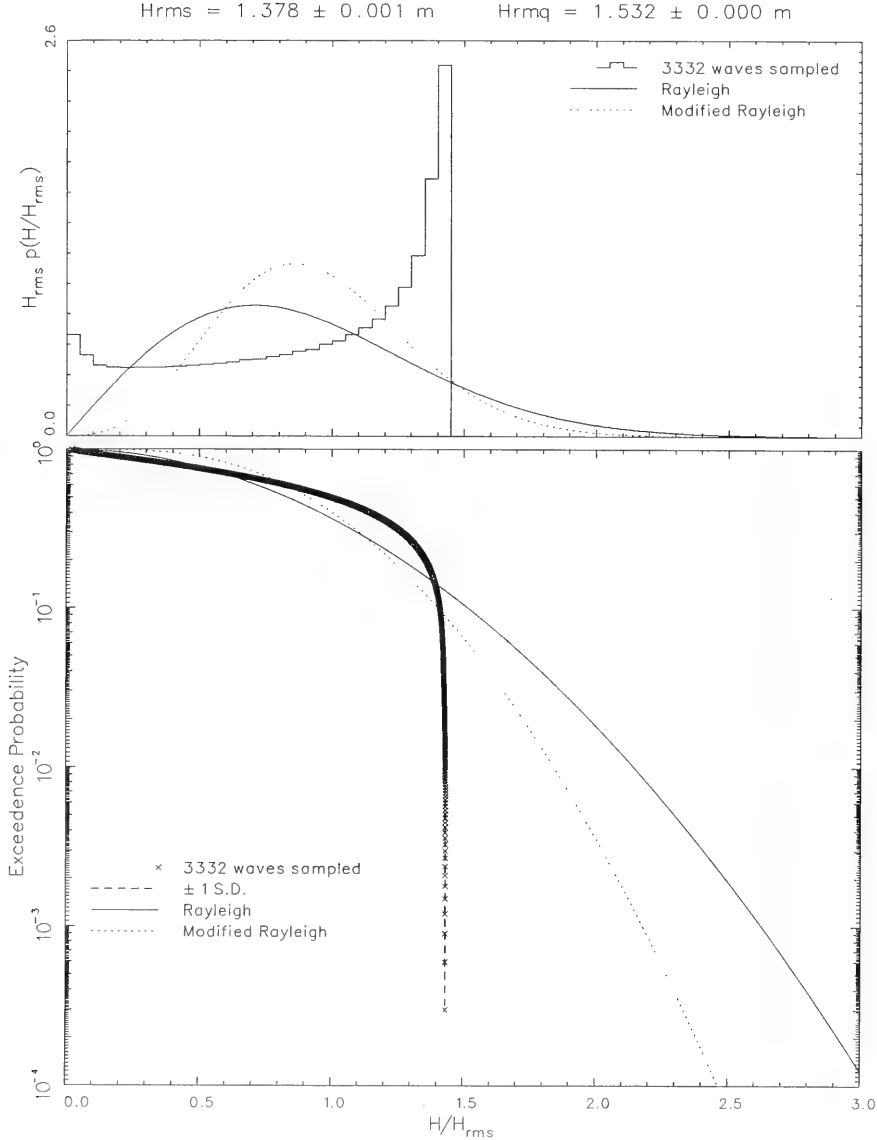


Figure 5. Probability and exceedence curves for two wave trains

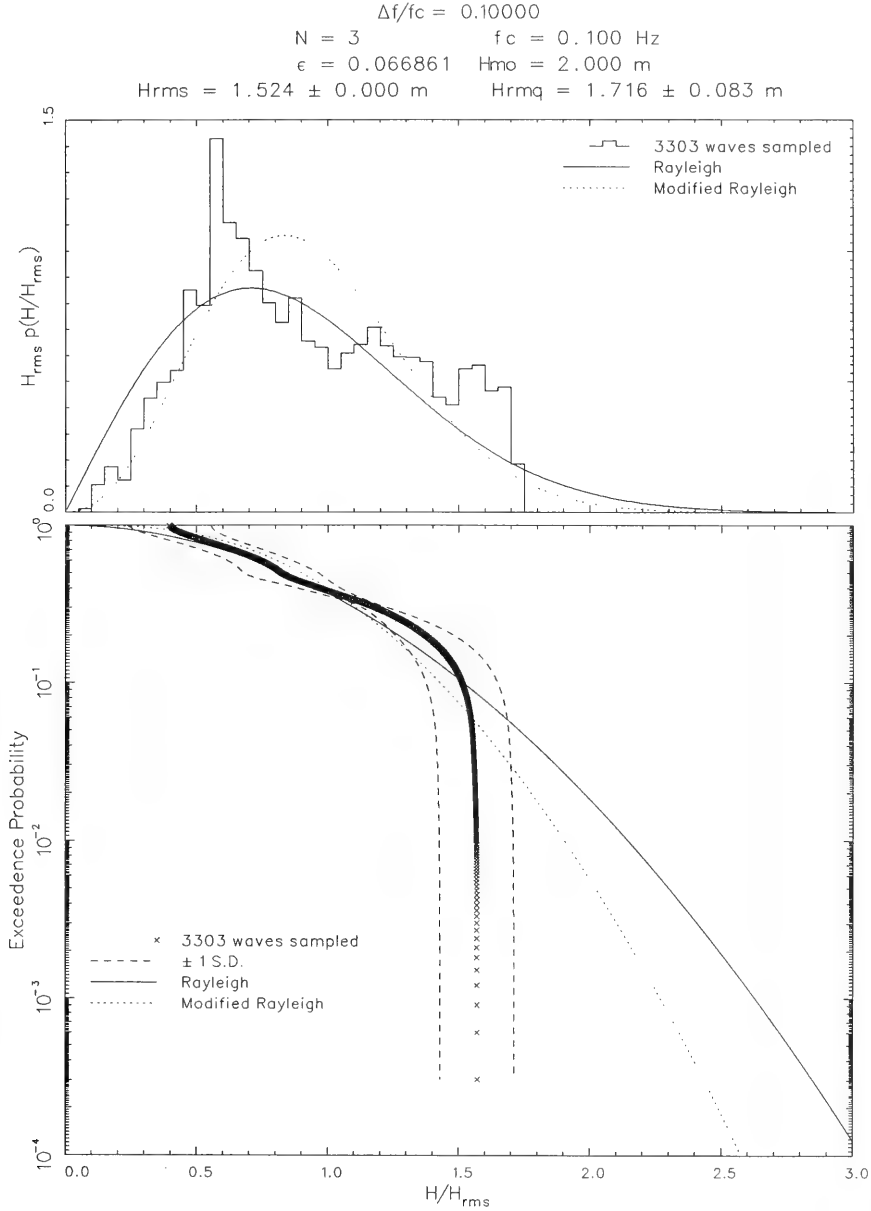


Figure 6. Probability and exceedence curves for three wave trains

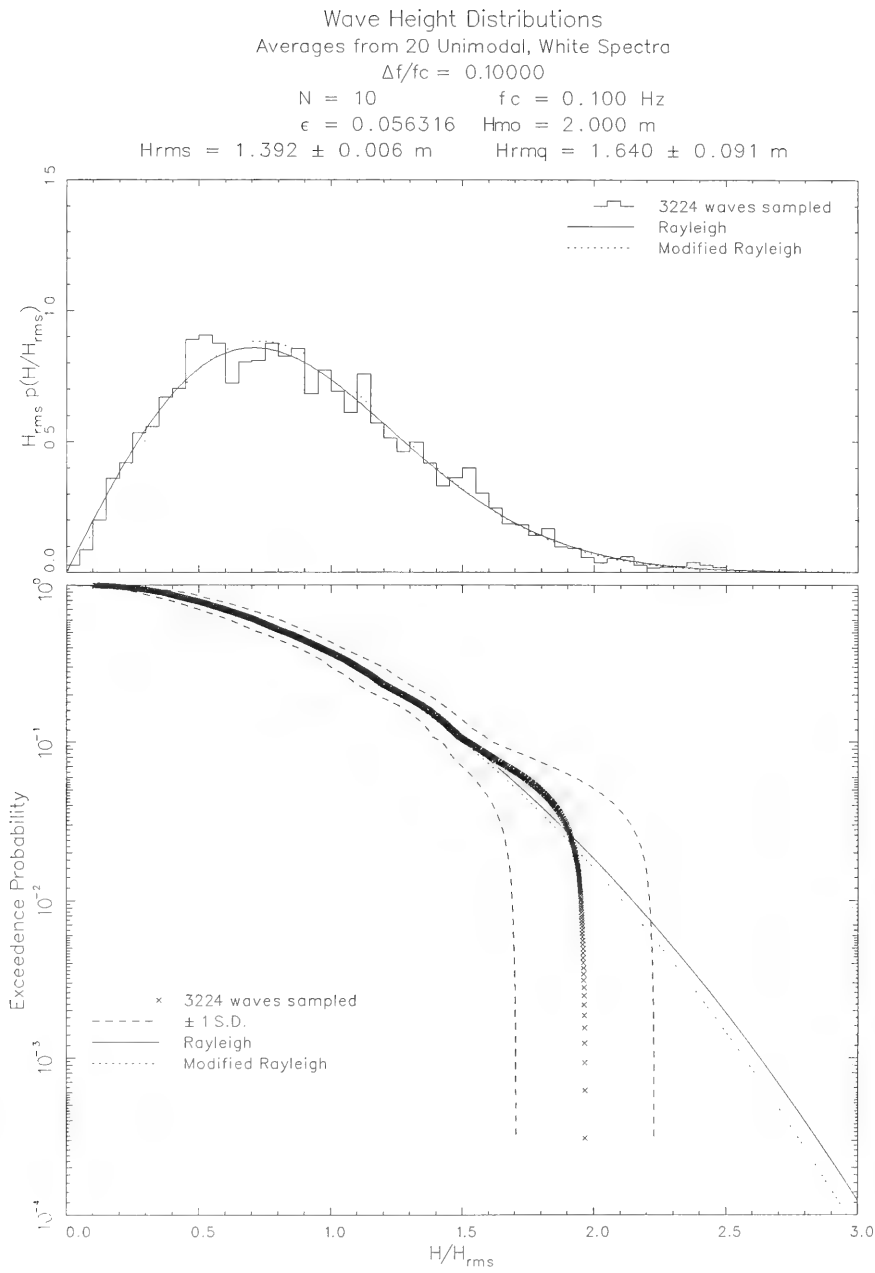


Figure 7. Probability and exceedence curves for 10 wave trains

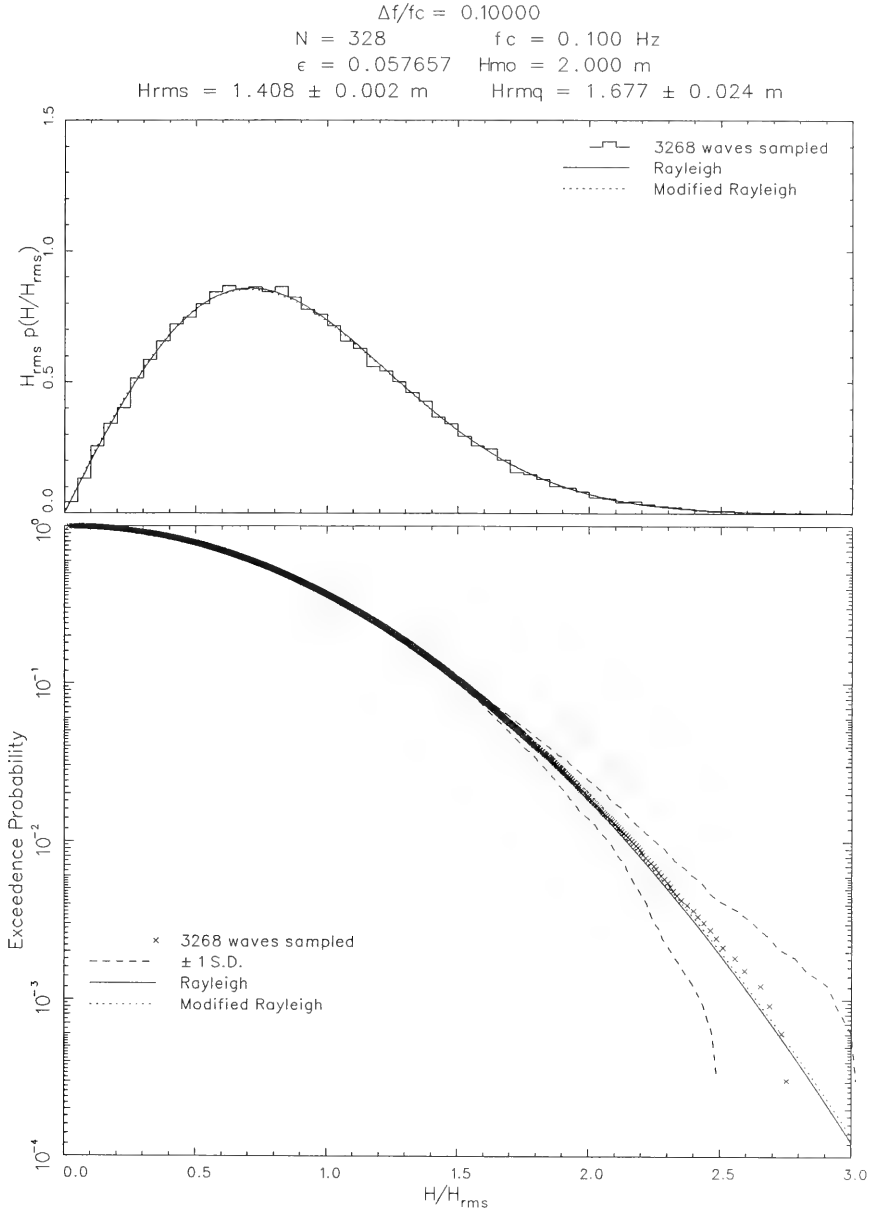


Figure 8. Probability and exceedence curves for 328 wave trains

is suggested in Figure 8 by the widely spaced and nearly vertical standard deviation curves at high H/H_{rms} and by extrapolation of the results shown in Figures 5, 6, and 7. However, this has not affected adversely the statistics chosen for the present tests.

62. A second observation of the results shown in Figure 3 is that beyond a certain bulk spectral bandwidth, the parameters do not appear to approach Rayleigh parameters no matter how many component waves are used in the spectral definition. This is most clear for the widest case, $\Delta f/f_c = 1.60$, but is also slightly evident in the next-to-widest case $\Delta f/f_c = 0.80$. It is notable that even the mean wave height $H^{(1/1)}$ and the average of the highest one-third waves $H^{(1/3)}$ also show some effect of spectral broadening. One of the most extreme examples of this is shown in Figure 9. Though there are over 5,000 component wave trains, the mean curve deviates everywhere from the Rayleigh curve and the deviation is significant at least at the level of the standard deviation of the synthetic data.

63. Also shown in Figure 9 are the Modified Rayleigh pdf and exceedence curves. These curves clearly represent the synthetic data better than the Rayleigh curves. This is evident in the broader sense in Figure 4, where all four measures of comparison are much closer than the Rayleigh comparison of Figure 3 as long as enough component wave trains are present. This result suggests that the Modified Rayleigh curve is a better model for broad spectra if the added parameter H_{rmq} can be determined directly from the spectra. It is not yet evident that this determination can be made.

64. These tests reveal certain properties about unimodal, band-limited, white spectra and the probable wave height distributions contained in the corresponding time series. If the spectra contain in excess of 100 conponent wave trains and do not have a width such that $\Delta f/f_c$ is greater than about 0.4, the resulting wave height distributions will conform very nearly to the Rayleigh model.

Implications for Spectral Filtering

65. A common way to remove unwanted noise or extraneous signal from a time series is by spectral filtering. In this method, a time series is Fourier transformed to obtain a spectrum after which energy in specified frequency bands is either reduced or eliminated according to the filter

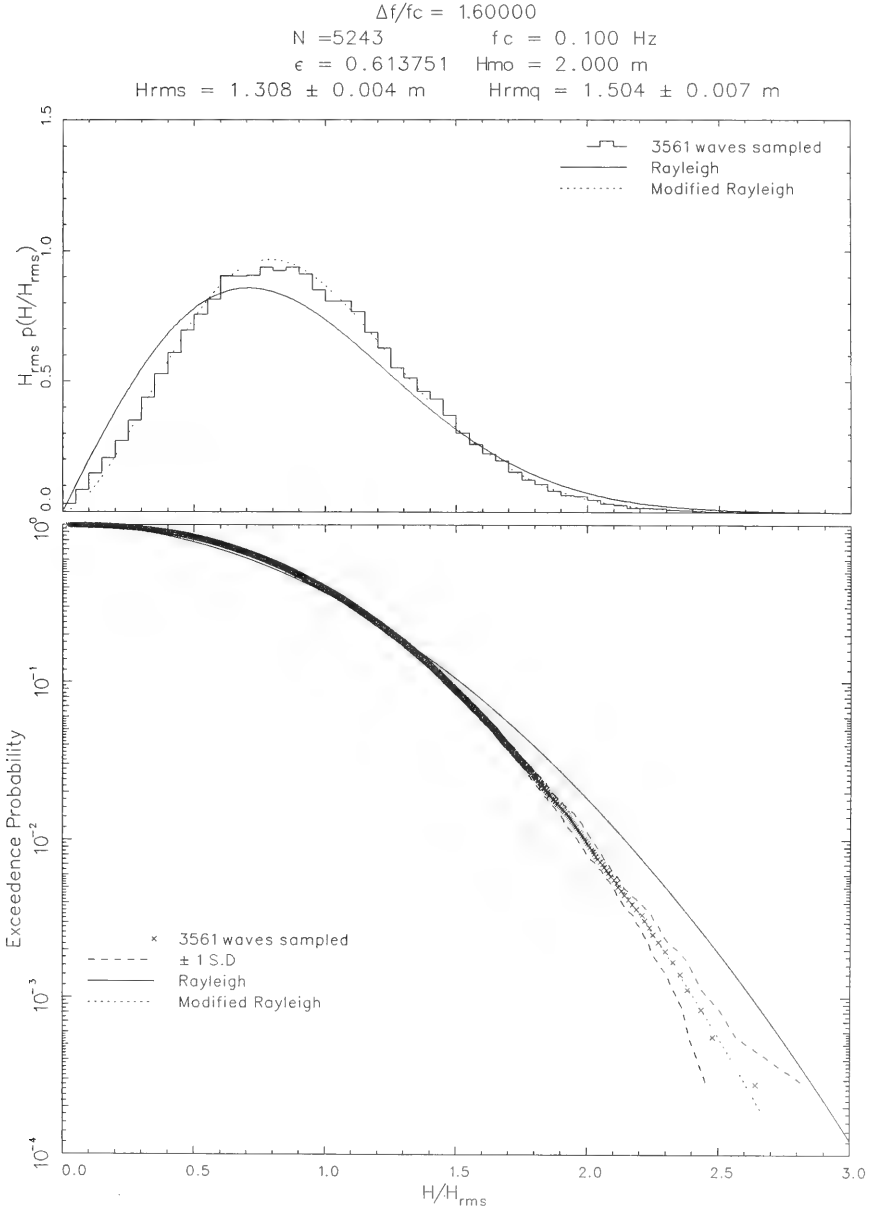


Figure 9. Probability and exceedence curves from a broad spectrum

characteristics. The filtered spectrum is then inverse Fourier transformed to generate a filtered time series.

66. A clear danger in this method is that the possibility arises to remove enough of the spectrum that insufficient energy-containing spectral lines remain to recover enough of the wave height distribution to have reasonably good statistics. This is especially true where the parent time series is short so that the raw spectral frequency bands are wide. Figure 3 indicates that, for spectra that are not too broad to begin with, at least 40 lines must remain to obtain $H^{(1/10)}$ to within about 2 percent and at least 150 lines must remain to obtain $H^{(1/100)}$ to within about 5 percent.

67. There may be advantages to spectral filtering as well. If there is truly noise in the time series such that the spectrum becomes broad in its presence, then filtering can remove the noise and make the spectrum narrower. If such a filter reduces spectral width from an equivalent $\Delta f/f_c$ of order 1.0 to something less than about 0.5, a clear improvement in all statistics is suggested by Figure 3.

PART V: BIMODAL TESTS

Test Conditions

68. A total of 55 cases have been examined for the bimodal tests. Table 2 lists the parameters used in the tests. The primary variables are modal separation, as defined by Equation 69, and relative modal energy, as defined by Equation 70. Each mode has been constrained to a width of 0.01 Hz by adjusting the parameter $\Delta f/f_c$ along with the parameter f_c so that their product yields the desired bandwidth. With a time step $\Delta t = 0.5$ sec and the time series length of 65,536 points, the raw spectral bandwidth is about 0.0000305 Hz. With this raw bandwidth, a modal bandwidth of 0.01 Hz contains about 328 spectral lines. Each mode is independently a Rayleigh process, as determined from the results of Part IV.

69. Modal separation has been allowed to vary from zero to about 1.33. At zero separation, the two modes coincide on the frequency axis. In this case, the resulting process should be Rayleigh for all values of relative modal energy because the effective spectrum is unimodal and narrow, and contains enough spectral lines. It is no longer a white spectrum, however, because there are now two randomly phased waves at each frequency. Two waves at the same frequency can interfere constructively or destructively, depending on their relative phases, so that the resulting wave amplitude can vary from zero (destructive interference with waves of equal amplitude) to the sum of the two component wave amplitudes (constructive interference with the waves in phase). Corresponding wave energy will vary from zero to a number proportional to the square of the sum of the two component wave amplitudes. In the derivation by Longuet-Higgins (1952), there was no constraint on the distribution of energy within the narrow band process. Hence, coincident modes in the test spectra are expected to yield wave height distributions that follow the Rayleigh model.

70. The largest modal separation has one mode center frequency at 0.05 Hz and the other at 0.25 Hz. These frequencies are near the limits of the overall band normally associated with wind waves, corresponding to wave periods of 20 and 4 sec, respectively. This separation is considered to be a practical limit for wind waves. Such a case might occur where young or short-fetch waves are riding on longer, swell-like waves.

Table 2

Parameters* of Bimodal Test Data

Case #	Mode 1 f_c (Hz)	Mode 1 $\Delta f/f_c$	Mode 2 f_c (Hz)	Mode 2 $\Delta f/f_c$
122	0.100	0.100	0.100	0.1000
123	0.095	0.105	0.105	0.0950
124	0.090	0.111	0.110	0.0909
125	0.085	0.118	0.115	0.0870
126	0.080	0.125	0.120	0.0833
127	0.070	0.143	0.130	0.0769
128	0.060	0.167	0.140	0.0714
129	0.050	0.200	0.150	0.0667
130	0.050	0.200	0.160	0.0625
131	0.500	0.200	0.200	0.0500
132	0.500	0.200	0.250	0.0400
	$H_{mo,1}$ = 0.894 m		$H_{mo,2}$ = 1.789 m	
100	0.100	0.100	0.100	0.1000
101	0.095	0.105	0.105	0.0950
102	0.090	0.111	0.110	0.0909
103	0.085	0.118	0.115	0.0870
104	0.080	0.125	0.120	0.0833
105	0.070	0.143	0.130	0.0769
106	0.060	0.167	0.140	0.0714
107	0.050	0.200	0.150	0.0667
108	0.050	0.200	0.160	0.0625
109	0.500	0.200	0.200	0.0500
110	0.500	0.200	0.250	0.0400
	$H_{mo,1}$ = 1.414 m		$H_{mo,2}$ = 1.414 m	
111	0.100	0.100	0.100	0.1000
112	0.095	0.105	0.105	0.0950
113	0.090	0.111	0.110	0.0909
114	0.085	0.118	0.115	0.0870
115	0.080	0.125	0.120	0.0833
116	0.070	0.143	0.130	0.0769
117	0.060	0.167	0.140	0.0714
118	0.050	0.200	0.150	0.0667
119	0.050	0.200	0.160	0.0625
120	0.500	0.200	0.200	0.0500
121	0.500	0.200	0.250	0.0400
	$H_{mo,1}$ = 1.789 m		$H_{mo,2}$ = 0.894 m	
144	0.100	0.100	0.100	0.1000
145	0.095	0.105	0.105	0.0950
146	0.090	0.111	0.110	0.0909
147	0.085	0.118	0.115	0.0870
148	0.080	0.125	0.120	0.0833
149	0.070	0.143	0.130	0.0769
150	0.060	0.167	0.140	0.0714
151	0.050	0.200	0.150	0.0667
152	0.050	0.200	0.160	0.0625
153	0.050	0.200	0.200	0.0500
154	0.050	0.200	0.250	0.0400
	$H_{mo,1}$ = 0.667 m		$H_{mo,2}$ = 0.943 m	
133	0.100	0.100	0.100	0.1000
134	0.095	0.105	0.105	0.0950
135	0.090	0.111	0.110	0.0909
136	0.085	0.118	0.115	0.0870
137	0.080	0.125	0.120	0.0833
138	0.070	0.143	0.130	0.0769
139	0.060	0.167	0.140	0.0714
140	0.050	0.200	0.150	0.0667
141	0.050	0.200	0.160	0.0625
142	0.050	0.200	0.200	0.0500
143	0.050	0.200	0.250	0.0400
	$H_{mo,1}$ = 0.943 m		$H_{mo,2}$ = 0.667 m	

* 20 runs averaged for each case.

71. Variation of relative amounts of energy so that one or the other of the modes dominates suggests what happens to wave height distributions in such situations as high-energy short waves on low-level swell or the initial stages of a growing sea in the presence of an active swell. Five levels of relative energy have been used in the present tests. The ratio varies from 0.25 to 4.0 following a geometric progression. Clearly, if the ratio approaches zero or infinity, then one or the other of the two modes will become the only energy containing mode so that the process will revert to unimodal behavior. It is expected that the most interesting cases will be those with energy ratios of order unity.

72. Test criteria in this set of tests are the same as those in the unimodal tests. Wave height averages $H^{(1/100)}$, $H^{(1/10)}$, $H^{(1/3)}$, and $H^{(1/1)}$

from synthetic data are compared with Rayleigh and Modified Rayleigh model predictions of these statistics. Percentage differences are then computed following the pattern of Equation 71, and comparisons are done among these percentages.

Test Results

73. Figures 10 and 11 summarize the results of the bimodal tests. Figure 10 shows percentage differences of synthetic data from the Rayleigh model. Figure 11 shows synthetic data compared with the Modified Rayleigh model. In both figures, peak separation (defined by Equation 69) is along the abscissa. Symbols denote points of constant relative energy ratio (defined by Equation 70).

74. In Figure 10, the synthetic data clearly approximate the Rayleigh model for small modal separations, as expected. Synthetic data begin to deviate from the Rayleigh model when the peak separation parameter reaches about 0.40, especially for averages on the high-wave tails of the distributions, e.g., $H^{(1/100)}/H_{rms}$ and $H^{(1/10)}/H_{rms}$. Curves for all values of the energy ratio behave similarly up to peak separations of about 0.60, and then diverge radically depending on the energy ratio. At the largest peak separation, and for a small energy ratio, i.e., where the energy in the high-frequency mode is small relative to that in the low-frequency mode, the average of the highest 1 percent of the waves exceeds the Rayleigh prediction by almost 20 percent. This result suggests a rather severe problem in structural design if the Rayleigh model is used for estimating extreme waves, since it will be low by 20 percent under these conditions. For intermediate and large energy ratios and for large peak separations, the statistics of the higher wave averages are less, by up to 7 percent, than the Rayleigh model prediction.

75. The reasons for this severe dichotomy in the extreme average wave heights at large modal separations are not immediately obvious. Part of the explanation is how H_{rms} behaves relative to the absolute average heights in the various test conditions. If waves in the high-frequency mode are as energetic or are more energetic than those in the lower frequency mode (energy ratio ≥ 1), the effect is simply to broaden the spectrum. High-frequency waves are carried by lower frequency waves of comparable energy, and the

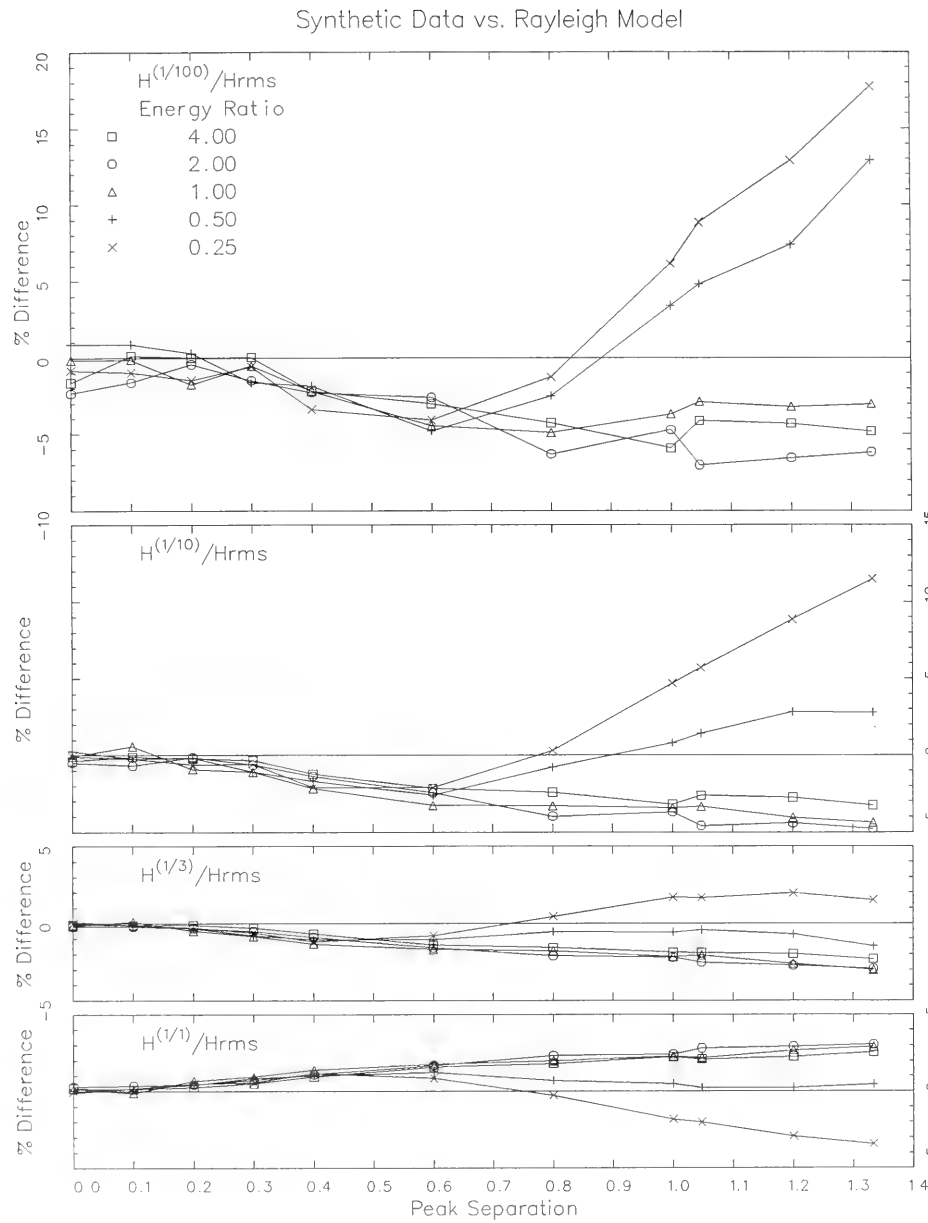


Figure 10. Wave height averages from bimodal spectra compared with Rayleigh wave height averages

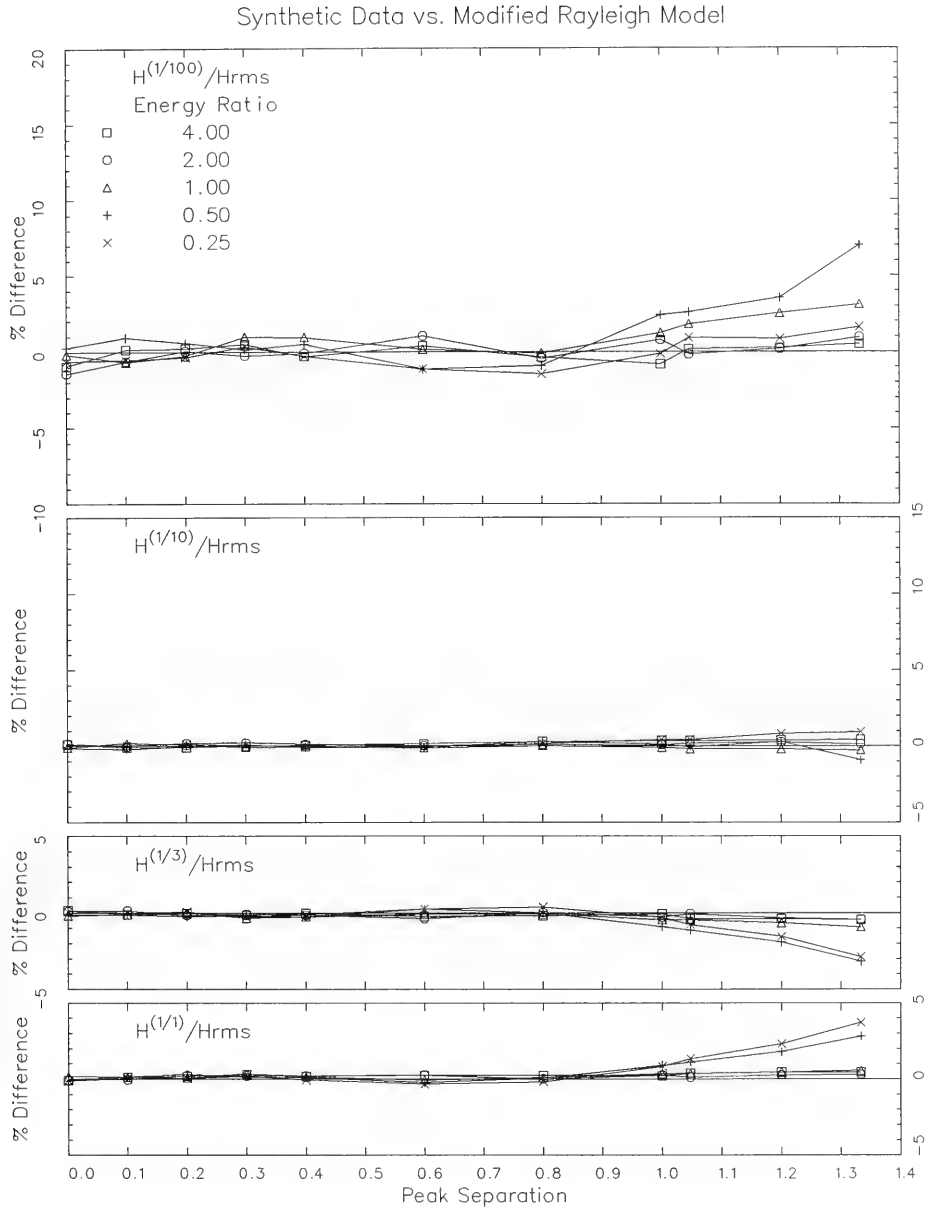


Figure 11. Wave height averages from bimodal spectra compared with Modified Rayleigh wave height averages

effect on average heights is similar to results from the broad, unimodal cases of Part IV. In fact, the cases in Figure 10 with energy ratios of 1.0, 2.0, and 4.0 have properties very much like the two broadest unimodal cases ($\Delta f/f_c$ = 0.80 and 1.60) in Figure 3. Mean wave heights $H^{(1/1)}$ are higher than Rayleigh predictions by 2 to 3 percent; the $H^{(1/3)}$ are smaller than Rayleigh by 2 to 3 percent; the $H^{(1/10)}$ are smaller by 3 to 5 percent; and the $H^{(1/100)}$ are smaller by 3 to 10 percent. The evident conclusion is that from some bimodal spectra with intermediate and large peak separation, the effect on wave height averages is the same as for broad, unimodal spectra.

76. For bimodal spectra with low energy ratios (energy in the high-frequency mode somewhat smaller than that in the low-frequency mode), average wave height behavior is very different at large peak separations. Here, again, short waves are being carried on longer waves, but now the short waves tend to have less amplitude than the long waves. This situation leads to the condition that at the crests and troughs of the long waves, there are a number of short waves whose extrema do not cross the line of zero displacement. There tend to be higher crests because short wave crests add to long wave crests, and deeper troughs because short wave troughs drop below long wave troughs, but fewer waves overall because the short waves do not have as many zero crossings as when they have higher energy. In regions of the time series near nodes (zero-crossings) of the low-frequency waves, the number of zero crossings is about the same as for higher levels of high-frequency energy. Hence, there is a relative reduction in the number of waves of intermediate height, those which would occur if the high-frequency waves reached zero level from extrema of the low-frequency waves. There becomes a relative abundance of both large and small waves, at the expense of the number of waves of intermediate height.

77. This redistribution of wave heights also affects the value of H_{rms}, a parameter used to normalize all distributions and parameters in this study. The net results of this alteration of H_{rms} (relative to H_{mo}, the parameter used to govern total spectral energy and to generate the time series) and the redistribution of wave heights are shown as the normalized wave height averages in Figure 10. The higher wave height averages are very much higher than the Rayleigh prediction. For the lower averages, $H^{(1/3)}$ hovers near the Rayleigh prediction, and the average wave height $H^{(1/1)}$ tends to be less than the Rayleigh value.

78. A remarkable improvement over the results shown in Figure 10 is the comparison of the same synthetic data with the Modified Rayliegh model as shown in Figure 11. This improved comparison could be expected for two reasons. One reason is that the Modified Rayleigh model is a two-parameter model and is therefore much more highly adaptable than the Rayleigh model. The second reason is that the effective broadening of the spectra by medium-to-large modal separations and medium-to-large energy ratios is rather like the broad unimodal cases for which the Modified Rayleigh model worked exceptionally well, as shown in Figure 4.

79. However, bimodal cases with large modal separations and small energy ratios were not entirely well-fitted by the Modified Rayleigh model. There is some improvement over the Rayleigh model as can be seen by comparing Figures 10 and 11 (they are plotted at the same scale), but this improvement is likely due to the improved adaptability of the Modified Rayleigh model rather than any inherent ability to represent the underlying process. This result suggests that for wide modal separations and small energy ratios, there is another, as yet undefined, distribution which better represents these processes.

80. Some examples will clarify the above arguments. Figures 12, 13, 14, and 15 show pdf and exceedence curves for time series derived from bimodal spectra with small, intermediate, and two cases of large modal separation. Figure 12 represents a case where modal separation is small (= 0.10) and energy ratio is also small (= 0.25). In this case, the synthetic data lie very close to the Rayleigh curve, and the Modified Rayleigh model is very near the Rayleigh model asymptotic shape. Hence, under these conditions, the process is well-modeled by the Rayleigh pdf. Figure 13 represents a case with the same energy ratio as in Figure 12, but with an intermediate modal separation (= 0.60). The effective spectral broadening is apparent, especially in the exceedence curves. Synthetic data and Modified Rayleigh results agree very closely, and both differ from the Rayleigh model on the high-wave tail of the distribution. This behavior is very much like that shown in Figure 9, representing broad-banded, unimodal spectra.

81. Figures 14 and 15 depict results from cases where the governing bimodal spectra have the largest separation parameter (= 1.33), but extremes of relative energy parameter (= 4.0 and 0.25, respectively). Figure 14 represents a case that follows the pattern of large relative modal energy.

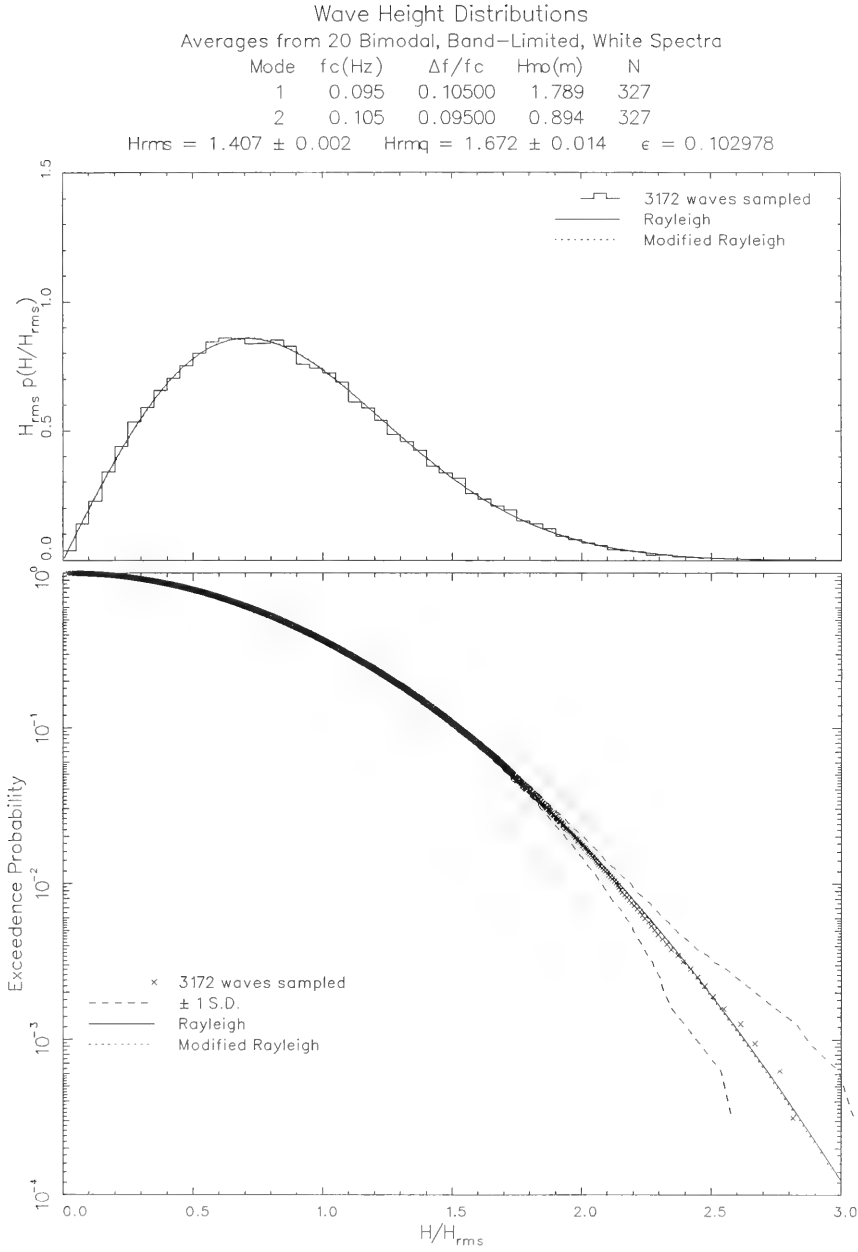


Figure 12. Probability and exceedence curves from bimodal spectra with small modal separation

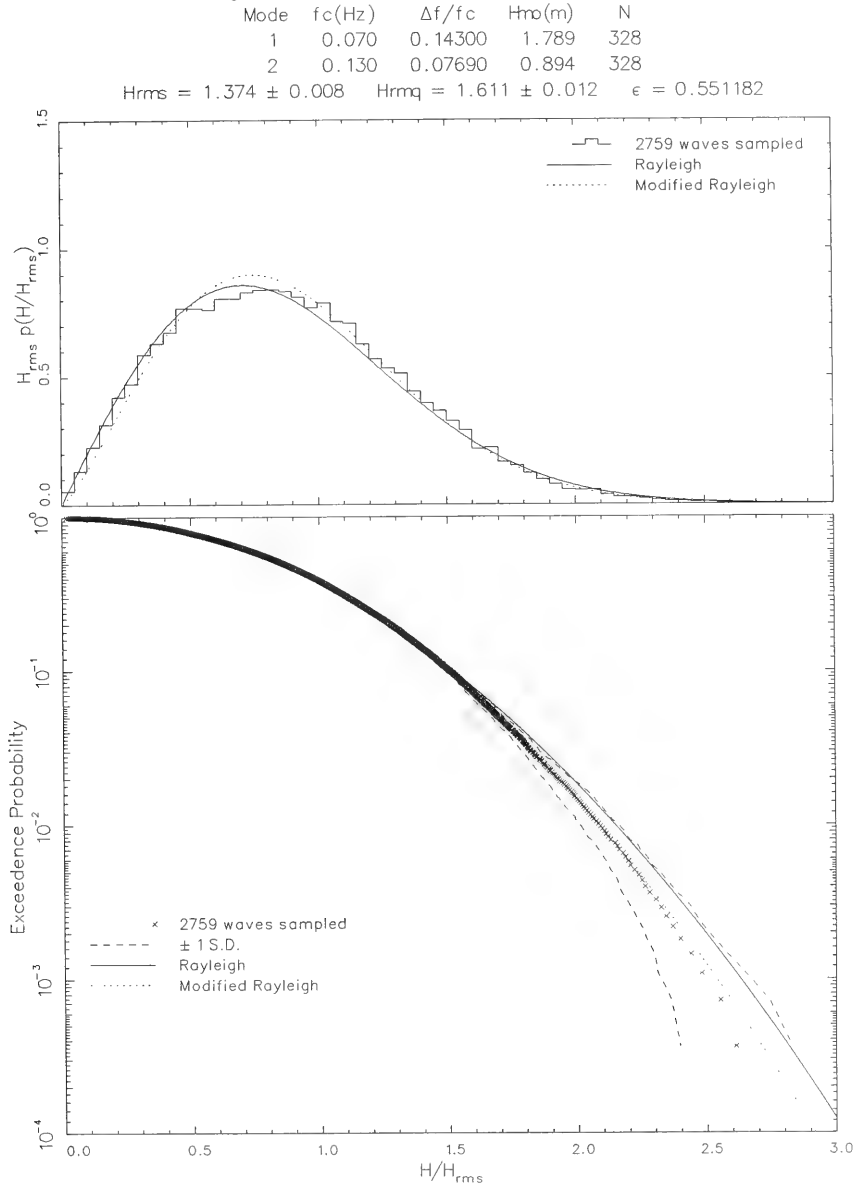


Figure 13. Probability and exceedence curves for bimodal spectra with intermediate modal separation

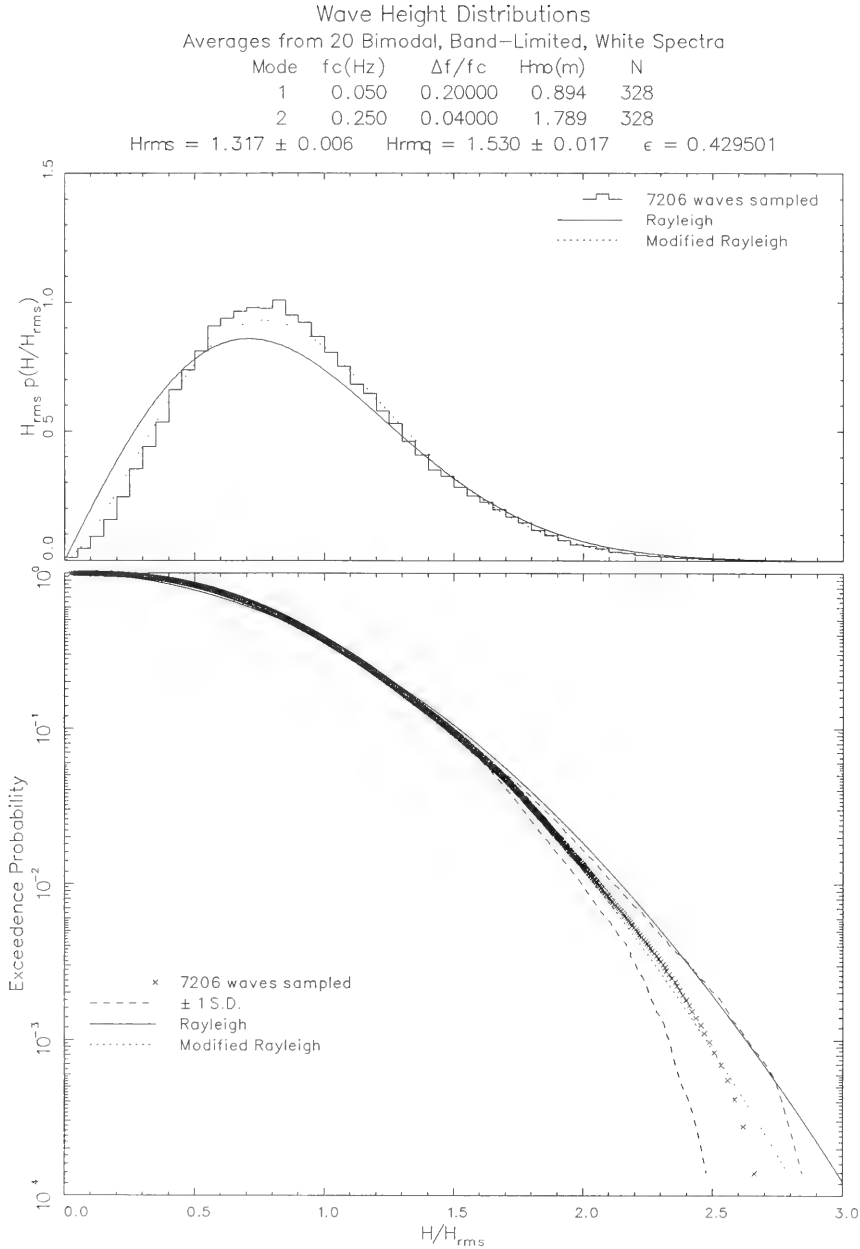


Figure 14. Probability and exceedence curves from bimodal spectra with large modal separation and high energy ratio

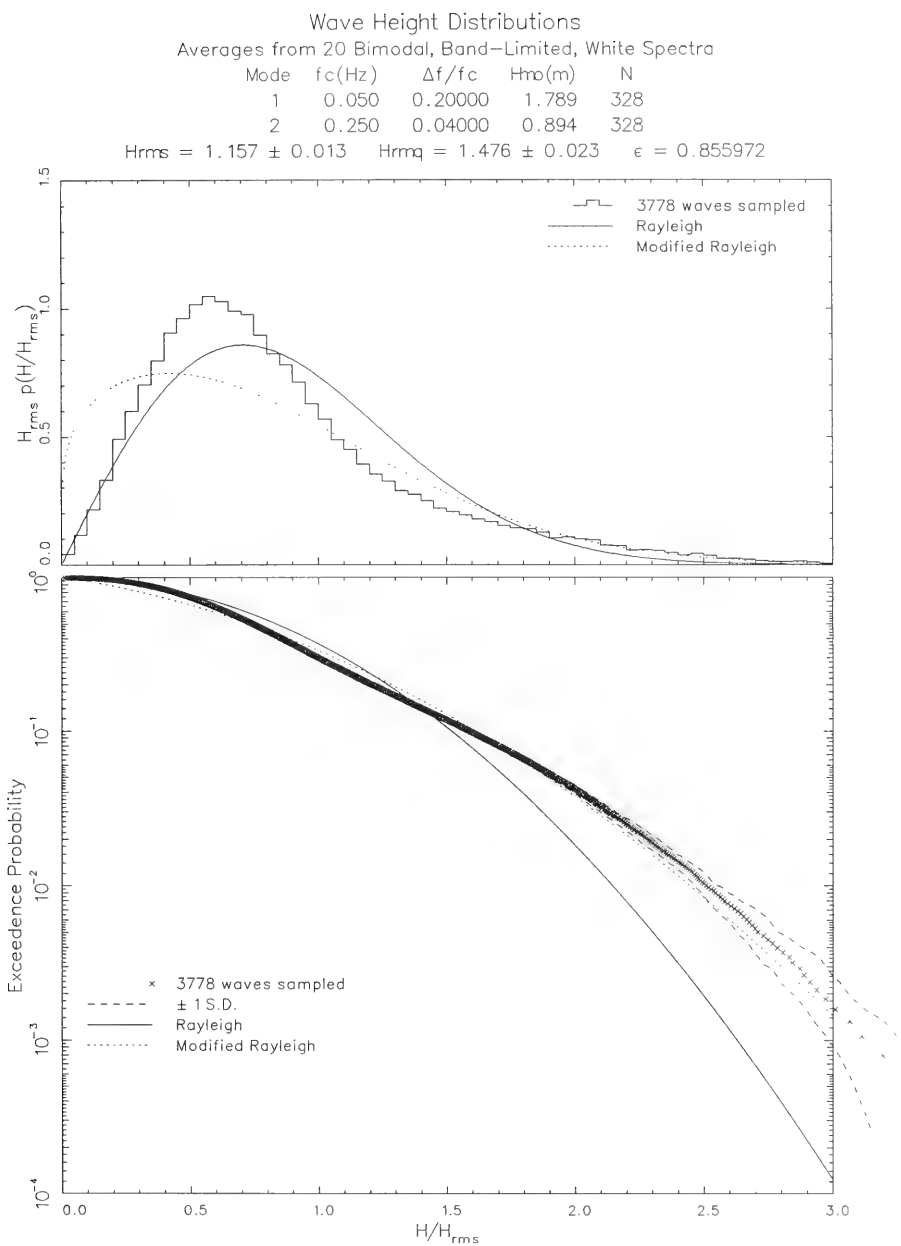


Figure 15. Probability and exceedence curves from bimodal spectra with large modal separation and small energy ratio

Synthetic wave heights average to an exceedence curve that is more nearly represented by the Modified Rayleigh model than the Rayleigh model. Dramatic differences from Figure 14 are seen in Figure 15, which represents a case with small relative energy. In Figure 15, synthetic data differ dramatically from both Rayleigh and Modified Rayleigh pdf curves. The exceedence estimate of the synthetic data clearly differs from the Rayleigh model, as shown in the lower part of Figure 15. Synthetic data also deviate by more than one standard deviation from the Modified Rayleigh curve at low and intermediate wave heights. The Modified Rayleigh model does appear to conform reasonably well with synthetic data in the high-wave tail of the distribution, but this may simply be a fortuitous circumstance, given the poor agreement elsewhere in the distribution.

82. To summarize, it appears that for spectra with broad ranges of modal separation parameter and with a moderate range of relative modal energies roughly greater than 1.0, derived wave height distributions tend to be well-represented by the Modified Rayleigh model. Most distributions in these ranges of parameters tend to be overpredicted by the Rayleigh model, as was found for time series derived from broad-banded spectra and as have been found frequently in natural observations (SPM 1984). For large modal separations and small relative modal energy, synthetic data deviate strongly from the Rayleigh model and deviate significantly from the Modified Rayleigh model. This result suggests the need for a third model to represent wave height distributions under conditions where governing spectra have wide modal separations and low relative energies. It also suggests that further study be performed to determine how frequently such conditions occur in nature, so as to determine their relative importance.

PART VI: CONCLUSION

83. In this report, a simple examination is made of wave height distributions in synthetic sea states characterized by energy spectra with multiple peaks, i.e., having energy centered at two or more distinct frequencies. Sea states with broad frequency spectra are included in this study since such seas are also composed of waves of diverse frequencies. These cases violate the assumptions of unimodal, narrow spectra that are formally required for the conventionally used Rayleigh distribution of wave heights to apply. Because of common usage of this distribution in engineering design, it is important to determine what errors are incurred by these violations and if an alternate model can compensate for them.

84. Reported here are tests of the Rayleigh and Modified Rayleigh wave height distribution models. These models are compared with wave heights from idealized, synthetic time series having spectra with variable widths, numbers of modes, and modal separations. The Modified Rayleigh model is the two-parameter, deepwater asymptotic form of the Beta-Rayleigh distribution introduced by Hughes and Borgman (1987). Synthetic time series are produced by inverse Fourier transform techniques and consist of 65,536 points at a nominal time step of 0.5 sec, so the record simulates in excess of 9 hr of sampling. Synthetic wave periods are maintained near 10 sec so the samples contain typically about 3,000 waves, enough to compute some reasonably concise statistics. Random phases are used in signal generation, so there is the possibility of some variation between runs with otherwise constant generating parameters. Hence, 20 runs were done for each case, and the results averaged to produce a stabler estimate of expected behavior.

85. The first part of this work involves determining some of the limits implied by the phrase "unimodal, narrow spectrum" in the synthesis of time series having Rayleigh distributions of wave heights. It is found that for unimodal, band-limited, white spectra, there are constraints on both the overall bandwidth and the number of component waves for a close approximation to a Rayleigh process to occur. Average heights $H^{(1/1)}$ and the average of the highest one-third waves $H^{(1/1)}$ are within 2 to 3 percent of Rayleigh estimates as long as bandwidths normalized by center frequencies do not exceed about $\Delta f/f_c = 0.4$. For $H^{(1/10)}$, it appears that about 20 component waves are necessary to differ from a Rayleigh $H^{(1/10)}$ by less than 10 percent. For

$H^{(1/100)}$, it takes about 100 component waves to stay within 10 percent of a Rayleigh $H^{(1/100)}$. Beyond a certain bulk spectral bandwidth, the parameters do not appear to approach Rayleigh parameters, no matter how many component waves are used in the spectral definition. This is most clear for the widest case investigated here, $\Delta f/f_c = 1.60$, but is also slightly evident in the next-to-widest case $\Delta f/f_c = 0.80$. It is notable that even the mean wave height $H^{(1/1)}$ and the average of the highest one-third waves $H^{(1/3)}$ show some effect of spectral broadening.

86. The Modified Rayleigh pdf and exceedence curves clearly represent the synthetic data better than the Rayleigh curves. With roughly the same constraints on numbers of component waves as for the Rayleigh comparison, averages of synthetic wave height distributions were within 2 percent of averages from the Modified Rayleigh model for all bandwidths and for averages out to $H^{(1/100)}$. This result suggests that the Modified Rayleigh curve is a better model for broad spectra if the added parameter H_{rmq} can be determined directly from the spectra. It is not yet evident that this determination can be made, and further research is required to characterize H_{rmq} in terms of spectral shape.

87. The primary part of this investigation is examination of wave height distributions in wave signals derived from multimodal spectra. To keep the study simple, only bimodal cases are considered. Using criteria established in the unimodal tests, bimodal spectra are constructed from two unimodal spectra, each of which yields a Rayleigh wave height distribution by itself. Given this characteristic, the details of the mode structure (bandwidth and number of component wave trains) are no longer important. The two primary variables in these bimodal tests are a measure of modal separation in the frequency domain and the relative amount of energy in one mode compared with the other.

88. Modal separation is characterized by differences between modal center frequencies normalized by the mean of the two modal center frequencies. In the present tests, the smallest modal separation is zero, corresponding to modes coincident in the frequency domain and representing a directional wave field where the modes have two different peak directions. The largest modal separation has one mode center frequency at 0.05 Hz and the other at 0.25 Hz, corresponding to wave periods of 20 and 4 sec, respectively, and considered to be a practical limit for wind waves. Five levels of relative energy have been

used in the present tests. The relative energy ratio varies from 0.25 to 4.0 following a geometric progression.

89. Results of this part indicate that for spectra with broad ranges of modal separation parameter and with a ratio of low-frequency modal energy to high-frequency modal energy roughly greater than 1.0, derived wave height distributions tend to be well-represented by the Modified Rayleigh model. The Rayleigh model tends to overpredict all the wave height averages for these parametric ranges in much the same way as was found for time series derived from broad-banded spectra and as have been found frequently in natural observations (SPM 1984). For large modal separations and small high-frequency modal energy relative to low-frequency modal energy, synthetic data deviate strongly from the Rayleigh model and deviate significantly from the Modified Rayleigh model. This result suggests the need for a third model to represent wave height distributions under conditions where governing spectra have wide modal separations and low relative energies. It also suggests that further study be performed to determine how frequently such conditions occur in nature, so as to determine their relative importance.

90. The two key findings in this study are, first, that spectra that are not narrow-banded or are bimodal with moderate-to-wide modal separations tend to have wave height distributions that are overpredicted by the Rayleigh wave height model, which may account for the common observation that the Rayleigh distribution is conservative for engineering design. Second, the Modified Rayleigh wave height model has a much better capability of characterizing distributions from broad-banded or multimodal wave fields. Further research is required to determine the parameters of the Modified Rayleigh model in terms of spectral parameters. Finally, these conclusions need to be tested with natural data to ensure that the artifice of synthesis has not biased the results.

REFERENCES

Abramowitz, M., and Stegun, I. A. 1970. Handbook of Mathematical Functions, Dover, New York.

Bendat, J. S., and Piersol, A. G. 1971. Random Data: Analysis and Measurement Procedures, Wiley-Interscience, New York.

Cartwright, D. E., and Longuet-Higgins, M. S. 1956. "The Statistical Distribution of the Maxima of a Random Process," Proceedings of the Royal Society of London, Series A, Vol 237, pp 212-232.

Earle, M. D. 1975. "Extreme Wave Conditions During Hurricane Camille," Journal of Geophysical Research, Vol 80, pp 377-379.

Forristall, G. Z. 1978. "On the Statistical Distribution of Wave Heights in a Storm," Journal of Geophysical Research, Vol 83, pp 2353-2358.

Hughes, S. A., and Borgman, L. F. 1987. "Beta-Rayleigh Distribution for Shallow Water Wave Heights," Proceedings of a Conference on Coastal Hydrodynamics, American Society of Civil Engineers, 28 June-1 July 1987, Newark, DE, pp 17-31.

Long, C. E. "Use of Theoretical Wave Height Distributions in Directional Seas," Technical Report in preparation, US Army Engineer Waterways Experiment Station, Vicksburg, MS.

Longuet-Higgins, M. S. 1952. "On the Statistical Distribution of the Heights of Sea Waves," Journal of Marine Research, Vol 11, pp 245-266.

_____. 1980. "On the Distribution of the Heights of Sea Waves: Some Effects of Nonlinearity and Finite Bandwidth," Journal of Geophysical Research, Vol 85, pp 1519-1523.

Shore Protection Manual. 1984. 4th Ed., 2 Vols, US Army Engineer Waterways Experiment Station, Coastal Engineering Research Center, US Government Printing Office, Washington, DC.

Thompson, E. F. 1980. "Energy Spectra in Shallow U.S. Coastal Waters," Technical Paper No. 80-2, US Army Corps of Engineers, Coastal Engineering Research Center, Fort Belvoir, VA.

Thornton, E. B., and Guza, R. T. 1983. "Transformation of Wave Height Distribution," Journal of Geophysical Research, Vol 88, pp 5925-5938.

APPENDIX A: NOTATION

A_{xk}	Fourier cosine coefficient at k^{th} frequency for time series x
A_{yk}	Fourier cosine coefficient at k^{th} frequency for time series y
B_{xk}	Fourier sine coefficient at k^{th} frequency for time series x
B_{yk}	Fourier sine coefficient at k^{th} frequency for time series y
C_{xk}	Modulus of sinusoidal component k of time series x
df	Infinitesimal increment of frequency
dH	Infinitesimal increment of wave height
dt	Time step
dx	Infinitesimal increment of x
f	Frequency
f_c	Center frequency of a spectrum
f_1 , f_2	Specific frequencies
H	Wave height
H_{mo}	Spectrum-based characteristic wave height
H_n	The n^{th} wave height in a set
$H^{(r)}$	Average of highest fraction r of all wave heights
H_{rmq}	Root-mean-quad wave height
H_{rms}	Root-mean-square wave height
$H^{(1/3)}$	Average of highest one-third of all wave heights
$H_{1/3}$	Alternate notation for $H^{(1/3)}$
$\hat{H}$	Randomly chosen wave height
i	Summing index
j	Index of an observed wave height; index of an element of a discrete function
J_u	Number of wave heights in accumulation bin u

k	Summing index; frequency index
m_n	The n^{th} moment of a frequency spectrum
m_0	Zeroth moment of frequency spectrum
m_2	Second moment of frequency spectrum
m_4	Fourth moment of frequency spectrum
n	Index of a set of discrete wave heights; index of a sequence of time steps
N	Number of observed waves in a record; number of observed water surface elevations
p	Probability density function
p_2	Two-wave probability density function
pdf	Probability density function
p_{MR}	Modified Rayleigh probability density function
p_R	Rayleigh probability density function
P	Cumulative probability function
P_2	Two-wave cumulative probability function
P_{MR}	Modified Rayleigh cumulative probability function
P_R	Rayleigh cumulative probability function
Prob[ ]	Probability that expression in [ ] is true
Q	Exceedence probability
Q_2	Two-wave exceedence probability
Q_D	Exceedence probability estimated from data
Q_{MR}	Modified Rayleigh exceedence probability
Q_R	Rayleigh exceedence probability
r	Fraction between zero and one
r_j	Element j of function representing fraction of highest observed waves
RMS	Root-mean-square

S	Sea surface variance spectral density
S_{xk}	Discrete spectral density at k^{th} frequency of time series x
S_{yk}	Discrete spectral density at k^{th} frequency of time series y
t_n	Element n of a sequence of N discrete times
u	Bin index for histogram of wave heights
U_{xk}	Uniform random deviate at k^{th} frequency of time series x
U_{yk}	Uniform random deviate at k^{th} frequency of time series y
x	Dummy integration variable; a time series
x_n	Element n of discrete time series x
X_k	Discrete Fourier transform element k of time series x
y	A time series
y_n	Element n of discrete time series y
Y_k	Discrete Fourier transform element k of time series y
Z_k	Discrete Fourier transform element k of complex time series x + iy
α	Parameter of Modified Rayleigh pdf
Γ	Gamma function
Δf	Discrete increment of frequency
ΔH	Discrete increment of wave height
Δt	Discrete time step
ϵ	Frequency spectral width parameter
ϕ_{xk}	Phase of sinusoidal component k of time series x