NUMERICAL VALIDATION OF SOME WAVE HEIGHT DISTRIBUTIONS

Mahmoud A. Sharaki

Department of Transportation Engineering, Alexandria University,
Alexandria, Egypt

ABSTRACT

This paper studies, numerically, the validity of different wave height distributions and the minimum number of components needed to validate the distribution under study. Also, the effect of bandwidth of the wave spectrum is considered. Rayleigh distribution, the modified Rayleigh distribution and the Beta-Rayleigh distribution are considered in this study

Keywords: Wave height distributions, Wave statistics, Rayleigh distribution, Modified Rayleigh distribution, Beta-Rayleigh distribution.

INTRODUCTION

It was proved analytically that, if the sea surface elevation can be represented by a sum of an infinite number of harmonic components uniformly distributed about a central frequency and having the same amplitude., then the wave height can be represented by a Rayleigh distribution, assuming that the bandwidth of the wave energy spectrum is narrow. It is not clear what is meant by narrow band spectrum. More clearly, what is the limit of the bandwidth that allows Rayleigh distribution to represent the wave height distribution accurately.

In actual situations, we restore to numerical methods, where the energy spectrum is divided into finite number of components. In this paper we try to find the minimum number of components, if any, that makes Rayleigh distribution a good representative of the wave heights.

We numerically synthesize a collection of wave trains from a uniform spectrum using different numbers of components and bandwidths and compare the different statistical quantities of the synthesized data with the Rayleigh distribution. which is a one-parameter model. We also introduce two proposed distributions to represent the wave heights., namely the Beta-Rayleigh distribution, and the modified Rayleigh distribution. The former was introduced for the first time by Hughes and Borgman [1], for the distribution of the shallow water wave heights. It is a three-parameter distribution. The later is a two-parameter distribution .

The statistical quantities used to validate any distributions are H_1, H_3, H_{10} and H_{100}, which are the average of the whole train, the highest one third, one tenth and one hundredth respectively.

The importance of this study comes from the fact that for engineering design purposes it is useful to have a statistical description of the heights of the sea waves so that a probability can be assigned to a particular water level. This will help in the successful design of shore protection projects and find the likelihood of a project survival.

STATISTICAL MODELS

As was mentioned before, we will use three different Probability Density Functions, PDFs, as candidates to represent wave height distributions in the numerical analysis. Those PDFs are

1-Rayleigh distribution, which is a one parameter model. It depends on the root mean square H_{rms}, defined by

$$H_{rms} = \left[\frac{1}{N} \sum_{n=1}^{N} H_n^2 \right]^{0.5} \qquad (1)$$

where N is the number of waves in the record. In this study we will adopt the zero up crossing definition for the wave height. The Rayleigh PDF is given by

$$P_R(H) = \frac{2H}{H_{rms}^2} \exp -\left(\frac{H}{H_{rms}}\right)^2 \qquad (2)$$

It is preferred to put all quantities in a dimensionless form that is to normalize H with respect to H_{rms}. Writing

$$H = H^* H_{rms} \qquad (3)$$

where H* is dimensionless

Equation 2 has the dimension (1/L), where L is the length, then we can write

$$P_R(H^*) = 2H^* \exp - H^{*2} \qquad (4)$$

Which gives a dimensionless quantity.

As can be seen from Equations 2 and 4, Rayleigh PDF has the disadvantage that, there is no upper bound for the wave height. In fact there is always a very small probability that a very large wave may occur.

2-The Beta-Rayleigh PDF was introduced for the first time by Hughes and Borgman (1987) [1] to describe the distribution of the wave heights in shallow water, taking the breaking effect into account. It is a three-parameter PDF, besides it has the advantage that there is an upper limit for the wave height. It takes the form

$$P_{BR}(H) = \frac{2\,\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)} \frac{H^{2\alpha-1}}{H_m^{2\alpha}} \left[1-\frac{H^2}{H_m^2}\right]^{\beta-1} \qquad (5)$$

where

$$\alpha = K_1 (K_2 - K_1) / (K_1^2 - K_2) \qquad (6)$$

$$\beta = (1-K_1)(K_2 - K_1) / (K_1^2 - K_2) \qquad (7)$$

$$K_1 = H_{rms}^2 / H_m^2 \qquad (8)$$

$$K_2 = H_{rmq}^2 / H_m^4 \qquad (9)$$

$$H_{rmq} = \left[\frac{1}{N} \sum_{i=1}^{N} H_i^4\right]^{0.5} \qquad (10)$$

In the original work H_m is defined as the maximum height for the breaking wave. In this study we define it as the maximum wave height.. Two alternatives are introduced to represent H_{in} Firstly, we will take it as the maximum wave height in the record. Secondly, we will take it as the maximum wave height as predicted by Rayleigh distribution that is given by H_{1000}, according to Chakrabarti [2].

A dimensionless form of Equation 5 takes the form

$$P_{BR}(H^*) = \frac{2\,\Gamma(\alpha-\beta)}{\Gamma(\alpha)\,\Gamma(\beta)} \frac{H^{*2\alpha-1}}{H_m^{*2\alpha}} \left[1-\left(\frac{H^*}{H_m^*}\right)^2\right]^{\beta-1} \qquad (11)$$

Where

$$H_m = H_m^* H_{rms} \qquad (12)$$

The modified Beta-Rayleigh distribution will be tested once with H_m given by H_{1000} and once by the maximum wave height in the record. This gives 2 alternatives.

3-The modified Rayleigh distribution given by, in dimensionless form

$$P_{mr}(H^*) = \frac{2\,H^{*\,2\alpha-1}}{b_o^{\alpha ao}\ \Gamma(\alpha_0)} \exp\left(\frac{-H^{*2}}{b_o}\right)^2 \qquad (13)$$

where

$$\alpha_o = \frac{1}{H_{rms}^{*2} - 1} \qquad (14)$$

$$b_o = H_{rmq}^{*2} - 1 \qquad (15)$$

Equation 13 is nothing but Equation 11 when H_m tends to infinity. It is a two-parameter PDF.

Data Generation

A computer program is written in MATLAB to generate the required wave train and carry out the statistical analysis and comparison processes. To check the validity of the program two methods are utilized. Firstly, a wave train with just two components is generated and its distribution is compared with the theoretical distribution, given by Longuet Higgins [3]

$$p(H^*) = \begin{cases} \dfrac{2}{\pi\,(2 - H^*)^{0.5}} & H^* \leq \sqrt{2} \\ 0 & H^* > \sqrt{2} \end{cases} \qquad (16)$$

Secondly, H_{rms} as given by Equation 1, is checked versus its value as calculated from the wave spectrum by

$$H_{rms} = 2\sqrt{m_o} \qquad (17)$$

where m_i is the i th spectrum moment, given by

$$m_i = \int_0^{\infty} f^i\ S(f)\ df \qquad (18)$$

where f is the cyclic frequency, cycles per second,

To measure the bandwidth of the wave spectrum used to generate the wave train, two methods are in use. That due to Vanmarcke mentioned in Reference 2:

$$\varsigma_v = \left(1 - \frac{m_1^2}{m_o\ m_2}\right)^{0.5} \qquad (19)$$

and that due to Cart wright and Longuet Higgins mentioned in Reference 2:

$$\varsigma_L = \left(1 - \frac{m_2^2}{m_o\ m_4}\right)^{0.5} \qquad (20)$$

The wave train is generated using the formula

$$\eta_k = \sum_{i=1}^{n} C_i \cos(2\pi f_i\ k\,\Delta t + \phi_i\ ) \qquad (21)$$

where

$C_i = (2\,\Delta f\, S_i)^{0.5}$ is the amplitude of the i th component with cyclic frequency f_i

$\phi_i = 2\pi$

U[0,1] is a phase angle uniformly distributed in the interval [0, 2]

η_k is the surface elevation at time $t = k\,\Delta t$

S_i is the amplitude of the wave spectrum at frequency f_i

Δt =

T/20 is the time step, and

T is the period of the shortest wave in the record

Δt is selected equal T/20 so that the maximum error in calculating the crest or trough height is not more than 1.23% = [cos (0) − cos (9)] 100 %. This comes from the fact that the crest or trough is located not more than 9 degrees from the nearest point in the time series.

The simulation time is selected so that the longest wave component in the train appears several times. In this study we selected it to appear 200 times. This allows a long number of wave heights to be obtained for statistical analysis.

In this study we selected a narrow band white spectrum centered at f_c = 0.125 Hz (T = 8 sec.) and different values for $\Delta f = f_k - f_l$ as shown in Table 1. Where f_l and f_k are the smallest and the largest frequencies of the train, The number of wave components for each case is also shown in Table 1.

To compensate the effect of the random phase angle ϕ used in the analysis, each rum, given f_c, Δf, and the number of components is repeated 16 times and the analysis is carried out for the total sum of them. As shown in the table, the period of the longest wave is 10 sec., so the maximum simulation time is about 535 min.

Table 1 Properties of the wave trains used in the analysis

f	Δf/f_c	Fi	f_k	T_{max}	T_{min}	L	# of Comp
0.003125	0.025	0.1234	0.1265	8.1	7.901	0.0014	8, 16, 32, 64, 128, 256, 512, 1024
0.00625	0.05	0.1218	0.1281	8.205	7.805	0.0054	
0.0125	0.1	0.1187	0.1313	8.421	7.619	0.0212	
0.025	0.2	0.1125	0.1375	8.889	7.273	0.0775	
0.05	0.4	0.1	0.15	10.00	6.667	0.2312	

NUMERICAL TEST RESULT

As mentioned before, Equation 16 is used to validate the program. It is shown that there is good agreement between both the theoretical distribution and the one obtained from the synthesized wave train with two components. This is true when the difference between the frequency of the two components is small. Also, comparing the value of H_{rms} as calculated by Equation 1 and Equation 17 shows a very close agreement, especially when both the simulated time is long and the number of components is large.

Figures 1 and 2 show bar graphs for the wave heights distribution versus Rayleigh distribution. The number of components used in the analysis is 8 and the bandwidth equals 0.0013626 as calculated using Equation 20. The figures show that there is a pronounced deviation from Rayleigh distribution, in spite of the limited bandwidth. This may be attributed in part to the limited number of components used in the analysis and in part to effect of the random phase angle ϕ. To compensate for the second reason, a total of 16 runs for each case are carried out and the sum of them is used for the statistical analysis.

Figure 3 shows the statistics of a wave train with 1024 components and bandwidth 0.23123 versus Raleigh distribution Figure 3 shows less deviation between the actual distribution of the wave heights and the Rayleigh distribution in spite of the larger bandwidth. This reflects the importance of the number of components used in the analysis.

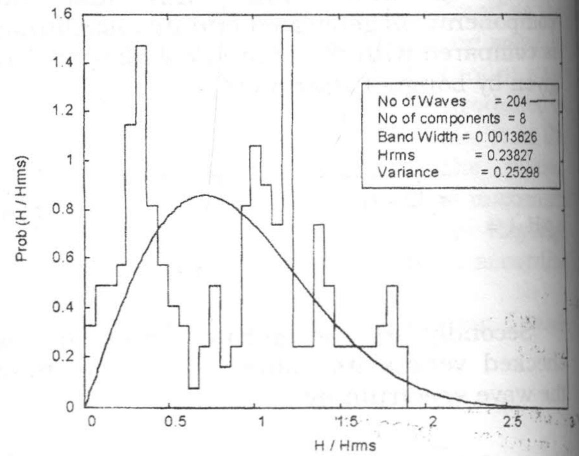

Figure 1 Wave height distribution versus Rayleigh distribution

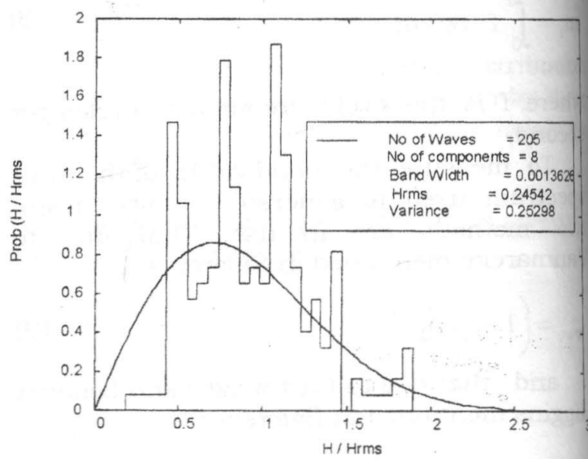

Figure 2 Wave height distribution versus Rayleigh distribution

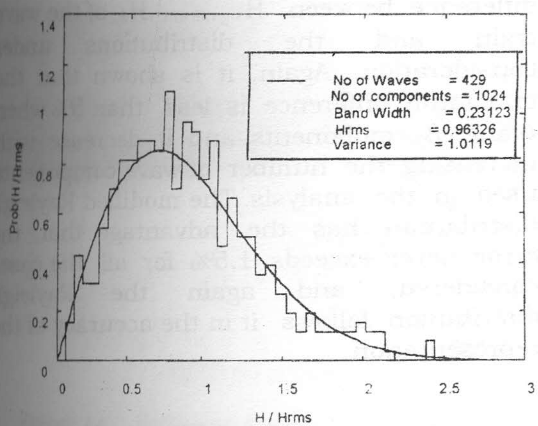

Figure 3 Wave height distribution versus Raleigh distribution

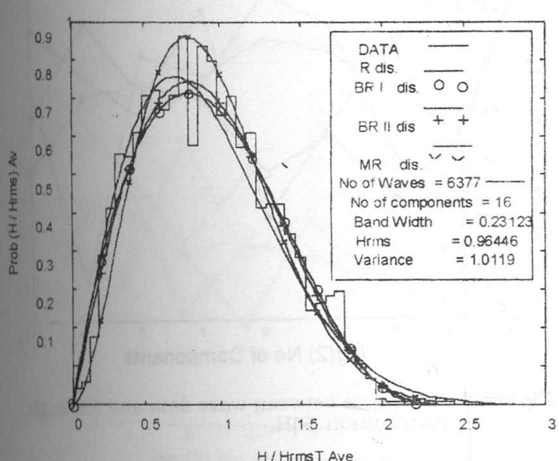

Figure 4 Typical Wave height distribution versus different theoretical distribution

Figure 4 shows a typical distribution of the wave heights versus the different theoretical distributions used in this study. Here, R stands for Rayleigh distribution. BRI for Beta-Rayleigh distribution using the H_{max} from the record, BRII for Beta-Rayleigh distribution using H_{max} as calculated by Rayleigh distribution, and finally, MR stands for modified Rayleigh distribution.

A total of 40 cases are studied. Table 1 shows the parameters used in the test. The primary variables are the bandwidth and the number of components. While the central frequency is kept constant at $f_c = 0.125$. All wave heights are normalized with respect to H_{rms}. The test is based on the quantities H_1, H_3, H_{10}, H_{100} where H_i stands for the average of the highest 1/i waves. The comparison of the different distributions is based on the following equation

$$\% \, difference = \frac{H_i^D - H_i^M}{H_i^D} \qquad (22)$$

where,
H^D the generated wave height
H^M the distribution wave height

The results of the comparisons based on Equation 22 are shown in Figures 5 through 21. The ordinate axes show the outcome of Equation 22, while the abscissa show the log, for base2, of the number of components used in the analysis.

Figure 5 through Figure 8 show the difference between H_{mean} of the wave train and the distributions under study. From the figures it is clear that the deviation is less than 5% for all bandwidths and number of wave components. It is less than 1.5% for the modified Rayleigh distribution. For wide band white spectrum, it is clear from the figures that the modified Rayleigh distribution represents the wave data better than the other three distributions and the Rayleigh distribution gives the second best results.

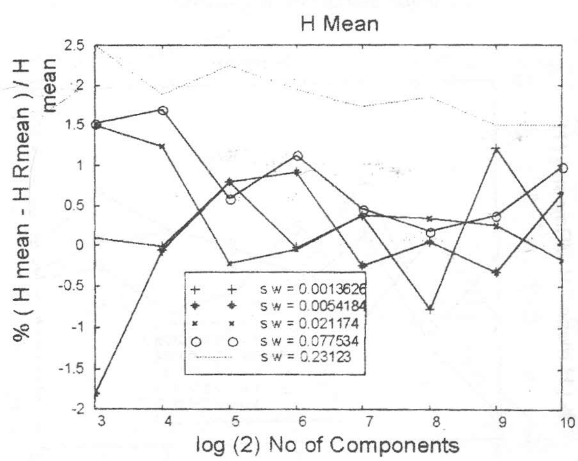

Figure 5 Difference between wave data and Rayleigh distribution mean

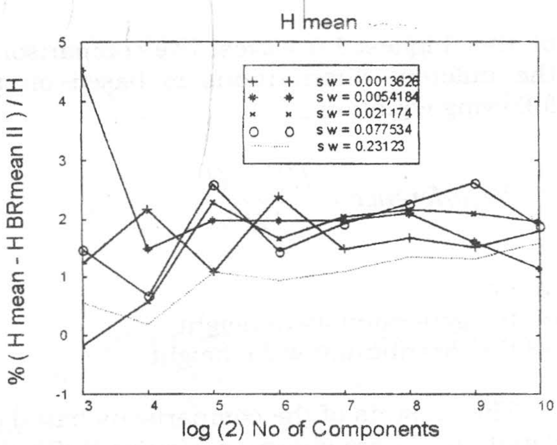

Figure 6 Difference between wave data and Beta-Rayleigh distribution (BRI) mean

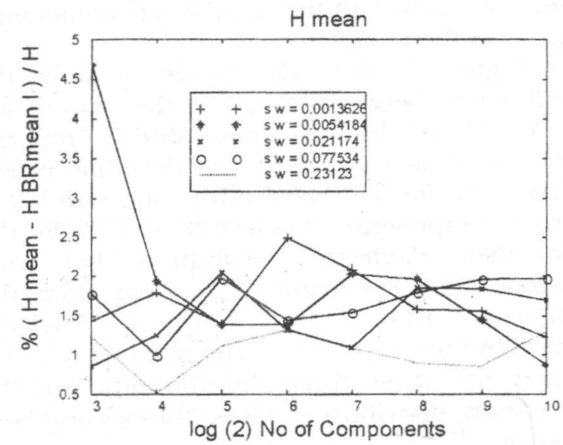

Figure 7 Difference between wave data and Beta-Rayleigh distribution (BRII)H_m

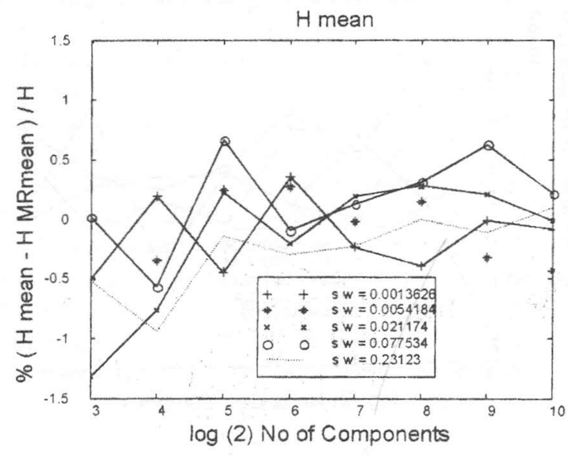

Figure 8 Difference between wave data and modified Rayleigh distribution (MR)H_m

Figure 9 through Figure 12 show the difference between $H_{significant}$,H_3. of the wave train and the distributions under consideration. Again, it is shown that the maximum difference is less than 5% when using 8 components and it decrease with increasing the number of wave components used in the analysis. The modified Rayleigh distribution has the advantage that, the error never exceeds 1.5% for all the cases considered, and again the Rayleigh distribution follows it in the accuracy of the representation

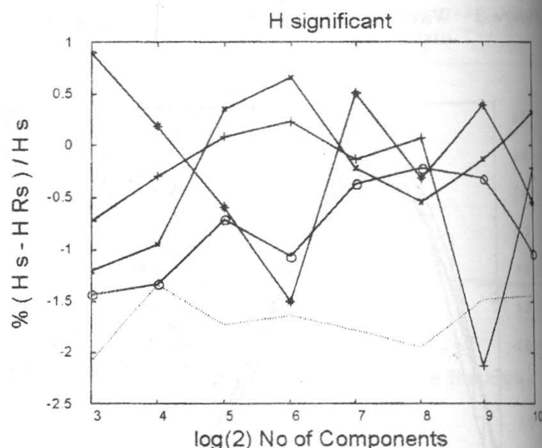

Figure 9 Difference between wave data and Rayleigh distribution (R)H_s

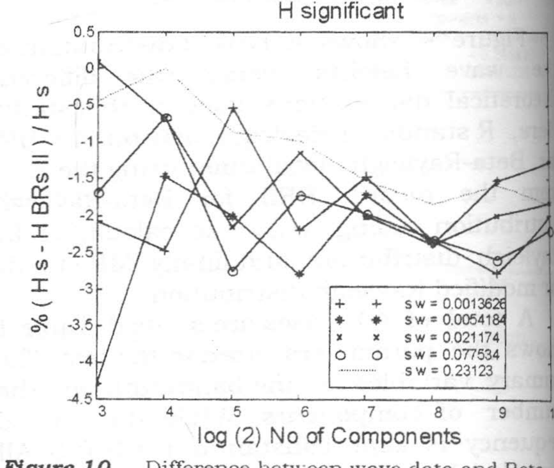

Figure 10 Difference between wave data and Beta-Rayleigh distribution (BRI)-H_s

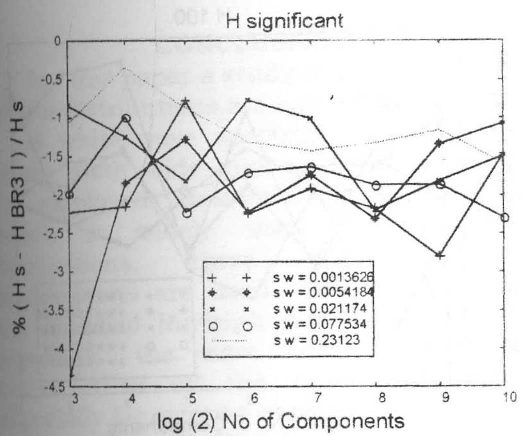

Figure 11 Difference between wave data and Beta-Rayleigh distribution (BRII) H_s

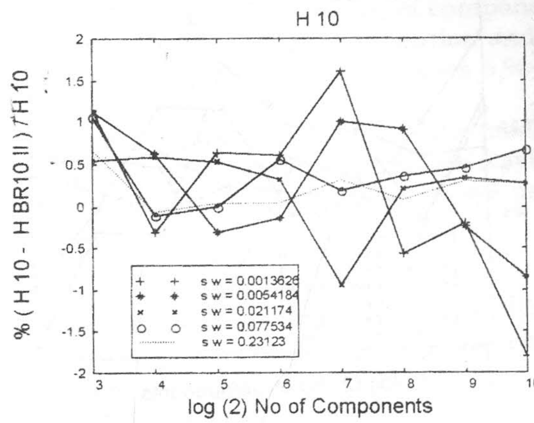

Figure 13 Difference between wave data and Rayleigh distribution (R) H_{10}

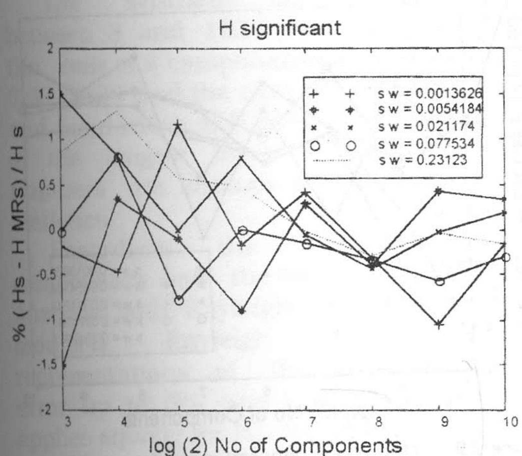

Figure 12 Difference between wave data and modified Rayleigh distribution (MR) H_s

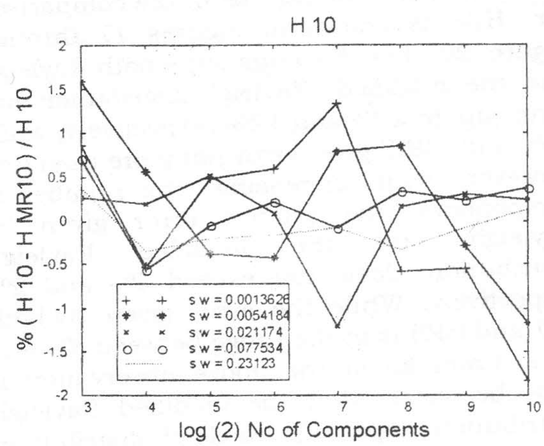

Figure 14 Difference between wave data and Beta-Rayleigh distribution (BRR) H_{10}

Figure 13 through Figure 16 show the difference between H_{10} of the generated wave train and the distributions under study. Excluding the results given by Rayleigh distribution, which gives a deviation up to 6%, all other distributions give deviation not more than 1.5% of the actual wave data. For BRI, and BRII distributions, this may be attributed to the use of H_{max} as a parameter in the distribution.

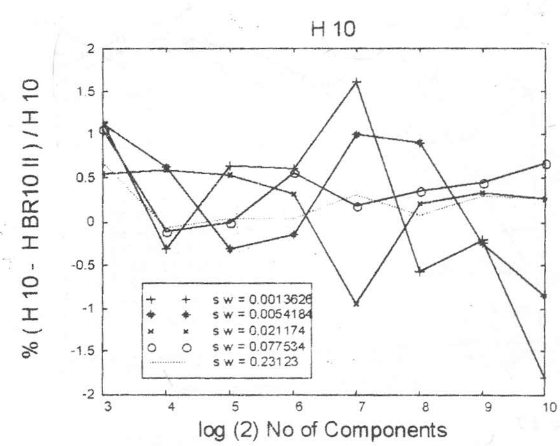

Figure 15 Difference between wave data and Beta-Rayleigh distribution (BRII) H_{10}

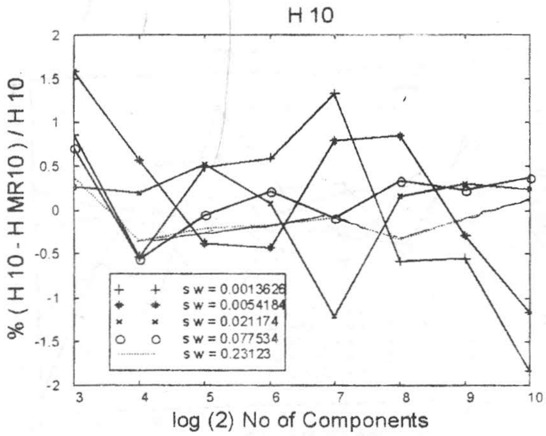

Figure 16 Difference between wave data and modified Rayleigh distribution (MR) H_{10}

Finally, the last set of the comparison for H_{100} is shown in Figures 17 through Figure 20. For 8 components both Rayleigh and the modified Rayleigh distribution give error up to 20% and 12% respectively, while BRI and BRII give error not more than 8%. However, with increasing the number of components the relative error given by Rayleigh and the modified Rayleigh distribution does not exceed 5% and 2% respectively. While the error given by both BRI and BRII is in the range between 8% and 10%. From all of the above observation, it may be clear that the modified Rayleigh distribution supercede Rayleigh distribution especially for extreme events representation.

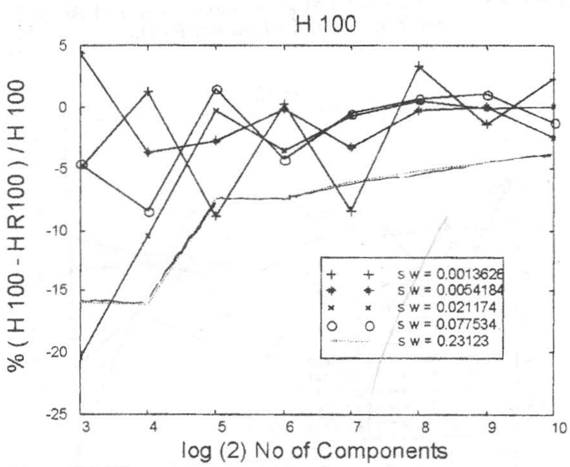

Figure 17 Difference between wave data and Rayleigh distribution (R) H_{100}

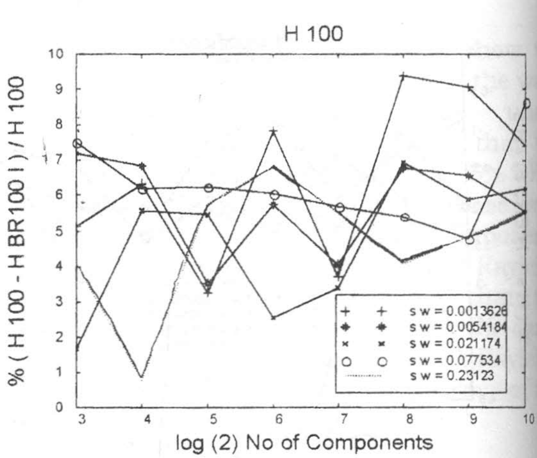

Figure 18 Difference between wave data and Beta-Rayleigh distribution (BRI) H_{100}

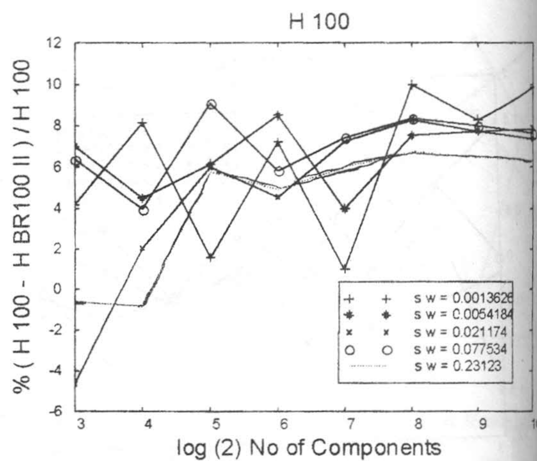

Figure 19 Difference between wave data and Beta-Rayleigh distribution (BRII) H_{100}

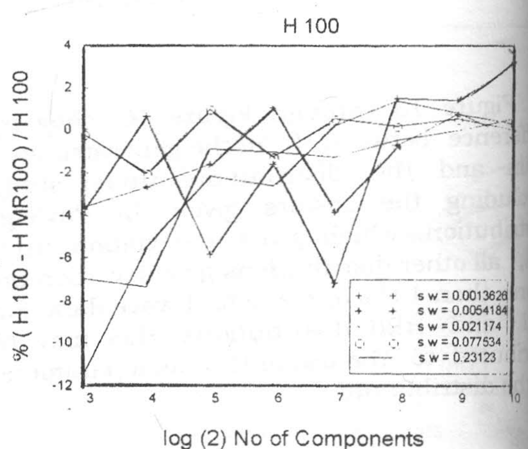

Figure 20 Difference between wave data and modified - Rayleigh distribution (MR) H_{100}

CONCLUSIONS

In this paper a study of a synthetic wave height distributions are carried out.

A white wave spectrum is used in this study. Both wide and narrow band spectra are employed. Besides Rayleigh distribution which is widely used in Engineering Applications, three others proposed distributions are used. Among all of them, the modified Rayleigh distribution seems to represent the wave heights distribution better than all the other distributions, especially for extreme events.

The modified Rayleigh distribution is a two parameter model and it is the asymptotic form of the Beta-Rayleigh model, introduced by Hughes and Borgman [1].

The synthetic time series contain between 8 and 1024 wave components. For the case of 8 components and band width of 0.0013626 and the case of 1024 components and band width of 0.23123, this means 3200 of the longest wave in each run. Those numbers are enough to computer concise statistics.

Regardless the number of wave components and the bandwidth used in the analysis both Rayleigh distribution and the modified Rayleigh distribution give representations of the wave heights with error up to 5% and 1.5% respectively. This applies equally for H_1, H_3 and H_{10}.

For H_{100} and for number of components more than 64, Rayleigh distribution and the modified Rayleigh distribution gives 5 % and 1.5 % error, respectively.

Those results suggest that the modified Rayleigh distribution must replace Rayleigh distribution in representing the wave height distribution.

REFERENCES

1. S.A Hughes and L.E. Borgman, "Beta-Rayleigh Distribution for Shallow Water Wave Heights", Proceedings of Conference on Coastal Hydrodynamic, ASCE Newark. DE, June 28 – July1 , pp. 17-31, (1987).
2. S.K. Chakrabarti, "Hydrodynamic of Offshore", Structures, Computational Mechanics Publications, Southampton Boston, (1994).
3. M.S Longuet Higgins, "On the Statistical Distribution of the Heights of sea Waves", Journal of Marine Research, Vol 11, pp. 245-266, (1952).

Received September 6. 1999
Accepted October 24. 1999

التحقق العددى من صلاحية بعض توزيعات الأمواج

محمود عبد المنعم شراكى
قسم هندسة المواصلات - جامعة الاسكندرية

ملخص البحث

يهدف هذا البحث الى التحقق من مدى صلاحية بعض التوزيعات النظرية مثل توزيع ريلاى وتوزيع ريلاى المعدل لتمثيل أرتفاع الأمواج وإيجاد الحد الأدنى لعدد المركبات المستخدمة وأيضا أقصى عرض لمنحنى الطيف المستخدم فى تمثيل الأمواج حتى يمكن استخدام هذه التوزيعات النظرية.