Empirical Moment-Based Estimation of Rayleigh-Stokes Distribution Parameters

Amir Hossein Izadparast
Research and Development Group, SOFEC Inc.
College Station, Texas, USA

John Michael Niedzwecki
Zachry Department of Civil Engineering, Texas A&M University.
College Station, Texas, USA

ABSTRACT

The Rayleigh-Stokes model has been widely applied to represent the probability distribution of crests and troughs of weakly non-linear random processes. In this study, the parameter estimates for the three-parameter Rayleigh-Stokes probability distribution model are obtained from application of two moment-based empirical parameter estimation methods, i.e. conventional method of moments and method of linear moments. Monte-Carlo simulations are utilized to compare the performance of these parameter estimation approaches in estimating the parameters of the Rayleigh-Stokes distribution and also to evaluate the uncertainty of the extreme statistics. Additionally, the effect of sample size on the uncertainty of the model statistics is evaluated. Finally, the Rayleigh-Stokes model is utilized to estimate the probability distribution of disturbed wave crests in beneath a mini-TLP and the model performance is evaluated.

KEY WORDS: Rayleigh-Stokes model, weakly non-linear process, probability distribution, method of moments, method of linear moments.

INTRODUCTION

The Rayleigh-Stokes model is a well-known probability distribution in the field of ocean wave mechanics and is widely utilized to estimate the probability distribution of weakly non-linear wave crests. The model initially developed by Tayfun (1980) for offshore wave crests was based on the assumptions that: 1) the waves can be modeled as a narrow-banded random process and consequently the wave crests of the linear waves follow the Rayleigh law (Longuet-Higgins 1952), and 2) wave elevations can be approximated by second-order Stokes wave theory. Tayfun's Rayleigh-Stokes model was a theoretical model in which the distribution structure was derived analytically and the underlying model parameter were obtained from its theoretical relation with the significant wave height and mean wave period. The mode structure and the estimate of model parameter were subsequently modified by other researches (Arhan and Plaisted 1981, Kriebel and Dawson 1991 and 1993, Tung and Huang 1985, Tayfun 2006). However, the different representations of the second-order Rayleigh-Stokes model converge to almost identical results for deepwater condition. The original one-parameter Rayleigh-Stokes model was reasonably successful in estimating the probability distribution of extreme offshore wave crests. However, the model did not consider the effects of interaction between incident, diffracted, and radiated waves and therefore the original model is not appropriate for wave crests close to or beneath an offshore structure.

Following an analogous methodology, Kriebel (1993) developed a two-parameter Rayleigh-Stokes model for probability distribution of non-linear wave run-up interacting with a fixed vertical column in deep water. The parameters of Kriebel's model were estimated from their relation with the significant wave height and mean period, and application of the linear diffraction theory. Kriebel's Rayleigh-Stokes model was compared with experimental data and was observed to underestimate the large crests (Stansberg and Nielsen 2001, Izadparast and Niedzwecki, 2009b and 2010). Similarly, Fedele and Arena (2005) developed the general two-parameter Rayleigh-Stokes distribution for the crests and troughs of second-order process and derived the theoretical estimates of the model parameters for a) the surface displacement and fluctuating wave pressure in an undisturbed field, b) waves in front of a rigid wall. Later, Izadparast and Niedzwecki (2009) developed a three-parameter Rayleigh-Stokes model for wave crests in the vicinity or beneath of an offshore platform. Izadparast and Niedzwecki (2009b and 2010a) utilized the three-parameter Rayleigh-Stokes model to estimate the probability distribution of wave run-up over vertical columns of offshore structures. In both studies, the model parameters were estimated empirically utilizing the method of linear moments (L-moments). It was shown that the empirically estimated Rayleigh-Stokes model is considerably successful in capturing the probability distribution of complex non-linear random variables.

In this study, the parameter estimates for the three-parameter Rayleigh-Stokes probability distribution model are obtained from application of the conventional method of moments (MoM). The conventional moments give more weight to the tail of the distribution and therefore they are more suited for use in the prediction of extremes than L-moments. However, the sample moments, especially the high-order moments, are more biased than the corresponding sample L-moments and the uncertainty of the small sample L-moments were shown to be less (Hosking 1990, Hosking and Wallis 1997). Here, Monte-Carlo simulations are utilized to investigate and compare the performance of both approaches in estimating the parameters of the Rayleigh-Stokes distribution model. Additionally, the uncertainty of the extreme statistics estimated from the two parameter estimation methods is evaluated and the effect of sample size on the uncertainty of the model statistics is studied. Finally the performance of Rayleigh-Stokes model is tested over measured disturbed wave crests beneath a mini-TLP model test.

MATHEMATICAL BACKGROUND

The wave crests and troughs of weakly non-linear and narrow banded waves ζ_n can be approximated from application of second-order Stokes wave theory as (Tayfun 1980)

$$\zeta_n = \alpha\zeta + \beta\zeta^2 + \gamma \tag{1}$$

where ζ is the narrow-banded linear random variable, α is the amplification of the linear term, β is the amplification of the quadratic term, and γ is the remaining shifting between linear and non-linear variables. In this model β and γ have real values, α is a positive real value, and in order to satisfy the weakly non-linear assumption $|\beta| << \alpha$.

In the Rayleigh-Stokes model it is assumed that the linear variable ζ follows a Rayleigh distribution law (Longuet-Higgins 1952) with probability density function (PDF) of

$$f_\zeta(x) = x\exp\left(-x^2/2\right) \tag{2}$$

From that and application of random variable transformation rule, the PDF, cumulative distribution function (CDF), and quantile distribution of non-linear variable for $\beta > 0$ is obtained as

$$f_{\zeta_n}(x) = \frac{\chi-\alpha}{2\beta\chi}\exp\left(-\frac{(\chi-\alpha)^2}{8\beta^2}\right)$$

$$F_{\zeta_n}(x) = 1-\exp\left(-\frac{(\chi-\alpha)^2}{8\beta^2}\right) \tag{3}$$

$$x_{\zeta_n}(u) = \gamma - 2\beta\ln(1-u) + \alpha\left(-2\ln(1-u)\right)^{1/2}$$

where for simplicity the following combination is utilized,

$$\chi = \left(\alpha^2 + 4\beta(x-\gamma)\right)^{1/2} \tag{4}$$

The probability distributions in Eq. (3) are defined for $x > \gamma$. In the case of $\beta < 0$, the probability distributions are derived in the form of

$$f_{\zeta_n}(x) = \frac{1}{2\beta\chi}\left[(\chi-\alpha)\exp\left(-\frac{(\chi-\alpha)^2}{8\beta^2}\right)H_\gamma(x) + -(\chi+\alpha)\exp\left(-\frac{(\chi+\alpha)^2}{8\beta^2}\right)\right] \tag{5}$$

$$F_{\zeta_n}(x) = \left[1-\exp\left(-\frac{(\chi-\alpha)^2}{8\beta^2}\right)\right]H_\gamma(x) + \exp\left(-\frac{(\chi+\alpha)^2}{8\beta^2}\right)$$

where $H_\gamma(x)$ is the step function and has a value of unity for $x \geq \gamma$ and is zero for $x < \gamma$. The distributions in Eq. (5) are defined for $-\infty < x < \gamma - \frac{\alpha^2}{4\beta}$ while only $\gamma < x < \gamma - \frac{\alpha^2}{4\beta}$ is physically justified. The exact analytical form of the quantile function is not available for the Rayleigh-Stokes model with $\beta < 0$. An approximate estimate for the quantile function can be obtained assuming that $(\alpha - \chi) << (\alpha + \chi)$ and consequently the third term in the $F_{\zeta_n}(x)$ can be ignored with respect to the second term; specifically,

$$x_{\zeta_n}(u) \approx \gamma - 2\beta\ \ln(1-u) + \alpha\ \left(-2\ln(1-u)\right)^{1/2} \tag{6}$$

In Fig. 1, the effects of model parameters on the quantile distribution of Rayleigh-Stokes model is shown. In this figure the linear amplification factor varies in the range of $1.0 \leq \alpha \leq 1.6$, β varies in the range of $-\alpha/10 \leq \beta \leq \alpha/10$, and γ is kept constant $\gamma = 0$. Note that positive and negative γ shift the entire quantile distribution up and down respectively. As shown in Eq.s 3 and 5, the second-order Rayleigh-Stokes model is a three-parameter probability distribution. In the following section the empirical estimation of the model parameters are discussed in detail.

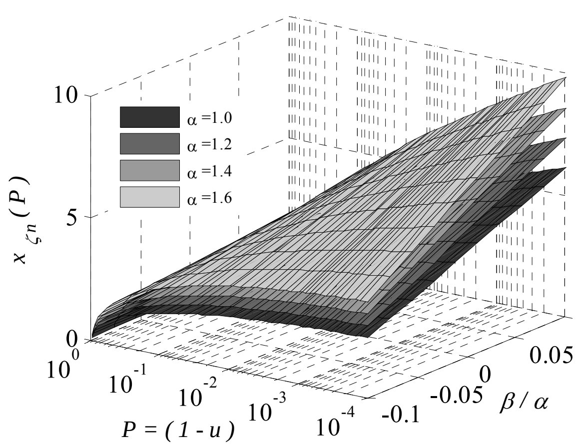


Fig. 1, Rayleigh-Stokes quantile distribution.

EMPIRICAL PARAMETER ESTIMATION

In this study the three parameters of the second-order Rayleigh-Stokes model is obtained from application of two moment-based parameter estimation methods, i.e. conventional method of moments (MoM) and method of L-moments (MoLM) (Hosking and Wallis 1997).

Method of Moments

Distribution moments have been widely used to characterize the probability distributions and these statistics are applied in method of moments to obtain estimates of the probability distribution model parameters. For a random variable X, the first moment, i.e. mean, is defined as (Hosking and Wallis 1997)

$$\mu_1(X) = E(X) \tag{7}$$

and the nth moment is defined as

$$\mu_n(X) = E(X-\mu)^n \tag{8}$$

where, $E(g(X))$ is the expectation of the function $g(X)$ and is obtained from

$$E(g(X)) = \int_0^1 g(x(u))\,du \tag{9}$$

From that, the distribution moments can be defined in the form of,

$$\begin{aligned} \mu_1(X) &= \int_0^1 x(u)\,du && n=1 \\ \mu_n(X) &= \int_0^1 \left(x(u)-\mu_1\right)^n du && n>1 \end{aligned} \tag{10}$$

The mean $\mu_1(X)$ represents the centroid of the distribution and the variance $\sigma_X{}^2 = \mu_2(X)$ is a measure of the distribution dispersion around its center. Other useful moments are the dimensionless third and fourth moments, respectively called skewness $s_X = \mu_3(X)/\sigma_X{}^3$ and coefficient of excess kurtosis $K_X = \mu_4(X)/\sigma_X{}^4 - 3$. Utilizing the quantile function of the Rayleigh-Stokes model given in Eq.s (3) and (6) in the integrals in Eq. (10), the relation between the first three distribution moments and the model parameters are derived in the form of

$$\begin{aligned} \mu_1(\zeta_n) &= \gamma + 2\beta + \alpha\sqrt{2}\,\Gamma(3/2) \\ \mu_2(\zeta_n) &= 4\beta^2 + \alpha\beta(2)^{3/2}\,\Gamma(3/2) + 2\alpha^2(1-\pi/4) \\ \mu_3(\zeta_n) &= 16\beta^3 + 9\sqrt{2}\,\alpha\beta^2\,\Gamma(3/2) + 12\alpha^2\beta\left(1-\Gamma^2(3/2)\right) \\ &\quad + \alpha^3(2)^{3/2}\left(2\Gamma^3(3/2) - \frac{3}{2}\Gamma(3/2)\right) \end{aligned} \tag{11}$$

where Γ is the well-known Gamma function. As shown in this equation, the nth moment is a polynomial of degree n of the parameters α and β, and as expected γ is only shifting the distribution mean.

For a data set $x_1, x_2, \ldots, x_{N_s}$ of size N_s, the unbiased estimates of the first three sample moments are obtained respectively from (Hosking and Wallis 1997)

$$\begin{aligned} \hat{\mu}_1(X) &= \frac{1}{N_s}\sum_{i=1}^{N_s} x_i \\ \hat{\mu}_2(X) &= \hat{\sigma}_X{}^2 = \frac{1}{(N_s-1)}\sum_{i=1}^{N_s}\left(x_i-\hat{\mu}_1\right)^2 \\ \hat{\mu}_3(X) &= \frac{N_s}{(N_s-1)(N_s-2)}\sum_{i=1}^{N_s}\left(x_i-\hat{\mu}_1\right)^3 \end{aligned} \tag{12}$$

In order to estimate the three model parameters of the Rayleigh-Stokes probability distribution, α, β, and γ with MoM, the first three distribution moments (Eq. (11)) are equated with their unbiased sample estimators (Eq. (12)), which gives a system of equations to be solved for the unknown parameters. The explicit relation of model parameters and the sample moments are not available and in order to obtain the solution an iterative numerical solver is required.

Method of L-Moments

Linear moments (L-moments) are developed from modifying the probably-weighted moments (PWM) formerly introduced by Greenwood et al. (1979). The main difference between ordinary moments and PWMs is that ordinary moments give greater weight to the extreme tails of the distribution. Therefore, distribution moments are more successful in representing the extreme values. However, the sample moments are highly affected by unexpectedly large observations and consequently high order moments are considerably more biased than the corresponding probability-weighted moments. PWMs are considered the desirable sample estimators for extreme analysis and when the sample sizes are limited.

PWMs, as alternatives to ordinary moments, have been used in the field of probability distribution parameter estimation (e.g. (Hosking, Wallis, and Wood 1985 and Hosking and Wallis 1987). However, it is difficult to directly connect PWMs to the characteristics of the probability distribution, e.g. shape and scale. Hosking introduced L-moments from a linear combination of PWMs to overcome this issue. For a random variable X with quantile function of $x(u)$, the distribution L-moments are obtained from integration of quantile function multiplied by an orthogonal function (Hosking and Wallis 1997), specifically

$$\lambda_n(X) = \int_0^1 x(u)\,P^*_{n-1}(u)\,du \tag{13}$$

where the orthogonal function is the shifted Legendre polynomials $P^*{}_n(u)$

$$P^*_n(u) = \sum_{k=0}^{r} p^*_{n,k}\,u^k \tag{14}$$

and the function coefficients are obtained from

$$p^*_{n,k} = \frac{(-1)^{n-k}(n+k)!}{(k!)^2(n-k)!} \tag{15}$$

By definition, λ_1 is the L-location or mean of the distribution, λ_2, $\tau_3 = \lambda_3/\lambda_2$ are L-scale, and L-skewness respectively, and are analogous to the ordinary standard deviation and skewness.

A more complete definition of the linear moments and their characteristics can be found in (Hosking and Wallis 1997).

Applying the quantile function of Rayleigh-Stokes model, Eq.s 3 and 6, in definition of linear moments, Eq. (13), the relations between the distribution moments and the model parameters are derived as

$$\begin{aligned}\lambda_1(\zeta_n) &= \gamma + 2\beta + \alpha(2)^{1/2}\,\Gamma(3/2)\\ \lambda_2(\zeta_n) &= \beta + \alpha\left(2^{1/2}-1\right)\Gamma(3/2)\\ \lambda_3(\zeta_n) &= \beta/3 + \alpha\left(2^{1/2}-3+(8/3)^{1/2}\right)\Gamma(3/2)\end{aligned} \tag{16}$$

As shown here, L-moments are linear function of parameters α and β and the shifting parameter γ only appeared in the L-location. The sample L-moment l_n of an ordered sample $x_{1:N_s} \le x_{2:N_s} \le \dots \le x_{N_s:N_s}$ of size N_s is defined as (Hosking and Wallis 1997)

$$l_{n+1}(X) = \sum_{k=0}^{n} p_{n,k}^{*}\,\varsigma_n \qquad n = 0,1,\dots,N_s-1 \tag{17}$$

where

$$\varsigma_n = N_s^{-1}\binom{N_s-1}{n}^{-1}\sum_{j=n+1}^{N_s}\binom{j-1}{n}x_{j:N_s} \tag{18}$$

and the brackets denotes binomial coefficients. Note that, the sample L-moment l_n is an unbiased estimator of the distribution L-moments λ_n. Equating the first three distribution linear moments with their corresponding sample statistics, the estimates of Rayleigh-Stokes model parameters are obtained in the form of

$$\begin{aligned}\hat{\alpha} &= 4.1394\left(l_2(\zeta_n) - 3l_3(\zeta_n)\right)\\ \hat{\beta} &= \left(l_2(\zeta_n) - \left(2^{1/2}-1\right)\hat{\alpha}\,\Gamma(3/2)\right)\\ \hat{\gamma} &= l_1(\zeta_n) - 2\hat{\beta} - \hat{\alpha}(2)^{1/2}\,\Gamma(3/2)\end{aligned} \tag{19}$$

EXTREME STATISTICS

The main purpose for including the effect of non-linearity in the probability distribution is to obtain more reliable estimation of the extreme statistics. Rayleigh-Stokes probability distribution in conjunction with the extreme value theory has been used in many research studies (Nerzic and Prevosto 1997, Dawson 2000, Prevosto and Krogstad 2000, and Krogstad and Barstow 2004) to estimate the ocean wave crests extreme statistics, e.g. expected crest maximum. In this approach, it is assumed that crests are independent random variables which is not theoretically well justified. However, the results of simulations by Krogstad and Barstow (2004) indicate that the approximation is reasonable accurate for large number of waves $N > 100$.

Assuming that crests ζ_n are independent identically distributed (i.i.d) random variables with PDF f_{ζ_n} and CDF F_{ζ_n}, the PDF and CDF of the crest maxima $\zeta_{\max}$ in N waves can be obtained from the ordered value statistics theory, specifically, (Leadbetter, Lindgren, and Rootzen 1983)

$$\begin{aligned}f_{\zeta_{\max}}(x) &= N\,f_{\zeta_n}(x)\left[F_{\zeta_n}(x)\right]^{N-1}\\ F_{\zeta_{\max}}(x) &= \left[F_{\zeta_n}(x)\right]^{N}\end{aligned} \tag{20}$$

For large number of waves N, it can be shown that the Rayleigh-Stokes probability distribution belongs to the Gumbel maximal domain of attraction and the asymptotic form of $x_{\zeta_{\max}}(u)$ can be represented by

$$x_{\zeta_{\max}}(u) = a_N - b_N\ln\left(-\ln(u)\right) \tag{21}$$

where a_N and b_N are the extreme distribution parameters and can be estimated from their relation with Rayleigh-Stokes parameters

$$\begin{aligned}a_N &= \gamma + 2\beta\ln(N) + \alpha\left(2\ln(N)\right)^{1/2}\\ b_N &= 2\beta + \alpha\left(2\ln(N)\right)^{-1/2}\end{aligned} \tag{22}$$

Assuming that $\zeta_{\max}$ follows the Gumbel probability distribution, the first three moments of $\zeta_{\max}$ are obtained in the form of,

$$\begin{aligned}\mu_1(\zeta_{\max}) &= E(\zeta_{\max}) = a_N + b_N\gamma_{EM}\\ \mu_2(\zeta_{\max}) &= \frac{\pi^2}{6}b_N^2\\ \mu_3(\zeta_{\max}) &= 2\zeta_R(3)b_N^3\end{aligned} \tag{23}$$

where $\gamma_{EM} \approx 0.5772$ is the Euler-Mascheroni constant and $\zeta_R(z)$ is the Riemann zeta function that is $\zeta_R(3) \approx 1.2021$ at $z = 3$. Similarly, the first three L-moments of $\zeta_{\max}$ are derived

$$\begin{aligned}\lambda_1(\zeta_{\max}) &= a_N + b_N\gamma_{EM}\\ \lambda_2(\zeta_{\max}) &= b_N\ln(2)\\ \lambda_3(\zeta_{\max}) &= b_N\left(2\ln(3) - 3\ln(2)\right)\end{aligned} \tag{24}$$

In Fig. 2, the effect of model parameter values on the quantile distribution of Rayleigh-Stokes maxima in $N = 1000$ waves is presented. The distributions in Fig. 2 are estimated from application of Rayleigh-Stokes distribution in Eq. (20). It should be noted that for $\beta < 0$ the Rayleigh-Stokes probability distribution has an upper bound which is considered in the calculation of distributions shown in Fig. 2.

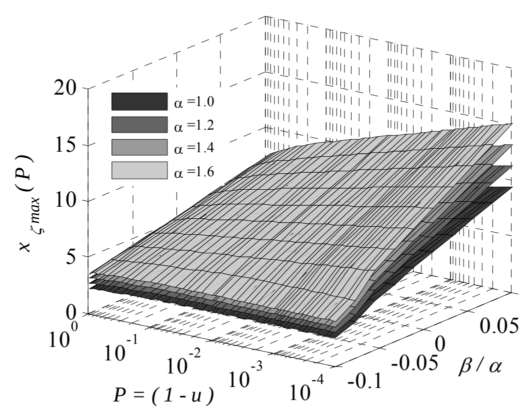


Fig. 2, quantile distribution of maxima in 1000 waves of Rayleigh-Stokes model.

UNCERTAINTY ANALYSIS

To evaluate the uncertainty of empirical parameter estimation and to compare the performance of MoM and MoLM on samples with different sizes, Monte-Carlo simulation technique is utilized. For this purpose, $100,000$ independent samples are generated from the Rayleigh-Stokes quantile with known parameters. The sample moments and linear moments are estimated for each sample and consequently the estimates of model parameters and the statistics are obtained from MoM and MoLM. For the uncertainty analysis purposes, the parameter α is varied in the range of $1.0 \le \alpha \le 1.6$, β is varied in the range of $-\alpha/10 \le \beta \le \alpha/10$, and γ is kept constant $\gamma = 0$. Samples with sizes $N_s = 2000$, 1000, 330, 100 are studied which respectively approximate the number of waves in 6hr, 3hr, 1hr, 17min ocean wave timeseries. Root-mean-squared error (RMSE) is utilized as metric to evaluate the uncertainty in the estimates. By definition, the RMSE of an estimate $\hat{\varphi}$ of the true value φ is estimated from

$$RMSE(\hat{\varphi}) = \left(E\left[(\varphi - \hat{\varphi})^2\right]\right)^{1/2} \tag{25}$$

The RMSE may be normalized by the true value φ to obtain the percentage error.

In Fig.s 3~5, respectively the RMSE distributions of estimates of α, β, and γ from both MoM and MoLM are compared. For all the three model parameters the RMSE of both methods varies linearly with α and therefore normalizing the error with α results in an identical distribution for different values of α. As expected and also shown in Fig.s 3~5, the error increases with the decrease in the sample size. The change in the error distributions from $N_s = 2000$ to $N_s = 1000$ is relatively mild while the error distributions for $N_s < 1000$ indicate drastic increases. The error of the estimates of MoM and MoLM for large samples $N_s > 1000$ is reasonably close. However, the difference between the error distributions is more sensible for smaller samples $N_s \le 330$. It is also observed that MoM is more sensitive to the sample size. The distributions in Fig.s 3~5 clearly indicate that error distributions of MoLM is a linear function of parameter β while that of MoM is a non-linear function of β. In general, MoM performs better for samples with negative β, while MoLM has smaller errors for samples with positive β. Regarding the distributions in Fig. 3, one can see that $RMSE(\alpha)$ is highly dependent to the value of β and the dependency increases with decrease in the sample size. The dependency to the value of β is much milder in case of $RMSE(\beta)$ and $RMSE(\gamma)$ shown respectively in Fig.s 4 and 5. As can be seen in Fig. 4, the error in the estimates of β from small samples is almost of the same order of magnitude as of β itself which indicates that large enough samples are required to obtain reliable estimates of the non-linear contribution.

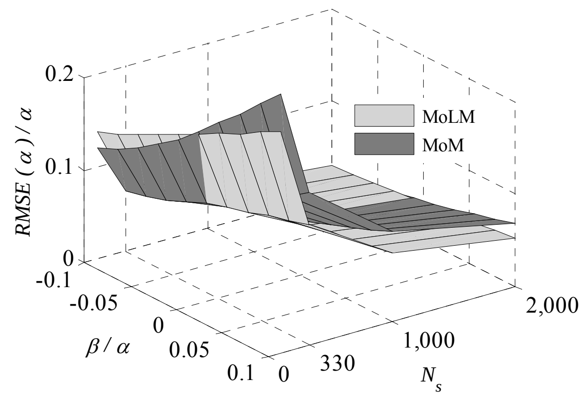


Fig. 3, RMSE distributions of estimates of parameter α.

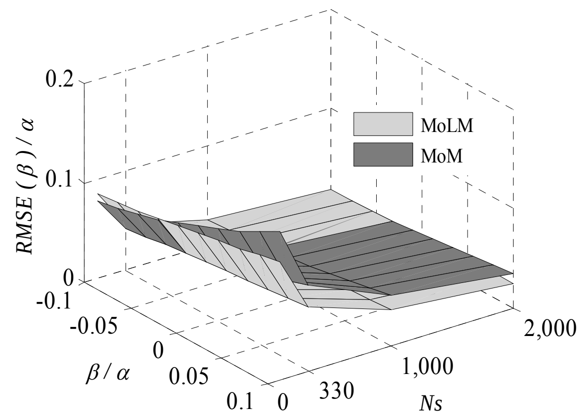


Fig. 4, RMSE distributions of estimates of parameter β.

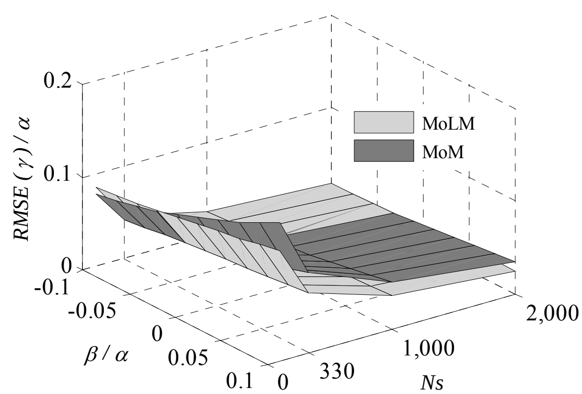


Fig. 5, RMSE distributions of estimates of parameter γ.

Utilizing the parameter estimates from MoM and MoLM in conjunction with the Eq. (23) and the extreme parameters Eq. (22), the estimates of expected crest maximum in $N = 1000$ waves are estimate and the error distributions are presented in Fig. 6. In this figure, the distributions are normalized with respect to the true $E(\zeta_{\max})$. The normalized RMSE distributions for different values of α converge to an identical distribution and as shown in Fig. 6, the dependency of the distributions to the value of β is relatively small. Estimates of both methods indicate an increase in the normalized error with the increase in the value of β which is more sensible in the case of MoM and smaller samples. More importantly, it is observed that MoLM is more efficient in estimating the extreme statistics. The difference between the performance of MoM and MoLM is relatively small for large samples $N_s \geq 1000$ while MoLM significantly performs better for small samples. For the studied range of parameters, the error in the estimates of $E(\zeta_{\max})$ in $N = 1000$ waves obtained from MoLM remains less than 4.5 percent of the true value for $N_s = 1000$ and the error remains less than 14 percent of the true value for $N_s = 100$. The extreme estimates of MoM are expected to have errors in the range (4.2-6.7) and (14 - 20) percent of the true value for $N_s = 1000$ and $N_s = 100$ respectively.

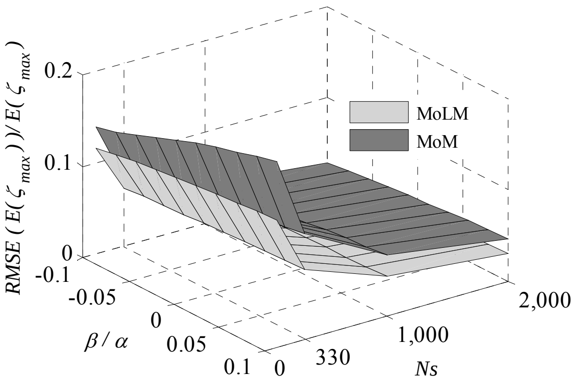


Fig. 6, RMSE distributions of estimates of parameter $E(\zeta_{\max})$ in $N = 1000$ waves.

ANALYSIS OF MEASURED DATA

Here, the performance of the Rayleigh-Stokes model with parameters estimated by MoM and MoLM are compared on measured data. The data sets utilized here are obtained from a mini-TLP model test investigating the behavior of the structure in extreme environment. The model tests were performed in the wave basin at Offshore Technology Research Center (OTRC). Details of this experiment can be found in the articles by Niedzwecki et al. (2001) and Teigen, Niedzwecki, and Winterstein (2001). In this study only the measurements of a 3-hr seastate generated from JONSWAP spectrum with significant wave height $H_s = 4.0m$, peak period of $T_p = 16s$ and peakedness factor $\gamma_s = 2.0$ are analyzed. The seastate represents the 100-year design wave conditions off the West Africa coastline for which the mini-TLP was originally designed. The Rayleigh-Stokes model is utilized to estimate the probability distribution of wave crests at a) over the front pontoon A2 and b) at the center of the moon-pool beneath the structure A3 (see Fig. 7 for probes location). In the studied cases, the long crested waves are attacking the structures in the heading seas direction. The wave crest is defined as the maximum wave elevation in between two successive zero-upcrossings and the wave crest samples are obtained by zero-crossing analysis of measured wave timeseries. The wave crest samples are normalized by the first order standard deviation of incident waves σ_η that is estimated from its relation with the significant wave height $\sigma_\eta = H_s/4 = 1.0$.

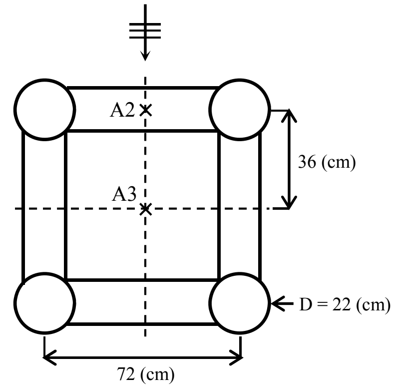


Fig. 7, wave probe location (dimension scale 1/40).

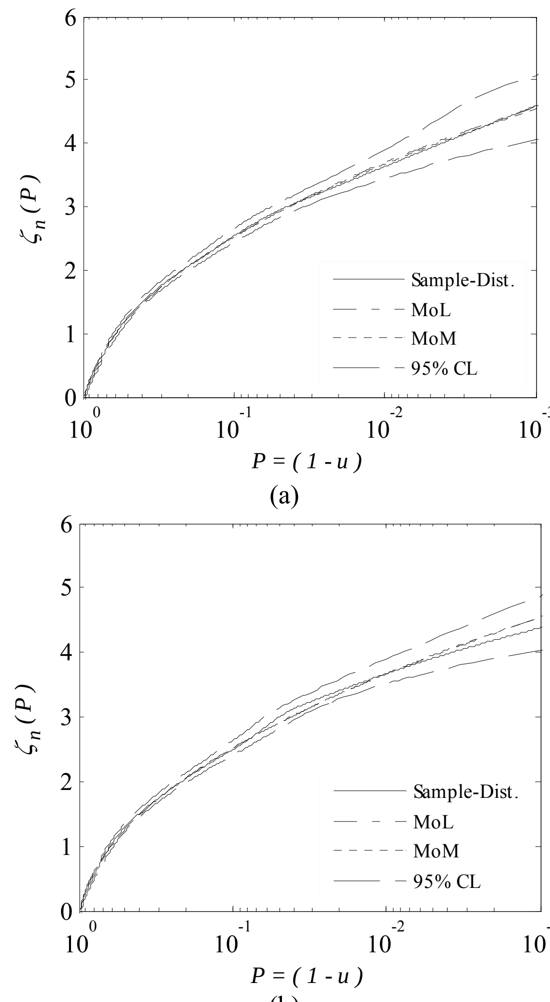
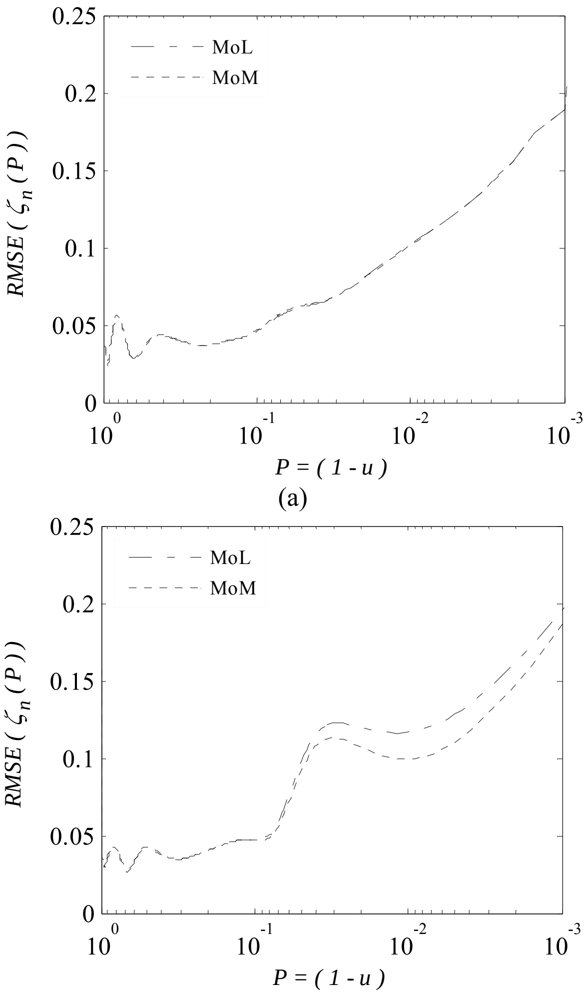


Fig. 8, Quantile distributions, (a) wave crests at A2, (b) wave crests at A3.

In Fig. 8, the sample quantile distributions are compared with those of the Rayleigh-Stokes distribution with the parameters estimated from MoM and MoLM. The sample distribution is estimated from application of semi-parametric approach in which the major part of the distribution is estimated by Kernel non-parametric probability estimation method and the tail distribution is modified with generalized-Pareto distribution. The sample distribution is then utilized in bootstrap analysis to obtain the 95 percent confidence limits (CL) and the estimates of the RMSE. More details about the semi-parametric probability estimation and bootstrap analysis can be found in the articles by Efron 1979, Silverman and Young 1987, and Caers and Maes 1998. As shown in Fig. 8, the Rayleigh-Stokes model is successful in capturing the probability distribution of data and the estimates of MoM and MoLM converge considerably well.



(a)

(b)

Fig. 9, RMSE distribution of quantile estimates for, (a) wave crests at A2, (b) wave crests at A3.

In Fig. 9, the RMSE of the quantile distributions of Rayleigh-Stokes models are presented. The RMSE distributions are estimated from utilization of 20,000 bootstrap samples. As shown in Fig. 9, the error in the estimates of Rayleigh-Stokes models remain below 0.2 (equivalent to 0.2m) in these example which once more indicates that Rayleigh-Stokes model is robust in estimating the probability distribution of data. The distributions in Fig. 9 indicate a reasonable agreement between the estimates of MoM and MoLM. The difference between estimates of MoM and MoLM is more sensible on the tail distribution of the wave crests at A3 (see Fig. 9 (b)). As expected, the RMSE of the quantile estimates increases on the tail distribution which is mainly because of sparse data points in this area. The sample distribution in Fig. 8 (b) indicates a sudden slope decrease in the tail distribution with probabilities of $P \leq 8\times10^{-2}$ which has not been fully captured by the Rayleigh-Stokes model. Consequently, the RMSE distributions in Fig. 9 (b) show some unexpected fluctuations for the probabilities $P \leq 8\times10^{-2}$.

The estimates of expected crest maximum in $N = 1000$ waves are compared in Table 1. The MoM and MoLM are estimated almost identical extreme statistics for the studied cases. Based on the statistics given in Table 1, the maximum wave crest in $N = 1000$ waves close to the structure is expected to be $E(a_{\max}) \approx 4.72m$ which is 1m (~26 percent) higher than that of the linear ocean waves $E(a_{\max}) \approx 3.72m$. This shows the importance of including the non-linear terms as well as the contribution of diffracted and radiated waves in estimation of the air-gap demand.

Table 1, estimates of expected maximum crest in $N = 1000$ waves.

Method	Crests at A2	Crests at A3
MoLM	4.76	4.72
MoM	4.72	4.73

SUMMARY AND CONCLUSIONS

Three-parameter Rayleigh-Stokes model was studied as an semi-empirical probability distribution in which the structural form was developed analytically from the second-order Stokes wave theory of weakly non-linear and narrow-banded processes and the estimates of the underlying model parameters were obtained empirically. For empirical parameter estimation purposes, two moment-based models, i.e. method of moments (MoM) and method of L-moments (MoLM) were utilized. The relations between the distribution moments and model parameters were derived. Consequently the relations were used to directly connect the model parameters to the sample statistics. The application of Rayleigh-Stokes model in extreme analysis was briefly discussed and the asymptotic form of the crest maxima distribution and statistics for large number of waves are derived.

The uncertainty of the model statistics estimated from MoM and MoLM was evaluated and the performance of these methods on samples with different sizes was compared. For this purpose, numerous Monte-Carlo samples were generated from Rayleigh-Stokes quantile with various sets of initial parameter values. The results of the uncertainty analysis indicated that the MoM and MoLM perform similarly well for relatively large samples $N_s \geq 1000$, while method of L-moments found to be the better alternative for small samples. It was observed that the root-mean-squared error distributions of the model parameters α,

β, and γ estimated from MoM and MoLM are linear functions of the parameter α. In the case of MoLM, the error distributions of parameter estimates are linear functions of parameter β, while those of MoM vary nonlinearly with β. The parameter estimates of MoM found to have less error for $\beta < 0$ than those of MoLM while the error of the parameter estimates of MoLM were less $\beta > 0$. It was concluded that, in order to have reasonably accurate estimates of the non-linear term, large enough $N_s > 330$ samples are required for both methods. The uncertainty analysis of extreme statistic indicated that MoLM is more robust than MoM in estimating the extreme statistics.

In the final section of this article, Rayleigh-Stokes model was utilized to estimate the probability distribution of wave crests beneath a mini-TLP model test. The data sets used here were measured during a relatively benign environmental condition in which wave timeseries can be approximated as weakly non-linear processes. It was observed that Rayleigh-Stokes model with empirically estimated parameters is considerably accurate in estimating the probability distribution of complex wave crests over the pontoon and inside the moon-pool. In the studied cases, the estimates of MoM and MoLM are found to be in a close agreement for the major part of the distribution, while the difference between their estimates of the tail distribution is more sensible. It should be noted that, the Rayleigh-Stokes model is not sensitive to the local changes of the tail distribution which causes uncertainty to the model estimates of extreme statistics.

Rayleigh-Stokes model was utilized to estimate the extreme wave crests in the vicinity of the Mini-TLP during the 100-year design wave conditions off the West Africa coastline. The expected maximum wave crest elevation in the vicinity of the structure was calculated to be about 4.72m which is significantly higher than the expected maximum crest elevation of linear ocean waves 3.72m. The results of this study, once more, confirmed that the contribution of non-linear term and the wave-structure interaction effects should be considered in the deck elevation design of offshore platforms.

REFRENCES

Arhan M, Plaisted, RO (1981). "Non-Linear deformation of sea wave profiles in intermediate and shallow water", *Oceanologica Acta*, Vol 4, No 2, pp 107-15.

Caers J, Maes MA (1998). "Identifying tails, bounds and end-points of random variables", *J of Structural Safety*, Vol 20, pp 1-23.

Dawson TH (2000). "Maximum wave crests in heavy seas", *J of Offshore Mechanics and Arctic Engineering*, OMAE, Vol 122, pp 222-4.

Efron B (1979). "Bootstrap methods: Another look at the jackknife", *The Annals of Statistics*, Vol 7, No1, pp 1-26.

Fedele F, Arena F (2005). "Weakly nonlinear statistics of high random waves", *J of Physics of Fluids*, Vol 17, No 1, pp 1-10.

Greenwood JA, Landwehr JM, Matalas NC, Wallis JR (1979). "Probability weighted moments: Definition and relation to parameters of several distributions expressible in inverse form", *J of Water Resources Research*, Vol 15, No 5, pp 1049-54.

Hosking JRM (1990). "L-moments: Analysis and estimation of distributions using linear combinations", *J of the Royal Statistical Society. Series B Methodological*; Vol 52, No 1, pp 105-24.

Hosking, JRM, Wallis JR (1987). "Parameter and quantile estimation for the generalized Pareto distribution", *Technometrics*, Vol 29, No 3, pp 339-49.

Hosking JRM, Wallis JR (1997). *Regional frequency analysis an approach based on L-moments*. Cambridge University Press.

Hosking JRM, Wallis JR, Wood EF (1985). "Estimation of the generalized extreme-value distribution by the method of probability-weighted moments". *Technometrics*, Vol 27, No 3, pp 251-61.

Izadparast AH, Niedzwecki JM (2009a). "Estimating wave crest distributions using the method of L-moments", *J of Applied Ocean Research*, Vol 31, pp 37-43.

Izadparast AH, Niedzwecki JM (2009b). "Probability distributions for wave run-up on offshore platform columns", *Proc 28th Int Conference of Ocean, Offshore and Arctic Engineering*, Honolulu, USA, OMAE, pp 607-15.

Izadparast AH, Niedzwecki JM (2010). "Probability distributions of wave run-up on a TLP model", *J of Marine Structures*, Vol 23, No 2, pp 164-86.

Kriebel DL (1993). "Non-linear run-up of random waves on a large cylinder". *Proc. 12th Int Conference of Ocean, Offshore and Arctic Engineering*, Glasgow, U.K., OMAE, Vol 1. pp 49-55.

Kriebel DL, Dawson TH (1991). "Nonlinear effects on wave groups in random seas", *J of Offshore Mechanics and Arctic Engineering*, Vol 113, pp 142-7.

Kriebel DL, Dawson TH (1993). "Nonlinearity in wave crest statistics", *Proc 2nd Int Symposium on Ocean Wave Measurement and Analysis*, pp 61-75.

Krogstad HE, Barstow SF (2004). "Analysis and applications of second-order models for maximum crest height", *J of Offshore Mechanics and Arctic Engineering*, OMAE, Vol 126, pp 66-71.

Leadbetter MR, Lindgren G, Rootzen H (1983). *Extremes and related properties of random sequences and processes*, Springer Series in Statistics, Springer-Verlag.

Longuet-Higgins MS (1952). "On the statistical distribution of the heights of sea waves", *J of Marine Research*, Vol 11, No 3, pp 245-66.

Nerzic R, Prevosto MA (1997). "Weibull-stokes model for the distribution of maximum wave and crest heights", *Proc. 7th Int Offshore and Polar Engineering Conference*, Honolulu, USA, ISOPE, Vol 3. pp 367-77.

Niedzwecki JM, Liagre P-YJ, Roesset JM, Kim MH, Teigen P. "An experimental research study of a Mini-TLP", *Proc 11th Int Offshore and Polar Engineering Conference*, Stavanger, Norway, ISOPE, pp 631-4.

Prevosto M, Krogstad HE, Robin A (2000). "Probability distributions for maximum wave and crest heights", *J of Coastal Engineering*, Vol 40, pp 329-60.

Tayfun MA (1980). "Narrow-band nonlinear sea waves", *J of*

Geophysical Research, Vol 85, No C3, pp 1548-52.
Tayfun MA (2006). "Statistics of nonlinear wave crests and groups", *J of Ocean Engineering*, Vol 33, pp 1589-1622.
Teigen PS, Niedzwecki JM, Winterstein SR (2001). "Wave interaction effects for non-compliant TLP", *Proc 11th Int Offshore and Polar Engineering Conference*, Stavanger, Norway, ISOPE, pp 453-61.
Tung CC, Huang NE (1985). "Peak and trough distribution of nonlinear waves", *J of Ocean Engineering*, Vol 12, No 3, pp 201-9.
Silverman BW, Young GA (1987). "The bootstrap: To smooth or not to smooth?", *Biometrika,* Vol 74, No 3, pp 469-79.
Stansberg CT, Nielsen FG (2001). "Non-linear wave-structure interactions on floating production systems", *Proc of 11th Int Offshore and Polar Engineering Conference*, Stavanger, Norway, ISOPE, pp 363-372.