

**UNCERTAINTIES IN EXTREME VALUE ANALYSIS  
AND THEIR EFFECT ON LOAD FACTORS**

**J.G. Rutten<sup>‡</sup>**

Delft University of Technology  
Offshore Engineering  
Stevinweg 1  
2628 CN Delft  
The Netherlands

**K.C. Ewans**

Shell International Exploration and Production  
P.O. Box 60  
2280 AB Rijswijk  
The Netherlands  
Kevin.Ewans@shell.com

**P.H.A.J.M. van Gelder**

Department of Civil Engineering  
TU Delft  
Stevinweg 1  
2628 CN Delft  
P.H.A.J.M.vanGelder@CiTG.TUdelft.nl

**M. Efthymiou**

Shell International Exploration and Production  
P.O. Box 60  
2280 AB Rijswijk  
The Netherlands  
Mike.Efthymiou@shell.com

**ABSTRACT**

A method for deriving site-specific load factors for fixed offshore structures that include uncertainties in extreme values is described. The method consists of an extreme value analysis, to derive the extreme value behaviour and the uncertainties involved, and a model, based on the Load Resistance Factor Design of the API rules, to derive site-specific load factors with or without the inclusion of uncertainties in the long term load distribution. The application of site-specific load factors could lead to a harmonised calculation of the probability of failure for platforms around the world.

The extreme value analysis that is applied is a threshold based Generalized Pareto Distribution in combination with Maximum Likelihood Estimation. This is applied to a Peaks-over-Threshold data selection from hindcast studies. The Extreme Value Analysis is used to derive four extreme value characteristics (shape and scale parameter, the threshold and the number of data points,) that are used to quantify the extreme value behaviour. The North Sea, West African Coast and Gulf of Mexico are analysed. These are three ocean basins that capture the main characteristic climates around the world: extra-tropical storms, swell dominated extremes and tropical storms, respectively.

The uncertainties in the extreme value analysis are dependent on the four characteristics. A simulation study is used to extract the confidence intervals of the parameters (shape, scale and significant wave height) as a function of the characteristics. The simulation of extreme value data sets gives a bias in the results, and therefore a practical method is proposed to correct for the bias and to extract confidence intervals. The bias correction is used to derive the shape and size of the confidence intervals for the extreme significant wave heights at the North Sea, West African Coast and Gulf of Mexico.

The load factors are computed on the basis of the Reserve Strength Ratio requirement. Only wave loads are incorporated in the determination of the loading on a fixed platform; the current and wind loading are not considered. Three computation methods for load factors are investigated; a traditional approach without the implementation of uncertainties, and two new methods that include the uncertainties derived with simulation studies. The three methods show that there are regional differences in load factors and that the implementation of uncertainties leads to an increase in the factors.

**INTRODUCTION**

---

<sup>‡</sup> Author performed work at Shell EP Project as part of a TU Delft Msc thesis. He is currently affiliated with BP Exploration Ltd and can be contacted at job.rutten@bp.com

One of the major issues with regard to the design of offshore structures is the determination of reliable metocean criteria (wind, waves and current). Metocean criteria are derived from data gathered by field measurements and hindcast studies and are used for the computation of the extreme environmental loading on a platform. Due to the limited amount of data records, there are uncertainties in, for instance, the determination of the 100 and 1000-year wave heights, winds and currents. These uncertainties can be translated into uncertainties in the corresponding maximum loading on a platform. The partial load factor of 1.35<sup>1</sup>, currently used for the design of fixed offshore structures in all ocean basins should capture these uncertainties. However the use of such a common load factor implies that the uncertainties in the metocean design criteria are assumed equal for all ocean basins.

For ocean basins around the world where the offshore industry is active, there is a considerable difference in the characteristics and knowledge of the oceanographic conditions. However, this is currently not reflected in the design of offshore structures since one common load factor is applied in all ocean basins. A better understanding of the wave climate and the uncertainties in the extreme wave conditions could allow for site-specific load factors for which this paper gives a proposal. This could be used to establish a globally harmonised probability of failure for offshore structures.

## NOMENCLATURE

EVA = extreme value analysis  
 cdf = cumulative distribution function  
 pdf = probability density function  
 GPD = generalized pareto distribution  
 MLE = maximum likelihood estimator  
 POT = peaks over threshold  
 CI = confidence interval  
 Hs = significant wave height  
 $\gamma$  = shape parameter  
 $\gamma^*$  = simulated shape parameter (biased)  
 $\gamma_{bc}$  = simulated shape parameter bias corrected  
 $\sigma$  = scale parameter  
 nov = number of values  
 Q100 = 100-year wave height  
 Q1000 = 1000-year wave height  
 LRFD = load resistance factor design  
 GLM = Generic Load Model  
 RSR = reserve strength ratio  
 BPI = bayesian predictive integral  
 DLF = distributed load factor  
 $\gamma_E$  = environmental load factor  
 $E_{100}$  = the 100 year environmental loading  
 $E_{2000}$  = the 2000 year environmental loading  
 $E_{33,000}$  = the 33,000 year environmental loading

## EXTREME VALUE ANALYSIS

### Theory

Fisher and Tippet [1928] were among the first to develop the extreme value theory with the derivation of the asymptotic distribution of the sample maxima. If applied to wave heights, a threshold is chosen in the record and only values above are used in the extreme value analysis. The modelling of these extreme waves is performed with a Generalized Pareto Distribution (GPD) theory as shown by Pickands [1975].

Let  $X_1, X_2, \dots, X_n$  be a series of independent random observations of a random variable  $X$  with the distribution function  $F(x)$ . To model the upper tail of  $F(x)$ , consider  $k$  exceedances of  $X$  over a threshold  $u$  and let  $Y_1, Y_2, \dots, Y_k$  denote the excesses (or peaks), i.e.,  $Y_i = X_i - u$ . Pickands [1975] showed that, in some asymptotic sense, the conditional distribution of excesses follow the generalized Pareto distribution. Thus the distribution function of  $Y_i = [(X_i - u) | X_i > u]$ ,  $i = 1, 2, \dots, k$ , is given as

$$GPD(y) = 1 - \left( 1 + \frac{\sigma(y-u)}{\gamma} \right)^{-1/\gamma} \quad 1$$

where  $u$ ,  $\sigma$  and  $\gamma$  denote the location, scale and shape parameters, respectively. In this paper the location parameter is taken as zero. The distribution is unbounded, i.e.,  $0 < y < \infty$  if  $\gamma \geq 0$  and bounded as  $0 < y < \sigma/\gamma$  if  $c < 0$ .

The parameter selection for fitting a GPD cdf to a dataset is a critical process in the extreme value analysis. This is the modelling step between measurements and the artificial representation of the data with a distribution function. The method used in this study is the Maximum Likelihood Estimator (MLE) used by Smith [1987]. The estimation of the shape and scale parameters from the likelihood function is done with the conjugate gradient method.

The EVA is applied to the datasets created based on the selection of Peaks over Threshold (POT). The selection of the threshold is not a straightforward procedure. Based on the data itself, it is difficult to determine which threshold (and respective number of values) gives the best fitting result. The threshold is determined based on the goodness-of-fit to the data. The relationship between the threshold and the number of values is dependent on a data set. For a method applicable to many data sets, as developed in Rutten [2003], both characteristics are needed.

The Extreme Value Analysis describes the extreme value behaviour with four characteristics: the shape and scale parameter of the GPD, the threshold and the number of values in the dataset. With this parameterisation the uncertainties and load factors are derived in section "Uncertainties in Extreme Value Analysis" and section "Effect of uncertainties on load factors".

<sup>1</sup> API Recommended Practise-2A LRFD (1933)

## North Sea, West Africa and Gulf of Mexico<sup>2</sup>

### North Sea

Omni directional analysis results in extreme value datasets created, regardless of the direction of the waves. However, in the North Sea data from different directions belong to different populations, and mixing these is not valid in statistical analysis. Consequently a directional analysis, to study the influence of directionality on the behaviour of extreme events, was undertaken for the North Sea. Based on the direction of the extreme events, the data have been sectorised in different directional sectors.

### West African Coast

All extreme values in the hindcast data are generated by swell propagated from the southern Atlantic Ocean. There are two types of datasets, an operational hindcast of 15 years continuous data and a storm hindcast of swell events in 35 years. The extreme value analysis of both datasets sets showed a remarkable different outcome. Outliers in the operational hindcast corrupted the analysis, but when the outliers were removed, the 15 & 35-year data set matched closely. This points out that short data records should be treated with caution for the extrapolation of the 100-year and 1000-year wave height, regardless of the goodness of fit of the data.

### Gulf of Mexico

The Gulf of Mexico is an extremely difficult region to tackle with an extreme value study, due to the hurricane climate. Performing an extreme value analysis of one location is difficult due to a scarcity in data. A method of site averaging has to be used for the generation of an extreme value dataset. In Rutten [2003] an area selection has been made from which the highest wave record is taken, generated by each hurricane that travelled through the area in the past century. This results in an independent dataset of 80 points that can be used for an EVA, and the analysis does produce a good statistical fit. However, the dataset does not fulfil to the extreme value analysis assumptions of a common population and a sufficiently high threshold.



Figure 1 North Sea (1), Gulf of Mexico (2) and West Africa (3)

<sup>2</sup> This research is performed for Shell and for confidentiality of the data, the hard numbers of the extreme value analysis cannot be presented here.

## UNCERTAINTIES IN EXTREME VALUE ANALYSIS

### Setup of Simulation Study

The construction of confidence intervals for the parameters and return period significant wave heights can be performed by an analytical approach, on the basis of the Fisher Information Matrix, or a simulation study, with a parametric bootstrap (as also proposed by Pandey et al. [2003]). The CIs of the 100-year return period value (Q100 or 100-year quantile) derived from both methods are shown in Figure 2, the vertical axis shows the CI interval as a percentage of the mean value. There are two issues that raise questions on the analytical approach - the behaviour for a shape parameter around zero, and the contour of the confidence intervals.

The analytical CIs are derived under the assumption that they are normally distributed. For shape parameters close to zero, the analytical confidence intervals asymptotically tend to  $\infty$  as shown by Rutten [2003]. The parametric bootstrap does not confirm the asymptotic behaviour and does not support the normal distribution but gives stable CIs. Therefore the parametric bootstrap is used for the construction of CIs.

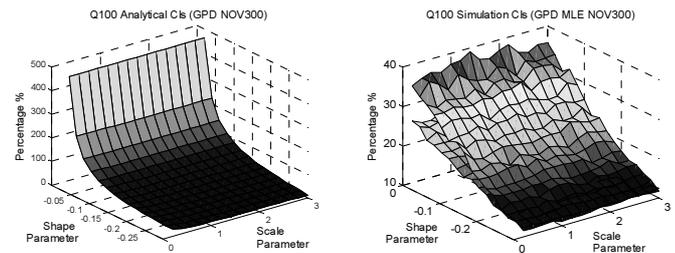


Figure 2 Analytical and Parametric Bootstrap confidence intervals

The parametric bootstrap is performed in a simulation study to derive confidence intervals for the various parameters of the Ocean Basins. In order to gather information about the confidence intervals as a function of the parameters (shape, scale and nov), this study is not limited to the parameter combinations found in the EVA study but utilizes a wide range of parameters. The range is shown in table 1.

Parameter	Min Value	Step Size	Max Value
Shape ( $\gamma$ )	-0.3	0.02	0
Scale ( $\sigma$ )	0 <sup>+</sup>	0.2	3
Nov	100	100	500

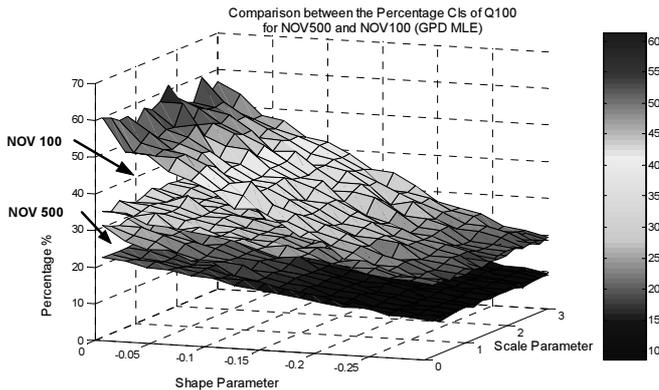
Table 1 Parameter Range of Simulation Study

With the specified parameters, 2500 datasets are generated and refitted with the GPD cdf and MLE parameter estimator. This results in a collection of 2500 sets of estimated shape and scale parameters and the corresponding values for the 10, 100, 2,000 and 33,000-year significant wave height.

## Results

### Simulation Study Variables

The purpose of the simulation study is the extraction of CIs for the extreme return periods. An overview of the 100-year return period main characteristics with respect to the simulation variables is given in figure 3. The axis show the shape and scale parameter, the surfaces show the upper and lower limit of the confidence interval for 100 and respectively 500 number of values.



**Figure 3 Simulation Study Variables**

The number of values is the first parameter that determines the size of the CI. The more data available, the better the extreme value behaviour is described. In the parametric bootstrap this relation is proven; the uncertainties decrease with an increase of the number of values.

The shape parameter is the most important parameter for the description of the GPD cdf. This parameter describes the behaviour of the tail of the distribution function. For shape parameters close to zero, the values are high and the distribution is steep. This results in a wider spreading of the extremes for shape parameters close to zero; this leads to larger CIs. When the shape parameter is close to  $-0.3$  the values are modest and the distribution is flat and the CI is smaller.

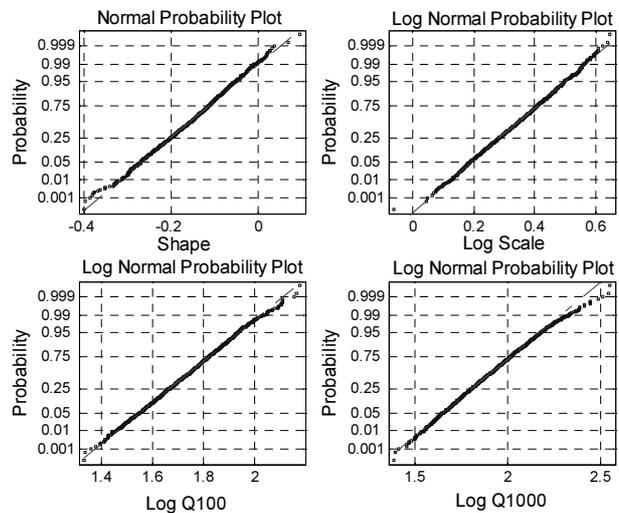
The size of the confidence interval is hardly dependent on the scale parameter for a given fixed shape parameter; this is shown in figure 3.

### Distribution of uncertainties

The simulation study provides 2500 simulations of each parameter combination. The histogram of the simulation shows the distribution of the results. If the centre 95% is used, a 95% CI can be constructed. The analytical approach assumes a normal distribution, but the simulation study shows a different distribution for the 95% CIs.

Figure 4 shows the probability plots for the distribution of the parameters shape, scale, Q100 & Q1000. These give the distribution of the uncertainty for the shape ( $\gamma = -0.14$ ) and scale ( $\sigma = 1.4$ ) parameter in a simulation with 200 values. The distribution of the shape parameter is best represented by the

normal cdf. However, those of the scale parameter and the quantiles are best represented by the lognormal cdf.

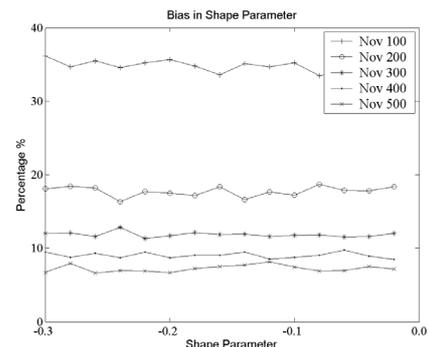


**Figure 4 Bias in Shape Parameter**

The distributions of the appropriate parameters of the CIs have proven that the assumption of the normal distribution is incorrect for the analytical confidence intervals. The aim of the research performed in Rutten 2003 was to extract appropriate CIs for ocean basins; therefore the simulation study are used.

### Bias

The results of the simulation study are biased. This bias is the difference between the value of the input parameter ( $\gamma$ ) and the mean derived from 2,500 simulation ( $\gamma^*$ ). The source of this error comes from the simulation process of generating data with the parametric bootstrap and the refitting process with the GPD cdf and MLE parameter estimator. The size of the error depends on the number of values used in the simulation and the parameter value itself.



**Figure 5 Bias in Shape Parameter**

The simulation study has shown a large bias in the shape parameter, see figure 5. For small number of values the bias can be more than 35%. The bias of the scale parameter is approximately 10 times smaller. In the GPD cdf the shape parameter is most influential, small differences in the shape parameter lead to different CIs as shown in figure 2. The bias

will therefore have a large effect on extracting confidence intervals, and needs to be corrected.

The bias of the shape and scale parameter, results in a biased computation of the quantiles, figure 6 shows the bias in Q100. The shape parameter has the largest effect in the computation of the quantiles. Although the shape parameter is severely biased and the scale parameter hardly at all the bias in the quantiles is small. A bias in the shape parameter of 30 % results in a bias in the quantiles of only 3 %, a factor of 10 less.

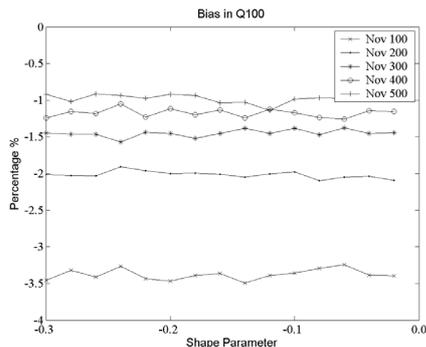


Figure 6 Bias in Q100

### Bias Correction

The shape parameter is the dominating influence on the bias. Therefore a bias correction is applied to the shape parameter in the simulation. The CIs of all shape and scale parameters and significant wave heights are extracted from a simulation study database on the basis of the shape parameter bias correction. The correction consists of two phases.

In phase one an interpolation method is applied that finds the simulated shape parameter ( $\gamma_{bc}$ ) corresponding to the input shape parameter. Due to the bias,  $\gamma^*$  is not equal to  $\gamma$ . With an iterative loop the input value for the shape is modified to till  $\gamma_{bc}$  matches  $\gamma$ . The modification of the input value that has to be used is referred to as the interpolation factor.

The second phase is based on the extraction of the confidence intervals with the interpolation factor. The CIs for the shape and scale parameter can be extracted and no further processing is required. The CIs of significant wave height have to be corrected because of the difference in bias between the shape parameter bias and the quantiles. The mean value from the quantiles is computed and the true quantile values are calculated on the basis of the EVA parameters. The CIs of the quantiles are transposed with the ratio of the calculated value over the mean of the simulated value. This produces quantile CIs with the simulated distribution and the calculated mean value.

### Confidence Intervals of locations

The interpretation of the simulation shows that there are two factors that have a dominating influence on the confidence interval of the Extreme Value predictions. These are the number of values used for the EVA and the shape parameter of the cdf. The number of values gives a side effect for the calculation of

the percentage CIs including the threshold. The percentage is smaller for a high threshold than for a low threshold. Since the height of the threshold is related to the number of values used in the EVA the effect of the number of values is reduced.

Table 2 lists the upper limit of the CIs of the three ocean basins. The trends of the CIs in the North Sea and West Africa are similar but the North Sea CIs are slightly higher. The Gulf of Mexico results show a very flat trend in the CI. The 10 and 100-year CI are similar and the 1000-year CI is only 50% higher. This shows that the physical meaning of the extreme value analysis of the Gulf of Mexico is questionable. We feel that these are too small, given the number of data at a particular point.

95 % Upper Confidence Limits			
North Sea	10 Year	100 Year	1000 Year
SNS	7 %	12 %	20 %
CNS	9 %	17 %	27 %
NNS	10 %	20 %	30 %
West Africa	10 Year	100 Year	1000 Year
Bonga	6 %	12 %	19 %
Girassol	7 %	14 %	22 %
Deep Water	7 %	13 %	23%
Gulf of Mexico	10 Year	100 Year	1000 Year
Area Selection	11 %	10 %	15 %

Table 2 CI-sizes (95 % Upper Confidence Limit / Mean value - 1)

## EFFECT OF UNCERTAINTIES ON LOAD FACTORS

### Load Factor Models

#### Reserve Strength Ratio

The load factors for fixed steel offshore structures are derived from a reliability assessment considering the structural strength and environmental loading. The type of platform that is designed determines the calculation of the partial load factors. An unmanned platform is designed for a probability of failure of  $5 \times 10^{-4} / year$ , the 2000-year return period, and a manned platform for  $3 \times 10^{-5} / year$ , the 33,000-year return period. These two different loading situations are used in this study.

The environmental load factor is derived from  $\phi_i R_i > \gamma_D D + \gamma_E E_{100}$ . The different elements of this equation consist of resistance factor  $\phi_i$ , characteristic resistance  $R_i$  for component  $i$ , dead load factor  $\gamma_D$ , dead load  $D$ , environmental load factor  $\gamma_E$  and the environmental loading  $E_{100}$ . For extreme storm check the dead load factor is 1.1. For a brace space frame collapse requires failure of a member in compression, for which  $\phi_i = 0.85$ .

$$R_i = \frac{1.1D}{0.85} + \frac{\gamma_E E_{100}}{0.85} \quad 2$$

For a brace the dead load contribution D is small and therefore ignored. The mean collapse strength of the platform,  $R_{mean}$ , is 10% greater than  $R_i$ , hence  $R_{mean} = 1.1R_i$ . Furthermore the yield strength is enhanced by another 5% due to the strain rate effect.

$$R_{mean} = \frac{1.1 \times 1.05 \times \gamma_E \times E_{100}}{0.85} = 1.37 \times \gamma_E \times E_{100} \quad 3$$

The reserve strength ratio, RSR, is defined as the mean strength (Rmean) divided by the 100-year load ( $E_{100}$ ).

$$RSR = \frac{R_{mean}}{E_{100}} \quad 4$$

$$RSR = 1.37 \times \gamma_E \quad 5$$

In equation 6 and 7 RSR targets are derived for unmanned and manned structures by considering an unmanned platform to collapse as soon as it experiences the 2000-year load and a manned structure when it experiences the 33.000-year load.

$$RSR_{unmanned} = \frac{E_{2000}}{E_{100}} \quad 6$$

$$RSR_{manned} = \frac{E_{33000}}{E_{100}} \quad 7$$

Once the RSR target has been established the environmental load factors are obtained from equation 3. The environmental loading is calculated with a stick model representation of a platform and using a combination of wave, current and wind forces. This is referred to as the Generic Load Model (GLM) developed by Shell, documented in Tromans and Vanderschuren [1995]. The current and wind loading have been left out of the GLM in this study. In equation 8 and 9 the reduced GLM formulae are given.

$$F = A_4 a^2 \Phi^2 + A_5 \frac{a^3 \Phi^2}{T_z^2} + A_6 a^2 \Phi^2 T_z^2 \quad 8$$

$$M = B_4 a^2 \Phi^2 + B_5 \frac{a^3 \Phi^2}{T_z^2} + B_6 a^2 \Phi^2 T_z^2 \quad 9$$

The environmental conditions are given by half the wave height  $a$ , the zero-upcrossing period  $T_z$  and the kinematics reduction factor for wave spreading  $\Phi$ . The coefficients  $A_{4-6}$ ,  $B_{4-6}$  have been determined by multiple wave loading simulations on a given structure for a physically realistic ranges of parameters. The indices of the coefficients  $A$  and  $B$  starts at 4 in order to keep in line with the paper by Tromans and Vanderschuren [1995] (coefficient 1-3 are current based).

The GLM formulae are used to calculate the  $E_{2000}$ ,  $E_{33000}$  and  $E_{100}$  loads. The loads can be determined by either base shear or overturning moments, the differences between the two methods are small. The results here are based on the base shear. The partial load factor for an unmanned and a manned platform can now be calculated with equation 10.

$$\gamma_E = \frac{RSR_{(unmanned/manned)}}{1.37} \quad 10$$

### Bayesian Predictive Integral

If we are involved with calculating the expected probability distribution for a random variable X, then the inferences we make on X should reflect the uncertainty in the parameters  $\theta$ . In the Bayesian terminology we are interested in the so-called predictive function:

$$F(x) = \int_0^{\infty} F(x|\theta) f(\theta) d\theta \quad 11$$

where  $F(x|\theta)$  is the probabilistic model of X, conditional on the parameters  $\theta$  and  $F(x)$  is the predictive distribution of the random variable x, now parameter free. In popular words: "the uncertainty in the  $\theta$  parameters has to be integrated out".

The predictive distribution can be interpreted as being the distribution  $F(x|\theta)$  weighted by  $f(\theta)$ . *In making inferences on a random variable it is important to use the predictive function for x*, as opposed to the probabilistic model for x with some estimator for the parameter set  $\theta$ , i.e.  $f(x|\theta^*)$ . This is because using point estimators for uncertain parameters underestimates the variance in the random variable X.

The Bayesian Predictive Integral (BPI) allows deriving the environmental loading on a platform taking account of the uncertainties in the extreme value prediction of the significant wave height in a rigorous way. The probability density function (pdf) of the extreme value distribution is known from the EVA Hindcast and the statistical variability is derived with EVA Simulation. The Bayesian Predictive Integral derives a new pdf that integrates the statistical uncertainty into the pdf of the extreme value distribution.

In  $p(H_s)$  the pdf of the extreme value distribution and  $p(H_{s_0} | H_s)$  the probability that the significant wave height is  $H_{s_0}$  instead of  $H_s$ . The integral in equation 12 calculates  $p(H_{s_0})$  that is the pdf of the extreme value distribution including the uncertainties. The integral of equation 13 calculates the probability that the wave height  $H_s$  is exceeded.

$$p(H_{s_0}) = \int_0^{\infty} p(H_{s_0} | H_s) p(H_s) dH_s \quad 12$$

$$P(H_s) = \int_0^{\infty} p(H_{s_0}) dH_{s_0} \quad 13$$

The pdf  $p(H_s)$  of the extreme value distribution is either the GEV or GPD pdf. The  $p(H_{s_0} | H_s)$  comes from the confidence interval extracted in the simulation study and is the lognormal pdf. The EVA LRFD module computes equation 12 and 13 and extracts the 100, 2.000 and 33.000-year return period  $H_s$  values from  $P(H_s)$ . These values are used in the GLM to compute the environmental load factor.

### Distributed Load Factor

The Distributed Load Factor (DLF) approach is less rigorous as the BPI, but still makes use of the distribution of the extremes. Instead of using the mean value of the significant wave height for the 100, 2000 and 33.000-year return period, the parameter histogram of the variability is used. The histogram of the Hs is transferred into a histogram of the load factor. The partial load factor for the environmental loading is now extracted by a weighted average computation. Figure 7 shows the results of the Distributed Load Factor approach for a manned platform.

The histogram of the 95% CI of Hs is extracted from the simulation study performed in Rutten [2003]. For the computation of the load factors for a manned platform, the histograms of Hs (33.000 year and 2.000 year) are entered in the GLM. The GLM computes the load factors for the 51 individual values that make up both histograms and shows the results as a histogram of the Load Factor. The environmental Load Factor is computed from the weighted average of the Load Factor histogram.

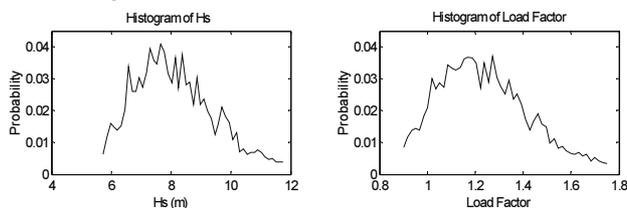


Figure 7 Distributed Load Factor

## Site-Specific Load Factors

### The effect of integrating environmental uncertainties in load factors

The partial load factors are derived with three methods; the RSR, the Bayesian Predictive Integral and the Distributed Load Factor. The results of the load factors are listed with respect to the RSR in table 3, the load factors of the RSR have been unified.

The RSR load factors have been calculated with the simplified GLM model without current and wind loading. Under the given simplifications the results are in accordance with the load factors used for actual designs in the Ocean Basins. Therefore, the traditional method can be used as a benchmark for the BPI and the DLF. The load factors, calculated on the basis of the RSR, depend on the shape parameter of the cdf and the water depth of each location. The current loading is proportional to the water depth. Therefore, the error of the simplified GLM (no current) is also proportional to the water depth.

Load Factors (M= Manned Platform, UM= Unmanned Platform)			
North Sea	Traditional	BPI	DLF
SNS	Unity	UM: 1.15 M: 1.40	UM: 1.04 M: 1.10
CNS	Unity	UM: 1.14 M: 1.35	UM: 1.03 M: 1.06
NNS	Unity	UM: 1.17 M: 1.40	UM: 1.03 M: 1.08
West Africa	Traditional	BPI	DLF
Bonga	Unity	Not Available	UM: 1.05 M: 1.06
Girassol	Unity	Not Available	UM: 1.02 M: 1.05
Deep Water	Unity	UM: 1.10 M: 1.30	UM: 1.03 M: 1.06
Gulf of Mexico	Traditional	BPI	DLF
Area	Unity	Not Available	UM: 1.05 M: 1.10

Table 3 Partial Load Factors

The new methods that implement the environmental uncertainties in the GLM model produce higher load factors. There are two relations that can be extracted from Table 3.

The increase in the load factors is larger for the manned platforms than for unmanned platforms. The simulation study has shown that the size of the CIs increases with the extrapolation, e.g. the 33,000 year CIs are larger than the 2,000 year CIs. Due to the difference in the CIs the ratio of the environmental loading with uncertainty over the environmental loading without uncertainty ( $E_{33.000}*/E_{33.000}$  and  $E_{2.000}*/E_{2.000}$ ) will be larger for the 33.000 than for the 2.000-year return period. Therefore the increase of the RSR is larger for the 33.000-year return period resulting in a higher increase of the load factor for the manned platforms.

The difference between the BPI and the DLF can be explained by the technique in which the CIs are used in both methods. The BPI uses the cdf of the CI (lognormal) while the DLF uses the histogram of the 95% CI of the 2.000 and 33.000-year return periods. The lognormal cdf is right skewed and its range is larger than the range of the histogram of 95% CI. The right skewness and larger range give ratios for  $E_{33.000}*/E_{33.000}$  and  $E_{2.000}*/E_{2.000}$  that are larger for the BPI than for the DLF. The use of the lognormal cdf gives errors for the numerical computations at certain parameter combinations (listed as "not available" in table 3).

### Recommendations on the use of environmental uncertainties

The load factors computed in Rutten [2003] cannot be used in design practice due to the simplified GLM. However, the results do show the differences of the extreme value behaviour and uncertainties in Ocean Basins and their effects on the partial load factors. Taking the uncertainties into account increases the load factors and results in stronger structures at higher costs. The cost effects start the discussion if increasing the load factors is necessary.

Current design experience, based on the number of platforms installed in the world and its failure rate, gives reason to believe that the methodology applied nowadays is correct and does not require increased load factors. This is valid for the thousands of small fixed platforms in water depths up to 50 or

60 meters in the North Sea or the South China Sea with a low value at risk. The question rises if this is also valid for the few platforms in deep water or relatively unexplored areas with a high value at risk and many uncertainties. For new designs with such a criteria, capturing the uncertainty of the environmental loading may be desired.

The BPI and DLF concepts have both shown to be systematic approaches towards the implementation of the environmental loading uncertainties. The implementation of the uncertainties results in load factors that increase with roughly 30% for BPI and 8% for the DLF. For design projects with a large environmental uncertainty and a high value at risk, the DLF gives a rational approach for the implementation of the environmental uncertainties. The BPI is more rigorous and can be used in most situations but results in higher but not necessarily conservative load factors. Both concept methods cannot be used in real design practice due to the simplifications in the GLM and the code of the API and ISO standard.

The design code by API does not consider environmental uncertainties ( $\gamma_E = 1.35$ ). The new standard by ISO (still under development) has different load factors for manned platforms (exposure Level 1) and unmanned platforms (exposure Level 2) based on their locations in the world (e.g. North Sea (L1  $\gamma_E = 1.40$ ) and Australia (L1  $\gamma_E = 1.59$ )). The new ISO standard will provide Ocean Basin averaged environmental load factors without taking account of the uncertainties.

There are other methods for the incorporation of the environmental uncertainties in the design of offshore structures. A more probabilistic approach is the first order reliability method (FORM). With this method, all loads and resistances are transformed to normally distributed variables and the design is based on the linearisation of the limit state function. The method also leads to partial load factors (as described in Baker [1982]). The use of a full probabilistic design (with Monte Carlo or Riemann integration) gives a new look at the design of offshore structures.

The partial load factors derived in Rutten [2003] have the same methodology as the ISO norms but cannot be compared with them due to the simplified GLM model. The two methods given do provide a transparent method to derive site-specific load factors. If the simplified GLM is extended with the current and wind loading, site-specific environmental load factors could be derived. In Rutten [2003] a systematic extreme value analysis in combination with a simulation study leads to the determination of uniform load factors around the world.

## CONCLUSIONS

The extreme value analysis performed with the GPD theory and MLE parameter estimator describe the characteristics of the extreme value behaviour in the North Sea and offshore West Africa well. However, the analysis appeared to be less appropriate for the Gulf of Mexico.

The simulation study provides reliable sizes of the CIs for the ocean basins; the upper CIs for the North Sea, offshore West Africa and Gulf of Mexico for the 1000-year return period are 30%, 20% and 15% respectively.

The simulation study showed the presence of a bias in the shape parameter. Appropriate CIs can be extracted from the simulation after the application of a bias correction.

The load factors calculated for the North Sea, offshore West Africa and the Gulf of Mexico on the basis of the ocean basin dependent extreme value behaviour are different, even without uncertainties.

Computing load factors including the uncertainty in the environmental loading leads to increased values. For the Distributed Load Factor method the increase is approximately 10% and for the Bayesian Predictive Integral approximately 30%.

## ACKNOWLEDGMENTS

The authors would like to express their gratitude to Jan Meek and Walt Massie from the Delft University of Technology for their support during the writing of this report. Furthermore we would like to thank George Forristall, Paul Verlaan and Phil Jonathan, for dedicating time to discussions about extreme values and uncertainties.

## REFERENCES

- Baker, M.J. and Thoft-Christensen, P. Structural Reliability theory and its applications. Berlin, Springer, 1982. ISBN 3-540-11731-8.
- Fisher, R.A. and L.H.C. Tippett 1928. Limiting forms of the frequency distribution of the largest or the smallest member of a sample. Proceeding of Cambridge Philosophical Society, 24, 180-190.
- Boendey, M.D. Van Gelder P.H.A.J.M. and J.K. Vrijling, Bootstrap simulations for evaluating the uncertainty associated with POT estimates of extreme wind velocity, January 2003, Environmetrics, Vol. 14, p.27-43.
- Pickands J. 1975, Annals of Statistics, 3, 119-131. Statistical inference using order statistics.
- Rutten, J.G., 2003, Uncertainties in extreme value analysis and their effect on partial load factors. Shell EP 2003-5377.
- Smith, R.L. 1987. Estimating tails of probability distributions. The Annals of Statistics, 1987, Vol. 15, No. 3, pp. 1174-1207.
- Tromans, P.S., Vanderschuren, L., 1995, Response based design conditions in the North Sea: Application of a new method, Offshore Tech. Conf., no. OTC 7683, pp. 387-397.