Open Access

# Bivariate Analysis of Extreme Wave and Storm Surge Events. Determining the Failure Area of Structures

Panagiota Galiatsatou* and Panagiotis Prinos

*Department of Civil Engineering, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece*

**Abstract:** In the present paper a bivariate process of extreme waves and storm surges at a Dutch station on the North Sea is considered. A bivariate logistic model and a sequential estimation procedure are used to extract joint exceedance probabilities of the two variables. The parameters of the margins of the bivariate distribution are defined by three different methods of estimation: a) the Maximum Likelihood Estimation (MLE) approach, b) a Bayesian procedure with flat prior distributions and c) the L-Moments (LM) estimation procedure. Comparison of the results of the three methods is performed and general conclusions are extracted. An approach to estimate the failure area of a particular structure under extreme sea conditions is presented, using the margins resulting from the three different estimation methods.

**Keywords:** Bivariate process, joint exceedance probabilities, Maximum Likelihood Estimation, Bayesian procedure, L-Moments, failure area.

## 1. INTRODUCTION

Offshore and coastal structures are exposed to extreme wave conditions and therefore an optimum design requires the estimation of factors affecting them and especially the estimation of extreme oceanographic data. For such systems, the notion of return period has to be extended to more than one dimension to represent multiple environmental variables. The aforementioned structures are often designed based on the estimation of their failure probabilities under extreme conditions of the wave climate. Long-term characteristics of the wave climate (such as the significant wave height or the wave period) and the storm surge can constitute the hydraulic boundary conditions related to the design of offshore and coastal structures.

Structures typically fail because of the occurrence of extreme values of a single environmental process or a critical extreme combination of constituent variables. Failure of a coastal structure is mainly caused by loadings arisen from extreme waves and water levels. The risk of local flood events is related straightforward to nearshore wave and water level conditions. However, the fact that nearshore waves can be affected by hydraulic events, such as the breaking of waves caused by the depth, enforces an extreme value analysis of offshore data. The results of the extrapolation process, deriving from the extreme value analysis of offshore waves and water levels, can then propagate shoreward using wave propagation and long-wave models.

Extreme offshore wave heights are often strongly dependent on high sea levels. The observed water level is the sum of a deterministic astronomical tidal component and a stochastic meteorologically induced component, the surge. Dependence between surges and waves is expected, since both are related to local weather conditions [1]. Especially at extreme levels strong dependence is likely, when meteorological systems which generate extreme surges also cause strong onshore winds from a direction having a long fetch. For a probabilistic design and optimization of the design process based on flood risk to be possible, loads imposed on coastal and offshore structures are described using a joint density distribution function.

The basic methodology for creating such a multivariate function starts with choosing independent multivariate observations, according to data availability in each particular case and the purpose of such an analysis. If the study variables are primary variables causing the phenomenon of coastal flooding (such as wave height and surge), different possible combinations of concomitant observations have to be studied to find the most conservative among them. To define these possible bivariate observations, thresholds for both primary variables have to be defined. After defining the extreme bivariate observations, dependence between these two variables is calculated. Based on the dependence function of the variables, appropriate bivariate models are chosen to simulate their extreme values. Following the estimation of an appropriate bivariate model, extrapolation to more extreme levels than those observed is conducted. To estimate the marginal parameters of the variables under study, the univariate estimation procedures can be utilized, such as the Maximum Likelihood Estimation, the Bayesian approach and the L-Moments procedure.

To calculate the joint probabilities of wave height and water level or of wave height and surge, different techniques of the Multivariate Extreme Value Theory are utilized. The Multivariate Extreme Value Theory is more complicated than the equivalent univariate analysis and the appropriate methodology to estimate the multivariate models and the

*Address correspondence to this author at the Department of Civil Engineering, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece; Tel: +30 2310 995 856; Fax: +30 2310 920 718;
E-mail: pgaliats@civil.auth.gr

extrapolation to levels more extreme than those observed, was developed during the last few years. Coles and Tawn [2, 3], Joe *et al.* [4], Zachary *et al*. [5], DEFRA [6], Butler [7] and Li and Song [8] present different aspects of the Multivariate Extreme Value Theory with applications to oceanographic and environmental datasets. Yeh and Ou [9] use the Multivariate Extreme Value Theory to estimate the joint exceedance probabilities of extreme significant wave height and water level, during hurricanes.

The implementation of the Multivariate Extreme Value Theory is based on the calculation of the dependence between the variables under study. Primitive methods for estimating the dependence of extreme values of different variables have been proposed by Buishand [10, 11], Dales and Reed [12] and Chilès and Delfiner [13]. Coles *et al.* [14] and Coles [15] use methods of Multivariate Extreme Value Theory to examine the existence of dependence between the bivariate extreme values of significant wave height and surge at a coastal site of southwestern England, while Ledford and Tawn [16] present a range of diagnostic tools based on tail characteristics of joint survivor functions for identifying a structure within extreme events and apply this technique to rainfall and exchange rate data. Svensson and Jones [17] implement the dependence measure $\chi$, in investigating the dependence between the sea surge and river flow, using precipitation as a proxy for river flow.

In the present work a simple methodology to estimate the failure area of adequately defined coastal structures is implemented. This approach selects the "worst case" combination of wave height and surge in terms of the response of the structure under study (e.g. overtopping or force). Another important aspect of the study is that the joint return levels that represent the worst case combination of wave height and surge are calculated using three different methods of estimation of the marginal distributions of both variables, namely the Maximum Likelihood Estimation (MLE) procedure, a Bayesian approach with flat priors and the L-Moments (LM) approach, and the results of these methods are compared. Galiatsatou and Prinos [18] study the bivariate process of extreme wave heights and storm surges, using different methods of selecting concurrent observations as well as different measures of extremal dependence of the two variables involved. Galiatsatou [19] uses the three abovementioned estimation procedures to calculate the parameters of the margins of a bivariate distribution of extreme wave heights and storm surges, considering the cases of full and asymptotic dependence. Galiatsatou and Prinos [20], Galiatsatou *et al.* [21] and Sánchez - Arcilla *et al.* [22] consider the effects of different methods of estimation of the marginal distributions of rainfall, surge and wave height data, respectively, on the return levels of the studied process.

The datasets used in the present paper are presented in Section 2, with special reference to the methodology used to "de-cluster" the available bivariate observations. In Section 3, a methodology to perform bivariate analysis of extreme wave heights and surges is developed. Techniques of selecting bivariate observations, fitting marginal distribution functions to wave heights and storm surges, estimating the dependence of the bivariate data, and selecting and fitting an appropriate model from the family of Multivariate Extreme Value (MVE) distributions to the bivariate data are presented and analysed further. Three different methods of estimation, the Maximum Likelihood Estimation (MLE) procedure, a Bayesian approach with flat priors and the L-Moments (LM) approach are used to calculate the parameters of the marginal distributions of both variables. At the end of the Section, an approach to determine the failure area of a structure and to estimate the design parameters under extreme sea conditions is presented. The characteristics of the structure used are those proposed by Coles and Tawn [3]. In Section 4, the methodology analyzed in the previous Section is implemented using wave height and storm surge data from a Dutch station on the North Sea. Joint exceedance probabilities resulting from the three different estimation methods of the marginal distributions of wave heights and storm surges (MLE, Bayesian approach, L-Moments) are compared and discussed. The three different estimation methods are also compared with regard to bivariate return levels of wave height and storm surge corresponding to the worst case in terms of overtopping of the studied structure. Finally, Section 5 summarizes the conclusions of the present work.

## 2. DATA BASE

The datasets used in the present work consist of a sequence of 23 years, over the period 1979-2001, from 9 locations along the Dutch coast (Fig. (**1**)). Two of the nine buoys lie far from the coast in the neighborhood of platforms, while of the other seven, four lie in the deep water some twenty kilometres off the coast, and the rest is in the neighborhood of the coast or the delta estuary. Three hourly data of wave height $H_{m0}$ and its standard deviation, average wave period $T_{m02}$, main wave direction $T_{h0}$, average wave height and period of the highest third part $H_{1/3}$ and $T_{H1/3}$ and the wave height of the low frequency waves $H_{TE3}$ are available. In addition to the wave parameters, data on wind speed and direction, water level relative to NAP/MSL, set-up or surge (the difference between the observed and the astronomical water level) and a column indicating the origin of the measured variables are included in the files. At most locations, two wave survey instruments, namely a main sensor and a secondary sensor, were used. If both sensors have registered values for a parameter, then the mean value is given. If, however, data are only present from one sensor, then this is used in the file and if neither of the sensors have values, then estimated values are retained. Missing records are patched by hindcasting. Details of these data sets are given in www.golfklimaat.nl of the Dutch National institute RIKZ. In Fig. (**1**) the coordinates of the nine stations are presented in the Dutch RD coordinate system (Rijksdriehoeksmeting), while Table **1** contains the names and the geographical coordinates of the stations, the duration of the measurements, as well as the water depth at the nine locations.

The theory described in the following assumes independence of successive observations of wave heights and storm surges. In practice, there is considerable short term dependence in metocean data. Extreme events are typically to be found in storms, which may last for many hours or even several days. In the present paper, an attempt to "de-cluster" extreme wave height and storm surge events similar to the most commonly used approach of explicitly identifying clusters of storm events (a cluster of storm events is a group of

### *3.1.2. Fitting Marginal Distributions for Wave Height and Surge*

If $X_1$, $X_2$,…, $X_n$ is a series of independent random observations of a random variable $X$ with common distribution function $F(x)$ and $Y_1$,$Y_2$,…, $Y_k$ ($Y_i$=$X_i$-$u$) are the excesses over a high enough threshold $u$, in some asymptotic sense, the conditional distribution of excesses follows the Generalized Pareto Distribution (GPD):

$$G(y) = 1 - (1 + \frac{\xi(y-u)}{\sigma})^{-1/\xi} \quad (1)$$

where $\sigma$ is the modified scale and $\xi$ is the shape parameter of the GPD distribution. An appropriate threshold $u$ is selected for both variables (wave heights and surges), which defines the level upon which an extreme event is defined. Two different methodologies are used for the selection of $u$ : (a) the mean residual life plot of the excesses of different threshold values and (b) the plots of parameters $\sigma$ and $\xi$ for a variety of possible threshold values. The mean residual life plot consists of the points: $\{(u, \frac{1}{n_u}\sum_{i=1}^{n_u}(x_{(i)} - u)$: $u < x_{max}\}$ where $x_{(1)},\ldots, x_{(n_u)}$ consist of the $n_u$ observations that exceed $u$ and $x_{max}$ is the largest of the $X_i$ [15]. An appropriate threshold value is the value of $u$ above which the mean residual life plot is approximately linear and estimates of $\sigma$ and $\xi$ are constant with $u$. Due to sampling variability, estimates of these parameters will not be exactly constant, but they should be stable after allowance for their sampling errors.

The choice of the threshold values, $u$, for both the variables of wave height and surge is also based, apart from the methodologies used in the univariate analysis and were previously mentioned, on methodologies used in the bivariate framework. The variables of wave height ($X_1$) and surge ($X_2$) are transformed to Fréchet margins, so that each have distribution function $F(z)$ = exp(-1/$z$) for $z$ >0, to highlight dependence of the two variables at extreme levels and then the radial and angular components, $R$ and $W$, respectively, are defined as:

$$R=(X_1+X_2)/n \text{ and } W_1=X_1/(X_1+X_2) \quad (2)$$

where $n$ is the number of bivariate pairs of observations. An extreme bivariate event is defined to be any value for which the radial component $R$ is sufficiently large. To extract such events out of the bivariate sample of wave height and surge data, histograms of the angular component $W_1$ are constructed for all bivariate pairs exceeding various choices of the radial component $R$ ($r_o$) . Then, $r_{min}$ is taken as the smallest value of $r_o$ above which there is apparent stability in the shape of the histograms [4]. After selecting $r_{min}$, the marginal thresholds of the variables involved are estimated using equation (3) :

$$u_j = \Psi_j^{-1}(nr_{min}), \text{ j=1, 2} \quad (3)$$

where $\Psi_j(X_{i,j}) = -\{\log F_j(X_{i,j})\}^{-1}$ for $X_{i,j} \le u_j$. Marginal thresholds, estimated utilizing the methodology based on the bivariate framework (equation (3)), are usually close to the unvariate thresholds or higher than them.

Thus, the assumption made about each marginal component (wave heights and storm surges) is that for high enough threshold $u_j$, the marginal distribution of $X_j-u_j$, for $X_j > u_j$ is the Generalised Pareto Distribution (GPD) $F_j(x)=1-\lambda_j\{1+\xi_j(x-u_j)/\sigma_j\}_+^{-1/\xi j}$, $x \ge u_j$, where $\lambda_j=1-F_j(u_j)$. The scale and shape parameters, $\sigma_j$ and $\xi_j$, respectively, can be calculated using three different methods of estimation: a) the Maximum Likelihood Estimation procedure, b) the Bayesian approach and c) the L-Moments estimation procedure.

#### ***3.1.2.1. The Maximum Likelihood Estimation Procedure***

The Maximum Likelihood Estimation (MLE) approach is a commonly used estimation procedure. Among others, Coles [15] used this approach to estimate the parameters of extreme value distributions fitted to sea level, rainfall as well as to financial data. The likelihood function gives the relative likelihood of the obtained observations, as a function of the parameters $\boldsymbol{\theta}=(\sigma, \xi)$: $L(\boldsymbol{\theta},x)=\Pi f(x_i,\boldsymbol{\theta})$ and where $L$ (or, for numerical convenience log$L$) is maximized with respect to the parameters $\sigma$ and $\xi$.

Maximization of $L(\boldsymbol{\theta},x)$ with respect to the set of parameters $\boldsymbol{\theta}$, is numerically straightforward and also has the convenience that various standard large sample theory results are available to enable the numerical calculation of standard errors and confidence intervals. If the available sample sizes are large, there seems little doubt that the Maximum Likelihood estimator is a good choice [26].

#### ***3.1.2.2. The Bayesian Approach***

In the Bayesian setting, parameters ($\sigma$, $\xi$) are treated as random variables and prior distributions on them are intended to represent beliefs about their values, prior to the availability of the data. The specification of information in the form of a prior distribution is regarded alternately as the greatest strength and the main pitfall of Bayesian inference [15]. The absence of genuine prior information leads to the use of priors $\pi(\boldsymbol{\theta})$ that have very high variance, or equivalently, are nearly flat. A trivariate normal distribution is used here, that enables the specification of independent parameters. Setting $\varphi$=log$\sigma$, a prior density function $f(\sigma, \xi)=f_\varphi(\varphi)\cdot f_\xi(\xi)$ is chosen, where $f_\varphi(\cdot)$, $f_\xi(\cdot)$ are normal density functions with mean zero and very high variances, corresponding to a specification of prior independence in the parameters $\varphi$ and $\xi$. The posterior density of $\boldsymbol{\theta}$ is given by [27]:

$$\pi(\theta|x) = \frac{\pi(\theta)L(\theta;x)}{\int_\theta \pi(\theta)L(\theta;x)\mathrm{d}\theta} \propto \pi(\theta)L(\theta;x) \quad (4)$$

where $L(\theta;x)$ is the likelihood function. Standard Markov chain Monte Carlo methods routinely allow the approximation of integrals such as the one in the denominator of equation (4). More details about the Bayesian methodology used here are given in Galiatsatou and Prinos [20], Galiatsatou *et al.* [21] and Sánchez-Arcilla *et al.* [22].

#### ***3.1.2.3. The L-Moments Approach***

L-Moments are analogous to ordinary moments. They provide measures of location, dispersion, skewness, kurtosis and other aspects of the shape of probability distributions or data samples, but are computed from linear combinations of the ordered data values [26]. They were introduced by Hosking [28], who assessed that L-moments weigh each element of a sample according to its relative importance.

The main advantage of L-Moments is that, being a linear combination of data, they are less influenced by outliers and the bias of their small sample estimates remains fairly small. It is therefore anticipated that L-Moments can provide reliable estimates of tail-index with a relatively small sample of the POT data. Furthermore, the required computation is quite limited compared with other traditional techniques, such as the Maximum Likelihood Estimation approach [29].

Using three sample L-Moments, the scale ($\sigma$) and shape ($\xi$) parameters of the GPD can be estimated as [29]:

$$\xi = \frac{(3\tau_3 - 1)}{(\tau_3 + 1)} \qquad (5.1)$$

$$\sigma = (1-\xi)(2-\xi)\lambda_2 \qquad (5.2)$$

where $\tau_3$ is the normalized L-Moment $\tau_3=\lambda_3/\lambda_2$. The L-Moments approach is a method requiring quite limited computation compared with other traditional techniques, such as the Maximum Likelihood Estimation approach.

### *3.1.3. Estimation of Dependence of the Bivariate Data*

Wave heights and storm surges are not independent variables, but they are certainly characterized by some form and some degree of dependence. The complete pair of measures of extremal dependence $\chi$ and $\bar{\chi}$, introduced by Coles *et al.* [14], is informative for both asymptotically independent and dependent variables. After transformation of the pair of variables ($X_1$, $X_2$) to ($U$, $V$)= $\{F_{X_1}(X_1), F_{X_2}(X_2)\}$ having Uniform marginal distributions, the one-dimensional function $\chi(u)$ is defined for a given threshold as:

$$\chi(u) = 2 - \frac{\log \Pr(U<u, V<u)}{\log \Pr(U<u)} \text{ for } 0 \le u \le 1 \qquad (6)$$

The measure $\chi$ can then be defined as:

$$\chi = \lim_{u\to 1} \chi(u) \qquad (7)$$

The one dimensional function $\bar{\chi}(u)$ is defined for $0\le u\le 1$:

$$\bar{\chi}(u) = \frac{2\log \Pr(U>u)}{\log \Pr(U>u, V>u)} - 1 \qquad (8)$$

It follows that:

$$\bar{\chi} = \lim_{u\to 1} \bar{\chi}(u) \qquad (9)$$

When used for bivariate random samples with identical marginal distributions, both measures provide an estimate of the probability of one variable (e.g. wave heights) being extreme, provided that the other one (e.g. surge levels) is extreme. The sign of $\chi(u)$ determines whether the variables are positively or negatively associated at the quantile level $u$. In the special case $\chi$=0 the variables are asymptotically independent. For asymptotically dependent variables $\bar{\chi}$=1. The complete pair of ($\chi$, $\bar{\chi}$) can give an impression of extremal dependence. The pair ($\chi$>0, $\bar{\chi}$=1) indicates asymptotic dependence, while the value of $\chi$ determines the strength of dependence and the pair ($\chi$=0, $\bar{\chi}$<1) signifies asymptotic independence, in which case the value of $\bar{\chi}$ determines the strength of dependence within the class [15]. Galiatsatou and Prinos [18] present plots of the pair ($\chi$, $\bar{\chi}$) for wave heights and surges at station Eld (Eierlandse Gat), 20km off the Dutch coast, identifying an increased degree of correlation of the two processes as $u \to 1$ suggesting a tendency for the most extreme levels to be correlated.

Another measure used in the present paper to examine dependence of the two processes at extreme levels is the coefficient of tail dependence, $\eta$. For estimating the coefficient of tail dependence, $\eta$, the structure variable $T$=min($X$, $Y$), where $X$ and $Y$ are the wave height and surge data transformed to have the standard Fréchet distribution, is defined. The coefficient of tail dependence, $\eta$, is estimated by using a point process approach, as the shape parameter of the univariate $T$ variable. The coefficient of tail dependence, $\eta$, can be related to the dependence measure $\bar{\chi}$ using the formula: $\bar{\chi} = 2\cdot\eta - 1$. If $\eta$=1 and $L(z)\to$c as z$\to\infty$ ($L(\cdot)$ is a slowly varying function, i.e. $\lim_{s\to\infty}\{L(st)/L(s)\} = 1$ for all fixed $t$>0), with 0<c≤1, then ($\chi$=c, $\bar{\chi}$=1) and the variables are asymptotically dependent of degree c [14].

### *3.1.4. Choosing an appropriate model from the family of MVE*

To determine an appropriate bivariate extreme value model (BVE), components $X_1$ (wave height) and $X_2$ (surge), having unit Fréchet margins, are transformed to radial and angular components, $R=(X_1+X_2)/n$ and $W_1= X_1/(X_1+X_2)$, where $n$ is the number of bivariate pairs. The $W_1$ versus log$R$ plot, for points exceeding at least one of the marginal thresholds defined can be used to determine whether there is mass near the boundaries of the one dimensional unit simplex $S_2$ ({$S_2$= ($W_1$,$W_2$): $\sum_{j=1}^{2} W_j = 1$, $W_j$>0, $j$=1,2}). If most of the points lie in the interval 0.1<$W_1$<0.9 and few points are near the boundaries of $W_1$, parametric models with all mass in the interior of the one dimensional unit simplex $S_2$ are quite appropriate, namely the Dirichlet and symmetric versions of the two logistic models (logistic and negative logistic) [2].

A more detailed inspection of the structure of the dependence function of the bivariate data requires the estimation of the parameter of asymmetry $\psi_\alpha$ for all candidate bivariate models. If $\psi_\alpha$ is the asymmetry parameter of a bivariate extreme value model, it is estimated using the formula [24]:

$$\psi_\alpha = \int_0^{0.5} 4(A(W_1) - A(1-W_1))/(3-2\sqrt{2})\,\mathrm{d}W_1 \qquad (10)$$

This measure lies in the interval [−1,1], with large absolute values representing stronger asymmetry. $A$ is the dependence function of the data. For the logistic and negative logistic models $A(W_1)=A(1-W_1)$ for all $0\le W_1\le 0.5$, so the value of $\psi_\alpha$ will be zero.

**Fig. (7).** Contours of design overtopping discharge $Q(\upsilon_p, z)$=0.002 $m^3$/sm for wave periods $T$ = 4, 6, 8, 10, 12, 14 s.

In Fig. (**8**), the intersection points of the contour lines are estimated for wave periods of $T$=8.7s for the Maximum Likelihood Estimation procedure, $T$=7.9s for the Bayesian estimation and $T$=8.4s for the L-Moments approach. The critical combinations of wave height ($x$) and surge ($y$), as estimated from Fig. (**8**), are for the Maximum Likelihood Estimation procedure $x_{4000\ \text{years}}$ =5.76 m and $y_{4000\ \text{years}}$ =2.88 m ($T$=8.7s), for the Bayesian estimation approach $x_{4000\ \text{years}}$ =5.90 m and $y_{4000\ \text{years}}$ =3.19 m ($T$=7.9s) and for the L-Moments approach $x_{4000\ \text{years}}$ =5.92 m and $y_{4000\ \text{years}}$ =2.97 m ($T$=8.4s). The Bayesian and L-Moments wave height estimates are almost identical, while regarding the surge, the Bayesian estimates are the highest. With regard to the Maximum Likelihood Estimation procedure, the values of wave height when marginal parameters from the Bayesian estimation approach are used, are estimated higher up to 2.5%, while the respective values of surge up to 10.7%. The wave period, $T$, in the latter case is lower up 9.2%. It can be noticed (Fig. (**8**)) that the equal overtopping curve lies towards the top left of the diagram for all three different methods of estimation of the marginal parameters of the two variables.

## DISCUSSION AND CONCLUSIONS

In the present work a bivariate process of extreme waves and storm surges at the Dutch station Schouwenbank (Swb) was considered. After the selection of an appropriate bivariate sample to be analyzed using extreme value methodologies, the analysis proceeded with the estimation of joint exceedance probabilities using the simple bivariate logistic model and a sequential estimation procedure, where the parameters of the margins of the bivariate distribution were defined by: a) Maximumum Likelihood Estimation (MLE), b) Bayesian estimation with flat prior distributions and c) the L-Moments estimation procedure. Using the margins resulting from the three different methods of estimation, an approach to estimate the failure region of a particular structure under extreme sea conditions was presented. The main conclusions of the paper can be summarized as follow:

a. The variables of wave height and storm surge are both related to local meteorological conditions. Thus, it is expected that dependence at their extreme levels can be high enough for the variables to be considered consistent with asymptotic dependence. This assumption leads to the use of a bivariate distribution from the family of MVE to model their dependence structure. In the present work, the symmetric bivariate logistic model was proven to be adequate.

**Fig. (8).** Estimation of the worst case combination of significant wave height and surge with return period of 4000 years for three different methods of estimation of their marginal distributions (station Swb).

b. The method of estimation of the marginal characteristics of wave height and storm surge has a large impact on joint exceedance probabilities of the two variables. The Maximum Likelihood Estimation (MLE) approach underestimates the probability of exceedance of extreme joint events of waves and surges compared to those of the Bayesian with flat prior distributions and of the L-Moments procedures. When the Bayesian estimation procedure is used for the marginal distribution of the storm surge, bivariate return levels are more conservative compared to the other two methods of estimation, while for wave height the Bayesian estimation and the L-Moments approach give really close results. It should be noted that in each case the method of estimation of the marginal parameters of the variables involved, which leads to the most conservative results in terms of joint exceedance probabilities, depends critically upon the characteristics of the extreme sample under study.

c. When overtopping of a coastal structure is considered as the main cause of flooding in a certain area, the worst case combination of wave height and storm surge, for a given design value of the height of the structure $\upsilon_p$, can be estimated as the intersection point of the contour line of joint exceedance probabilities of the variables with the contour line which corresponds to equal overtopping discharge rate per unit length of the structure, $Q(\upsilon_p, z)$.

d. The equal overtopping curve of the structure under study lies towards the area of the diagram of the two variables, where surge level is higher for all three different methods of estimation of the marginal parameters of the wave height and the storm surge.

e. Comparing the worst case combinations for the seawall under study, the Bayesian and L-Moments wave height estimates are almost identical, while the Bayesian estimates give the highest surge. Wave height and surge, when marginal parameters from the Bayesian estimation approach are used, are estimated up to 2.5% and 10.7% higher than those of the Maximum Likelihood Estimation procedure.

## ACKNOWLEDGEMENTS

The authors want to thank the Dutch National institute RIKZ for providing data of the Dutch coast. The work described in this publication was supported by the European Community's Sixth Framework Programme through the grant to the budget of the Integrated Project FLOOD*site*, Contract GOCE-CT-2004-505420. The paper reflects the authors' views and not those of the European Community. Neither the European Community nor any member of the FLOOD*site* Consortium is liable for any use of the information in this paper.

## NOTATION - GLOSSARY

| | | |
|---|---|---|
| MVE | = | Multivariate Extreme Value distributions |
| MLE | = | Maximum Likelihood Estimation procedure |
| LM | = | L-Moments approach |
| GPD | = | Generalized Pareto Distribution |
| $u$ | = | Appropriate threshold for variable $X$ |
| $n_u$ | = | Observations exceeding $u$ |
| $x_{max}$ | = | Largest value among all $X_i$ |
| $X_1, X_2$ | = | Representations of wave height and surge variables |
| $R$ | = | $(X_1+X_2)/n$ Radial component |
| $W_1$ | = | $X_1/(X_1+X_2)$ angular component |
| $F$ | = | Distribution function of data |
| $\boldsymbol{\theta}$ | = | $(\mu,\sigma,\xi)$= Vector of parameters (location, scale, shape) of distribution fitted to data |
| $L(\boldsymbol{\theta},x)$ | = | Likelihood function of data with respect to model parameters |
| $f(x_i,\boldsymbol{\theta})$ | = | Density functions of data |
| $f_\mu(\cdot), f_\varphi(\cdot), f_\xi(\cdot)$ | = | Normal density functions with mean zero and variances $\upsilon_\mu$, $\upsilon_\varphi$, $\upsilon_\xi$, respectively |
| $\pi(\theta\vert x)$ | = | Posterior density of parameters vector $\boldsymbol{\theta}$ |
| $\pi(\boldsymbol{\theta})$ | = | Prior density for vector of parameters $\boldsymbol{\theta}$ |
| $L(\boldsymbol{\theta};x)$ | = | A likelihood for $\boldsymbol{\theta}$ based on an observed set of exceedances $x$ |
| POT | = | Peaks Over Threshold |
| $\tau_3$ | = | The normalized L-Moment $\tau_3=\lambda_3/\lambda_2$ |
| $U, V$ | = | Transformed pair of variables having Uniform marginal distributions |
| $\chi$, $\bar{\chi}$ | = | Measures of extremal dependence introduced by Coles *et al.* [14] |
| $T$ | = | Structure variable defined to estimate the coefficient of tail dependence, $\eta$ |
| $\eta$ | = | Coefficient of tail dependence |
| $s$ | = | Transformation of variables to unit Fréchet margins |
| BVE | = | Bivariate Extreme Value distributions |
| $S_2$ | = | Unit simplex $\{(W_1,W_2): \sum_{j=1}^{2} W_j = 1$, $W_j>0$, $j=1,2\}$ |
| $\psi_\alpha$ | = | Asymmetry parameter |
| $A$ | = | Dependence function of the data |
| NLLH | = | Negative log-likelihood |
| $G(x_1, x_2)$ | = | Bivariate distribution (the symmetric logistic model) of $X_1$ and $X_2$ |
| $z_1, z_2$ | = | transformed GPD margins of the variables $X_1$ and $X_2$ |
| $r$ | = | Dependence parameter of bivariate logistic model |
| $A_{\mathbf{v}}$ | = | $\{\mathbf{x} \in \mathbf{R}^d: b(\mathbf{x};\mathbf{v}) > 0\}$ failure region of a coastal structure |
| $b$ | = | boundary function $\mathbf{R}^d$ x $V \rightarrow \mathbf{R}$, where $V$ is the design parameter space of $\mathbf{v}$ |
| $\upsilon$ | = | The height of the seawall |

$X_1^*, X_2^*, X_3$ = The inshore significant wave height, wave period and surge, respectively

$m$ = The mean sea level at the area under study

$z$ = The mean tidal level

$Q(\upsilon, z)$ = Expected overtopping discharge rate per unit length of the seawall

$\alpha_1, \alpha_2$ = Dimensionless constants depending on the form and characteristics of the seawall design

$Q(\upsilon_p, z)$ = Design overtopping discharge per unit length of the seawall

$x, y$ = Realisations of variables $X, Y$

## REFERENCES


[1] P. J. Hawkes, B. P. Gouldby, J. A. Tawn, and M. W. Owen, "The joint probability of waves and water levels in coastal engineering design", *J. Hydraulic Res.*, vol. 40(3), pp. 241-251, 2002.

[2] S. G. Coles, and J. A. Tawn, "Modelling extreme multivariate events", *J. R. Stat. Soc. Series B Stat. Methodol.*, vol. 53(2), pp. 377-392, 1991.

[3] S. G. Coles, and J. A. Tawn, "Statistical methods for multivariate extremes: An application to structural design", *Appl. Stat.*, vol. 43(1), pp. 1-48, 1994.

[4] H. Joe, R. L. Smith, and I. Weissman, "Bivariate threshold methods for extremes", *J. R. Stat. Soc. Series B.*, vol. 54, pp. 171-183, 1992.

[5] S. Zachary, G. Feld, G. Ward, and J. Wolfram, "Multivariate extrapolation in the offshore environment", *Appl. Ocean Res.*, vol. 20, pp. 273-295, 1998.

[6] DEFRA/ Environment Agency Flood and Coastal Defence R&D Programme, *Best Practice in Coastal Flooding*, R&D Technical Report FD2206/TR1, Workshop Version issued 24/09/03, HR Wallingford Report TR 132, 2003.

[7] A. Butler, "*Statistical Modelling of Synthetic Oceanographic Extremes*", Ph.D. thesis, Lancaster University, London, UK, 2005.

[8] C. W. Li, and Y. Song "Joint probability analysis of extreme waves and water levels at a typhoon invaded coastal region", In: *Proceedings of XXXI IAHR Congress*, 2005, pp. 3924-3934.

[9] S. P. Yeh, and S. P. Ou, "Joint probability analysis of waves and water level during typhoons", In: *Proceedings of the Third Chinese-German Joint Symposium on Coastal and Ocean Engineering*, 2006, [Online] Available: http://www.comc.ncku.edu.tw/joint/joint2006/pdf/68%20CHM%2006.pdf [Accessed 28th October 2009]

[10] T. A. Buishand, "Bivariate extreme-value data and the station-year method", *J. Hydrol.*, vol. 69, pp. 77-95, 1984.

[11] T. A. Buishand, "Extreme rainfall estimation by combining data from several sites", *Hydrol. Sci. J.*, vol. 36, pp. 345-365, 1991.

[12] M. Y. Dales, and D. W. Reed, *Regional Flood and Storm Hazard Assessment*, Institute of Hydrology no 102, Institute of Hydrology, Wallingford, Oxfordshire, OX10 8bb, UK, 1989.

[13] J. P. Chilès, and P. Delfiner, *Geostatistics: Modelling Spatial Uncertainty*, Wiley, New York, 1999.

[14] S. Coles, J. Heffernan, and J. Tawn, "Dependence measures of extreme value analysis", *Extremes*, vol. 2(4), pp. 339-365, 1999.

[15] S. Coles, *An Introduction To Statistical Modelling of Extreme Values*, Springer Series in Statistics, Springer, Berlin, 2001.

[16] A. Ledford, and J. A. Tawn, "Diagnostics for dependence within time series extremes", *J. R. Stat. Soc. Series B*, vol. 65(2), pp. 521-543, 2003.

[17] C. Svensson, and D. A. Jones, "Dependence between sea surge, river flow and precipitation in South and West Britain", *Hydrol. Earth Syst. Sci.*, vol. 8(5), pp. 973-992, 2004.

[18] P. Galiatsatou, and P. Prinos, "Analysis of dependence in a bivariate process of extreme waves and surges", In: *Proceedings of the 1st International Conference on Coastal Zone Management and Engineering in the Middle East*, Dubai 2005, pp. 221-225

[19] P. Galiatsatou, "Joint exceedance probabilities of extreme waves and storm surges", In: *Proceedings of XXXIII Congress of IAHR*, 2007, pp. 780 (abstract) - JFK Competition, (The full paper in the electronic proceedings of the conference)

[20] P. Galiatsatou, and Prinos, P., "Analysis of extreme rainfall events using a Poisson process", In: *Proceedings of the 10th National Conference E.Y.E*, Xanthi 2006, Greece, pp. 47-54.

[21] P. Galiatsatou, P. Prinos, and A. Sánchez-Arcilla, "Estimation of extremes. Conventional versus Bayesian techniques", *J. Hydraulic Res.*, vol. 46(2), pp. 211-223, 2008.

[22] A. Sánchez-Arcilla, J. Gomez-Aguar, J. J. Egozcue, M. I. Ortego, P. Galiatsatou, and P. Prinos, "Extremes from scarce data. The role of Bayesian and scaling techniques in reducing uncertainty", *J. Hydraulic Res.*, vol. 46(2), pp. 224-234, 2008.

[23] R. L. Smith, "Extreme value analysis of environmental time series: an example based on ozone data (with discussion)", *Stat. Sci.*, vol. 4, pp. 367-393, 1989.

[24] P. Galiatsatou, and P. Prinos, "Statistical models for bivariate extremal analysis of a spatial process", *J. Hydraulic Res.*, vol. 46(2), pp. 257-270, 2008.

[25] A. Repko, P. H. A. J. M. van Gelder, H. G. Voortman, and J. K. Vrijling, "Bivariate description of offshore wave conditions with physics- based extreme value statistics", *Appl. Ocean Res.*, vol. 26, pp. 162-170, 2004.

[26] P. H. A. J. M. van Gelder, "*Risk-based Design of Civil Structures*", Ph.D. thesis, University of Technology, Delft, The Netherlands, 1999.

[27] A. Stephenson, and J. Tawn, "Bayesian inference for extremes: Accounting for the three extremal types", *Extremes*, vol. 7(4), pp. 291-307, 2004.

[28] J. R. M. Hosking, "L-Moments: Analysis and estimation of distributions using linear combinations of order statistics", *J. R. Stat. Soc. Series B*, vol. 52, pp. 105-124, 1990.

[29] M. D. Pandey, P. H. A. J. M. van Gelder, and J. K. Vrijling, "The estimation of extreme quantiles of wind velocity using L-moments in the peaks-over-threshold approach", *J. Structural Safety*, vol. 23(2), pp. 179-192, 2001.

[30] J. A. Tawn, "Bivariate extreme value theory: Models and estimation", *Biometrika*, vol. 75, pp. 397-415, 1988.

[31] DEFRA/ Environment Agency Flood and Coastal Defence R&D Programme, *Joint Probability: Dependence Mapping and Best Practice*, Technical report on dependence mapping, R&D Technical Report FD2308/TR1, 2005.

[32] M. J. Tucker, A. P. Carr, and E. G. Pitt, "The effect of an offshore bank in attenuating waves", *Coastal Eng.*, vol. 7, pp. 113-144, 1983.


Received: November 20, 2010 Revised: November 20, 2010 Accepted: November 25, 2010