www.trjfas.org
ISSN 1303-2712
Turkish Journal of Fisheries and Aquatic Sciences 12: 523-531 (2012) DOI: 10.4194/1303-2712-v12_2_44

Analysis of Extreme Marine Events Causing Flooding in Varna Region

Panagiota Galiatsatou[1,*], Panayotis Prinos[1], Nikolay Valchev[2], Ekaterina Trifonova[2]

[1] Aristotle University of Thessaloniki, Hydraulics Laboratory, Department of Civil Engineering, 54124 Thessaloniki, Greece.
[2] Institute of Oceanology, BAS, Varna, Bulgaria.

* Corresponding Author: Tel.: +30.2310 995856; Fax: +30.2310 995672;
E-mail: pgaliats@civil.auth.gr

Received 15 March 2012
Accepted 28 July 2012

Abstract

In the present paper extreme wave height, wave period and storm surge events in the marine area of the Varna region are analyzed using Extreme Value Theory (EVT) to estimate return levels of the aforementioned variables corresponding to selected return periods. Both univariate and multivariate techniques are utilized. The GEV (Generalized Extreme Value) distribution is implemented to extrapolate the marine variables to levels more extreme than those observed. Joint probability analysis is also conducted. A bivariate extreme value model is used to produce estimates of joint probabilities of extreme wave heights and storm surges. For wave period extremes, a conditional distribution on extreme wave height is utilized. Joint estimates of wave height, storm surge and wave period are compared to the respective estimates of the univariate analysis, for selected return periods.

Keywords: EVT, GEV, univariate extremes, bivariate analysis, conditional distribution.

Introduction

Flood risk is defined as the combination of the probability of a flood event and of the potential adverse consequences to human health, the environment and economic activity associated with the event. For a risk to arise there must be a hazard that consists of a source or initiating event, a receptor (person or property) and a pathway that links the receptor to the source. For coastal areas, the hazards that contribute to the risk of flooding are related to the water surface elevation process. High wave conditions and high sea levels offshore and transformed to nearshore are typically considered as the “source” of coastal flooding. Description of risk sources will help to deliver the hydraulic boundary conditions which are needed to describe the loading of flood defence structures or will already be the key input for the probability of flooding.

Flood risk sources in coastal areas are associated to the occurrence of extreme marine events (wave climate and storm surge events). Some recent floods can be classified as extreme events with a recurrence interval of once in a hundred years or less. Extreme value methods are powerful statistical methods for drawing inference about the extremes of a process, using only data on relatively extreme values of the process. Extreme value methods are usually utilized for the purpose of extrapolation to levels more extreme than those which have been observed. The statistical methodology is motivated by a well-established mathematical theory (Extreme Value Theory - EVT), which relies on the assumption that the limiting models suggested by the asymptotic theory continue to hold at finite but extreme levels. Nevertheless, a crucial assumption in fitting distribution functions to data is that the data are independent and identically distributed (iid).

Univariate extreme value theory includes models for block maxima, as well as models for exceedances over appropriately defined thresholds known as peaks over threshold (POT) models. Use of the block maximum model for statistical applications seems to have started in the 1950’s. Gumbel promotes the methodology of using the Generalised Extreme Value (GEV) distribution to model componentwise maxima. Tawn (1992) applies the methodology to oceanographic data, while Walshaw and Anderson (2000) use it for wind field modeling. The statistical attributes of the approach of the problem of analysing extreme values using thresholds are studied in detail by Davison and Smith (1990), Walshaw (1994). The POT method is rather common today and it is considered, under conditions of course, advantageous in comparison to other techniques of analysis. Coles and Tawn (1996), Coles and Casson (1999) and Goda

(2000) contribute considerably to the use of the model in various applications. Coles (2001) introduces shortly the mathematical foundation of the model. Galiatsatou and Prinos (2006) and Galiatsatou *et al.* (2008) use the point process model with rainfall and storm surge extremes within a stationary context.

Where the source consists of more than one variables (e.g. coastal flooding caused by extreme wave heights and water levels), it is necessary to consider their joint probability. Joint exceedance probability refers to the chance of two or more partially related variables occurring simultaneously. Joint exceedance combinations of wave heights and sea levels with a given chance of occurrence are defined in terms of sea conditions in which a given wave height is exceeded at the same time as a given water level (or its surge component) being exceeded. Multivariate Extreme Value Theory (MVE) is used to describe the joint distribution of two or more variables and appropriate methodology has only been developed in recent years. Coles and Tawn (1991, 1994), Zachary *et al.* (1998) and Schlather and Tawn (2003) give the main aspects of MVE with applications to oceanographic and other sort of datasets. Athanassoulis *et al.* (1994), Morton and Bowers (1996), De Haan and De Ronde (1998), Ferreira and Soares (2002) and Repko *et al*. (2004) describe the joint probability distribution function of long-term hydraulic conditions. Galiatsatou (2007) compares different pairs of bivariate observations of extreme waves and surges with reference to joint exceedance probabilities, in order to find the most severe sea state caused by the two variables, while Galiatsatou and Prinos (2008) implement four bivariate models to the data of significant wave height and wave period from a field station off the Dutch coast and compare them mainly on their ability to correctly describe the data and on their behavior when extrapolation is of interest.

In the present paper analysis of extreme events, namely wave height, wave period and storm surge in the marine area of the Varna region is conducted using Extreme Value Theory (EVT) to estimate return levels of the aforementioned variables corresponding to selected return periods. Both univariate and multivariate techniques are presented. In Section 2, the datasets available are presented and described. In Section 3, univariate extreme analysis is conducted. More specifically, the GEV (Generalized Extreme Value) distribution is implemented to extrapolate the marine variables of wave height, storm surge and wave period to levels more extreme than those observed. The fitting of the distribution to all available datasets is judged and discussed. In Section 4, joint probability analysis is conducted for the studied variables. A bivariate extreme value model is used to produce estimates of joint probabilities of extreme wave heights and storm surges. Considering wave period extremes, a conditional distribution on extreme wave height is utilized. Joint estimates of wave height, storm surge and wave period are compared to the respective estimates of the univariate analysis, for selected return periods. In Section 5, the main conclusions of the analysis conducted are presented.

Datasets

The present work focuses on extreme value analysis of the marine climate in the coastal area of Varna. Varna coastal region is located in the Western Black Sea sector and it is mainly characterized by an eastern exposure. The average width of the accumulative coast is 63 m and the maximum width reaches up to 105 m (Keremedchiev *et al*., 2008). Winds from the NE prevail in the northern and middle sections of the shelf zone, while the impact of eastern winds increases southward. Usually, southeast winds are less significant in terms of storm intensity but are still of importance for the northern shelf in particular (Valchev *et al*., 2008). Following the wind patterns waves propagate most frequently from E, NE and SE.

The datasets considered and analysed in the present work are annual and monthly maxima of wave heights, storm surges and wave periods. Wave heights and wave periods are available for a time period of 61 years (1948-2008), while storm surge data cover a period of 80 years (1928-2007). Regarding storm surge, it should be noted that the available data are obtained from sea level measurements of the sea-level gauge in Varna. Main Administration for Geodesy and Cartography (now - Agency for Cadaster) maintains four sea level gauges along the Bulgarian Black sea coast. Sea-level gauge in Varna was put in operation in 1928. For this measuring location, there is no lack of data during the period of the Second World War.

Due to absence of regular long-term measurements of wind and wave conditions in the western Black Sea, the analysis of wave climate is based on continuous hindcast data series. The global sea level pressure reanalysis carried out by the NCEP/NCAR (Kalnay *et al*., 1996) is used in order to calculate the historical wind forcing for wave models. This is done by means of an atmospheric-ocean interaction model (Lavrenov, 1998). Since global fields have rather coarse resolution (2.5° spatial and 6 hours temporal), they are downscaled to the Black Sea domain. As a result, the obtained hourly gridded wind fields have a resolution of 0.5°.

Wave conditions are modeled using a coupled system of third generation spectral wave models. The WAM model (Günther *et al*., 1992) is run on a regular spherical grid covering the entire Black Sea basin at 0.5° spatial resolution. The deep-water settings are applied with source and propagation time steps set to 10 min and 20 min, respectively. The SWAN Cycle III model, version 40.72 (The SWAN Team, 2006) is set up for wave simulations in the western shelf zone. It is nested to WAM as its output provides conditions in terms of significant wave height, mean wave period, and mean direction of wave propagation on a number of points along the external boundary. Therefore wave height and wave period data from a SWAN model grid point in front of the study site at a depth of 20 m are used in the present work. More specifically, the variables used in the present paper to represent wave climate are those of the mean wave height (H_m) and the mean wave period (T_m). It should be noted that on the basis of measurements there is an established linkage between H_m, H_s (significant wave height), H_{rms} (root-mean square wave height) as well as between T_m, T_p (peak wave period) and T_z (mean zero crossing period), therefore the selected set of wave characteristics brings all information needed for future work that is not in the focus of the present study. To be consistent with the principles of a physically-based homogeneity of the data, wave maxima and associated maximum wave periods for three different directions affecting the coast are considered: a) the southeastern, b) the northeastern and c) the eastern.

Analysis of Univariate Extremes

The calculation of extreme quantiles is often applied to statistical models which use block maximum data. The GEV (Generalized Extreme Value) distribution is considered as an appropriate model for such data. It should be noted that the choice of the block size is critical, signifying that the block maxima should not violate the assumption of their identical distribution, where the model is founded. Pragmatic considerations often lead to the adoption of blocks of length one year. In the present work the GEV distribution function and the block maximum model are applied to available annual and monthly maxima of the marine variables. As the data sample consists of more than 60 years of observations, this method is considered to be reliable in terms of extrapolation results for return periods up to 100 years (used in the present work).

It is supposed that X_1, X_2,…X_n is a series of independent and identically distributed variables with a common distribution function F and that M_n=max(X_1, X_2,…, X_n). If there exist sequences of normalising constants a_n>0 and b_n that:

$$\mathrm{P}\left(\frac{M_n\text{-}b_n}{a_n} \le z\right) \rightarrow G(z) \quad \text{as } n\rightarrow\infty \qquad (1)$$

for all $z \in [z_-, z_+]$, where G is a non degenerate function supported in the interval $[z_-, z_+]$, then G is a member of the Generalised Extreme Value (GEV) distribution family, with distribution function:

$$G(z) = \exp[-\{1+\xi(\frac{z\text{-}\mu}{\sigma})\}_+^{-1/\xi}] \qquad (2)$$

where μ, σ >0 and ξ are location, scale and shape parameters of the distribution, respectively. The special case ξ=0 reduces the GEV to the Gumbel distribution function.

To estimate the parameters of the GEV distribution function, a common estimation procedure is applied, the Maximum Likelihood Estimation (MLE). The likelihood function is given with respond to acquired observations, and parameters $\theta = (\mu, \sigma, \xi)$:

$$\mathrm{L}(\boldsymbol{\theta}, x) = \prod \mathrm{f}(x_i, \boldsymbol{\theta}) \qquad (3)$$

and L (or, for numerical advantage logL) is maximized with regard to parameters μ, σ and ξ. Method MLE gives unbiased estimates of parameters and from all unbiased estimators it has the smallest mean square error (van Gelder, 2000). The maximization of L($\boldsymbol{\theta}$, x), with regard to all parameters $\boldsymbol{\theta}$, is numerically direct and allows easily the numerical calculation of standard errors and confidence intervals (Coles *et al*., 2003).

The GEV distribution function (Eq. (2)) assumes a homogeneous distribution for the extreme population data within a year. However, for monthly maxima the hypothesis of homogeneity is not adequately satisfied, since the effects of seasonality are evident. Therefore, when fitting the stationary GEV model to monthly wave height and storm surge extremes, the fitting is poor particularly at the most extreme tails and estimates of the model parameters lead to non-physically based extreme distributions (e.g. the Fréchet distribution which is lower-bounded). Therefore, a log-transformation is applied to the monthly maxima. The GEV model is fitted to the log-transformed sample and return level estimates are then back-transformed to the original data scale. This transformation seems to reduce the seasonal effects in the monthly samples, eliminating the main differences between seasons and enhancing the data homogeneity within each year. The fitting of the GEV distribution is significantly improved, when applying the log-transformation, while it is insured that the Weibull model (upper-bounded distribution) is a suitable distribution for the monthly maxima, as it was the case for annual maxima. For the datasets available, it is judged that the log-transformation is necessary for the monthly wave height and storm surge data, while it is not required for wave period monthly maxima, because seasonal effects in wave period extremes are not pronounced. However, it should be noted that in cases where seasonal data are available, a time-dependent GEV model should be applied. It is avoided in this paper, because it would perplex the fitting of the bivariate distribution function in Section 4.

Tables 1(a) and 1(b) present return level estimates of annual and monthly wave height extremes of southeastern, northeastern and eastern direction, resulting from the univariate analysis. The monthly data are first log-transformed to fit the GEV

Table 1(a). Return level estimates from fitting the GEV distribution to wave height annual maxima

Return period (years)	Return level estimates of annual maximum H_m (m)								
	Southeastern direction			Northeastern direction			Eastern direction		
	2.5%	MLE	97.5%	2.5%	MLE	97.5%	2.5%	MLE	97.5%
5	2.39	2.61	2.84	2.36	2.52	2.68	3.27	3.38	3.48
20	2.85	3.23	3.61	2.70	2.95	3.19	3.51	3.61	3.71
50	3.04	3.60	4.17	2.82	3.18	3.53	3.60	3.70	3.81
100	3.12	3.86	4.60	2.88	3.33	3.79	3.64	3.75	3.87

Table 1(b). Return level estimates from fitting the GEV distribution to wave height monthly maxima

Return period (years)	Return level estimates of monthly maximum H_m (m) (log-transformation is included)								
	Southeastern direction			Northeastern direction			Eastern direction		
	2.5%	MLE	97.5%	2.5%	MLE	97.5%	2.5%	MLE	97.5%
5	2.44	2.66	2.91	2.42	2.54	2.66	3.23	3.48	3.75
20	2.99	3.39	3.84	2.76	2.93	3.10	3.60	4.01	4.47
50	3.28	3.81	4.42	2.91	3.11	3.33	3.75	4.26	4.84
100	3.46	4.09	4.84	3.00	3.22	3.46	3.83	4.41	5.07

distribution function and then back-transformed to their original scale. The central column for each direction corresponds to the MLE, while the other two columns represent the 95% confidence intervals, assuming that the maximum likelihood estimator is normally distributed.

Figure 1, representatively shows the fitting of the GEV distribution to the annual maxima of southeastern direction, by means of four diagnostic plots (p-p plot, q-q plot, return level plot and density plot). From Figure 1, as well as from the diagnostic plots of all three directions, it can be concluded that the GEV distribution provides a very good fit to the wave height annual extremes. Empirical observations (marked with circles in Figure 1) lay very close to the diagonal, representing the perfect fit, in both the probability (p-p plot) and the quantile (q-q plot) plots. All empirical observations lay also within the 95% confidence intervals in the return level plot, while in the density plot the fitted model density represents the histogram quite well. For wave height monthly maxima (after the log-transformation), the fitting of the stationary GEV distribution is not as good as for the annual maxima, as a result of remaining seasonal effects, especially for the eastern direction. However, it is judged satisfactory for the southeastern and the northeastern directions.

For smaller return periods the eastern direction gives higher wave height return level estimates compared to the other two directions, while for high return periods (e.g. 100 years) the southeastern direction wave height estimates are the highest (Table 1(a)). Indeed the distribution function for eastern wave height annual extremes is strongly upper-bounded. The range of the confidence interval for eastern wave height extremes is significantly narrower, compared to the other two directions. More specifically, the 95% confidence interval range for the eastern direction wave height corresponding to a return period of 5 years is up to 53% and 34% narrower than the one for southeastern and northeastern waves, respectively. For a return period of 100 years, these percentages increase to about 84% and 75%, respectively. For monthly maxima (Table 1(b)), the eastern direction gives the highest return level estimates. The northeastern wave direction presents the lowest return level estimates, together with the narrowest 95% confidence interval ranges for small as well as for large return periods. For a return period of 100 years the confidence interval range of the northeastern wave height is narrower compared to the southeastern and the eastern up to 67% and 63%, respectively. Comparing return level ML estimates from annual and monthly data, it can be concluded that the latest are larger for the southeastern and the eastern directions, while the opposite happens for the northeastern direction. For the southeastern, northeastern and eastern directions, differences in ML estimates reach 6%, 3.5% and 18%, respectively. The range of the 95% confidence intervals for the northeastern direction is reduced to half from fitting the GEV to monthly maxima. It should be noted that for the eastern direction any comparison is not trustworthy, due to the not so good fitting of the GEV model to the monthly maxima.

Tables 2(a) and 2(b) present return level estimates of annual and monthly wave period extremes for waves of southeastern, northeastern and eastern direction, resulting from the univariate analysis. The four diagnostic plots (not presented here for the sake of brevity) for annual wave period extremes show a very good fit of the GEV distribution function. It should be noted that the fitting is better for annual wave period extremes of southeastern and northeastern direction. For wave period monthly maxima, the fitting of the stationary GEV distribution is judged satisfactory.

Return level estimates for wave periods of

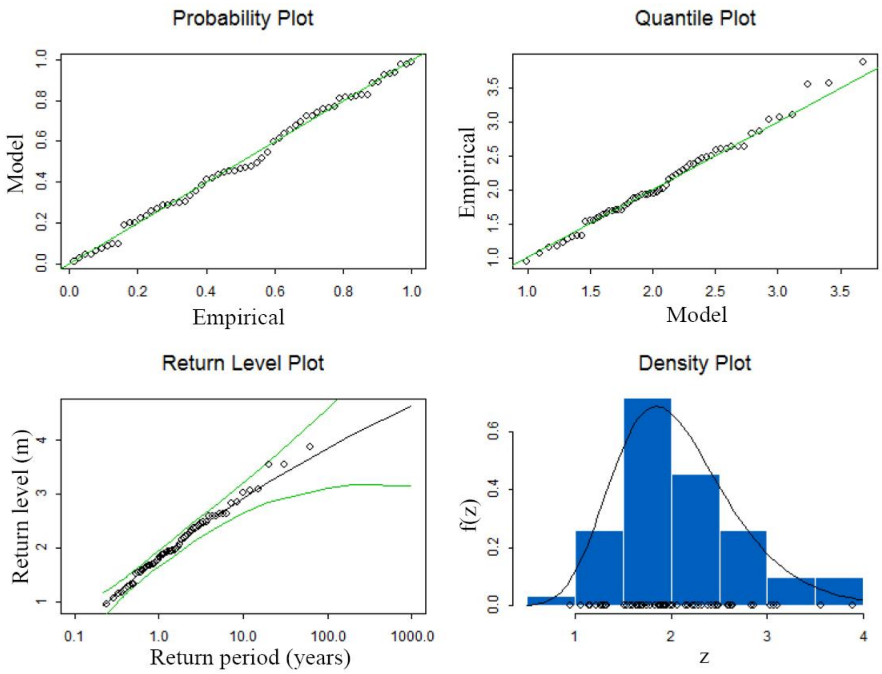

Figure 1. Diagnostic plots for GEV fit to the southeastern wave height annual maxima.

eastern direction are higher than the respective estimates corresponding to southeastern and northeastern direction (Table 2(a)). More specifically, MLEs for a return period of 100 years are higher for the eastern direction up to 10.5% and 30%, compared to the estimates of the southeastern and the northeastern direction, respectively. The range of the 95% confidence interval for the northeastern direction is narrower from the respective ranges of the southeastern and the eastern direction up to 47% and 33%, respectively (for the studied return periods). For monthly maxima (Table 2(b)), the eastern direction gives also the highest ML return level estimates. For a return period of 100 years, the confidence interval range of the southeastern wave period is narrower compared to the northeastern and the eastern up to 51% and 57%, respectively. Comparing ML return level estimates from annual and monthly data, it can be concluded that the latest are larger for the northeastern and the eastern directions, while the opposite happens for the southeastern direction. For the southeastern, northeastern and eastern directions, differences in MLEs reach 5%, 8% and 5%, respectively. The range of the 95% confidence intervals for the southeastern direction is reduced from fitting the GEV to monthly maxima more than three times.

Table 3 presents return level estimates of annual and monthly storm surge extremes, resulting from the univariate analysis. The sample that is used for analysis comprises only of the 61 years (1948-2007), for which observations of the wave climate are available. It should be noted that when the GEV distribution is fitted to the whole sample of storm surge data (80 years), the resulting model seems to describe the observations very poorly. The resulting distribution is a Fréchet one, with no upper-bound, which is also physically incorrect. The stationary GEV model seems not to be appropriate for the whole data set, maybe because of some trend inherent or some unusual records in the earliest data.

From the four diagnostic plots (omitted here for the sake of brevity), the GEV distribution seems to provide a very good fit for both the annual and the monthly (log-transformed) storm surge data. Return level estimates, resulting from annual and monthly (log-transformed) data seem to be close enough (Table 3). MLEs of storm surge return level resulting from annual maxima are larger than the resulting estimates from monthly maxima up to about 10% (for the studied return periods). Confidence interval ranges are reduced when monthly maxima are used for the fitting, up to about 40% for a return period of 100 years, hence uncertainty in return level estimates is significantly reduced.

Analysis of Multivariate Extremes

Where the source of risk consists of more than one variable, it is necessary to consider their combined probability. Dependence between surges and waves is expected, since both are related to local weather conditions (Hawkes *et al*., 2002). 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. Hence, joint probability analysis is necessary to assess the combined effect of

Table 2(a). Return level estimates from fitting the GEV distribution to wave period annual maxima

Return period (years)	Return level estimates of annual maximum T (sec)								
	Southeastern direction			Northeastern direction			Eastern direction		
	2.5%	MLE	97.5%	2.5%	MLE	97.5%	2.5%	MLE	97.5%
5	5.94	6.14	6.34	5.24	5.40	5.57	6.68	6.86	7.05
20	6.33	6.70	7.07	5.59	5.83	6.06	7.09	7.41	7.74
50	6.45	7.03	7.60	5.72	6.04	6.37	7.29	7.76	8.24
100	6.49	7.26	8.03	5.77	6.18	6.59	7.41	8.02	8.64

Table 2(b). Return level estimates from fitting the GEV distribution to wave period monthly maxima

Return period (years)	Return level estimates of monthly maximum T (sec)								
	Southeastern direction			Northeastern direction			Eastern direction		
	2.5%	MLE	97.5%	2.5%	MLE	97.5%	2.5%	MLE	97.5%
5	6.05	6.18	6.31	5.26	5.46	5.67	6.82	7.05	7.28
20	6.40	6.58	6.76	5.73	6.06	6.38	7.37	7.74	8.11
50	6.58	6.79	7.00	5.99	6.40	6.82	7.67	8.14	8.62
100	6.68	6.92	7.16	6.16	6.65	7.14	7.87	8.43	8.99

Table 3. Return level estimates from fitting the GEV distribution to storm surge annual and monthly maxima

Return period (years)	Return level estimates of maximum SS (m)					
	Annual maxima			Monthly maxima ((log-transformation is included)		
	2.5%	MLE	97.5%	2.5%	MLE	97.5%
5	0.46	0.51	0.55	0.45	0.49	0.53
20	0.57	0.65	0.73	0.55	0.61	0.68
50	0.60	0.73	0.85	0.59	0.67	0.76
100	0.62	0.78	0.95	0.62	0.71	0.82

these variables associated with selected return periods. Considering wave period extremes, a conditional distribution on extreme wave height is used in the present work.

The basic methodology for creating a bivariate distribution function starts with selecting independent bivariate observations, according to data availability in each particular case and the purpose of the analysis. The selection of block maxima (annual or monthly) to perform such an analysis is a way of ensuing independence in the univariate data and thus satisfying the crucial assumption of EVT. But it should also be noted that the selection of the block maximum model, instead for example to the POT, was forced in the present study due to the conducted bivariate analysis. The non-availability of sea level data during significant storms, for which wave climate data were available, was an obstacle to using the more detailed POT model.

After defining the extreme bivariate observations, dependence between the variables of wave height and storm surge is calculated. Based on the dependence function of the variables, appropriate bivariate models should be selected to simulate their extreme values. 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 χ and $\overline{\chi}$, introduced by Coles *et al*. (1999), is informative for both asymptotically independent and dependent variables. 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 complete pair of (χ, $\overline{\chi}$) can give an impression of extremal dependence (Coles, 2001). Estimating the pair of dependence measures (χ, $\overline{\chi}$) for all different pairs of wave height and storm surge data, it can be observed that there is not significant evidence that the two variables are not consistent with asymptotic dependence at extreme levels.

Modelling approaches for multivariate extremes are analogous to block maximum, threshold and point process results, derived for univariate extremes. Provided that the bivariate pairs of directional wave heights and storm surges are consistent with asymptotic dependence, their dependence function can be well represented by the dependence function of a Bivariate Extreme Value distribution (BVE). The most widely used parametric model, which is utilized to simulate the joint distribution of wave height and storm surge extreme events in the present work, is the BVE exchangeable logistic model with distribution function (Tawn, 1988):

$$G(x_1, x_2) = \exp[-(z_1^{1/r} + z_2^{1/r})^r] \quad (4)$$

with dependence parameter r ($0 < r \leq 1$). In Eq. (4) Z_1 and Z_2 are the transformed GEV margins of the variables X_1 (waves) and X_2 (surges). Complete dependence is achieved in the limit as r approaches zero, while independence is obtained when r=1.

Before fitting the bivariate logistic distribution function to pairs of wave height of one of the three selected directions (east, northeast, southeast) and of storm surge, the univariate GEV distribution is fitted to annual and monthly maxima of all variables (Section 3). MLE is applied to estimate the parameters of each such model ($\hat{\mu}, \hat{\sigma}, \hat{\xi}$). Assuming the margins of the logistic model are GEV distributed, the variables X_1 and X_2 are transformed:

$$z_1 = \{1 + \hat{\xi}_1 (\frac{x_1 - \hat{\mu}_1}{\hat{\sigma}_1})\}_+^{-1/\hat{\xi}_1} \quad (5a)$$

$$z_2 = \{1 + \hat{\xi}_2 (\frac{x_2 - \hat{\mu}_2}{\hat{\sigma}_2})\}_+^{-1/\hat{\xi}_2} \quad (5b)$$

Replacing Eq. (5) within Eq. (4) can yield selected joint exceedance probabilities of wave height and storm surge. In the present study, the bivariate logistic distribution function with GEV margins is utilized to calculate the contour levels of directional wave height and storm surge, corresponding to return periods of 5, 20, 50 and 100 years.

In view of the dependence structure between wave heights and wave periods, marginal analysis is in itself insufficient to come to an accurate description of the long-term wave climate. Hence in order to model wave period in a more reliable way, a conditional distribution of the variable on the wave height is utilized. Indeed the GEV distribution (Eq. (2)) is used for the wave period with the three parameters μ, σ and ξ. Annual and monthly maxima of southeastern, northeasten and eastern direction are analysed conditional on maxima wave heights (annual or monthly, respectively) of similar direction. It has been noted that log-transformed wave height and wave period data are highly correlated and therefore the GEV distribution is fitted to the log-transformed wave period data of the three directions. The location and scale parameters of the distribution are modelled as linear functions of the log-transformed wave height:

$$\mu = \mu_0 + \mu_1 \ln H \quad (6a)$$

$$\sigma = \sigma_0 + \sigma_1 \ln H \quad (6b)$$

where μ_0, μ_1, σ_0 and σ_1 are estimated using MLE.

Table 4 presents return level estimates of annual maximum wave heights and conditional wave periods of the three selected directions (SE, NE and E) and "concomitant" estimates of annual maximum storm surge, corresponding to return periods of 5, 20, 50 and 100 years. The combinations presented are selected out of an infinite number of different combinations of the two variables (wave height and storm surge), in the curved part of the contour plots, in a way to represent most properly the dependence structure of the data.

Comparing ML return level estimates of the univariate analysis of annual maxima of wave height and storm surge (Table 1(a), Table 3) with the estimates from the bivariate model (Table 4), it is obvious that in the latter all return levels are calculated significantly lower. More specifically for the southeastern wave direction, wave height is reduced up to almost 7% and storm surge up to 43%. For the northeastern direction, the respective proportions rise to 6% and about 51%, respectively. For the eastern direction, wave height is slightly reduced up to 3.5%, while the storm surge is estimated lower up to 45%. For wave period, differences between the estimates of the univariate stationary model (Table 2(a)) and the conditional one (Table 4) are not significant. Larger divergences are observed for the southeastern direction, where for a return period of 100 years the return level estimate resulting from the conditional model is higher than the one of the former model up to almost 4%.

As analysed in Section 3, due to the presence of strong seasonal effects in the monthly data both wave height and storm surge variables are first log-transformed to perform the fitting of the bivariate distribution function and at the end the contours are back-transformed to the original scale of the data. This transformation reduces the effects of seasonality on extremal analysis. Table 5 presents return level estimates of monthly maximum wave height and conditional wave period of the three selected

Table 4. Concomitant return level estimates of annual maximum wave heights, conditional wave periods and annual maximum storm surge

Return period (years)	Concomitant estimates of annual maximum H_m, T and Storm surge								
	Southeastern direction			Northeastern direction			Eastern direction		
	H_m (m)	SS (m)	T(s)	H_m (m)	SS (m)	T(s)	H_m (m)	SS (m)	T(s)
5	2.45	0.29	6.05	2.37	0.26	5.31	3.26	0.28	6.86
20	3.06	0.39	6.80	2.80	0.32	5.80	3.51	0.40	7.47
50	3.36	0.50	7.21	3.02	0.37	6.01	3.59	0.53	7.84
100	3.62	0.58	7.53	3.13	0.43	6.12	3.64	0.62	8.13

Table 5. Concomitant return level estimates of monthly maximum wave heights, conditional wave periods and monthly maximum storm surge

Return period (years)	Concomitant estimates of monthly maximum H_m, T and Storm surge								
	Southeastern direction			Northeastern direction			Eastern direction		
	H_m (m)	SS (m)	T(s)	H_m (m)	SS (m)	T(s)	H_m (m)	SS (m)	T(s)
5	2.34	0.25	6.55	2.29	0.26	5.32	3.16	0.25	7.08
20	2.86	0.41	6.94	2.72	0.42	5.78	3.57	0.42	7.44
50	3.35	0.50	7.12	2.95	0.47	6.01	3.92	0.48	7.60
100	3.65	0.58	7.21	3.07	0.52	6.13	4.11	0.54	7.68

directions (SE, NE and E) and concomitant estimates of monthly maximum storm surge for return periods of 5, 20, 50 and 100 years. The combinations presented are selected out of an infinite number of different combinations of the two variables (wave height and storm surge), in the curved part of the contour plots, in a way to represent most properly the dependence structure of the data.

Comparing ML return level estimates of the univariate analysis of monthly maxima of wave height and storm surge (Table 1(b), Table 3) with the estimates from the bivariate model (Table 5), it is again obvious that in the latter all return levels are reduced. More specifically for the southeastern wave direction, wave height is reduced up to almost 16% and storm surge up to 49%. For the northeastern direction, the respective proportions rise to 10% and about 47%, respectively. For the eastern direction, wave height is reduced up to almost 11%, while the storm surge is estimated lower up to 49%. For wave period, differences between the estimates of the univariate stationary model (Table 2(b)) and the conditional one (Table 5) are not very large. For the southeastern wave direction, wave periods are reduced up to 6% with the conditional model, for small return periods. For northeastern and eastern wave directions, wave periods are reduced up to 8% and 9%, respectively for large return periods.

Conclusions

In the present paper extreme wave height, wave period and storm surge events in the marine area of the Varna region were analyzed using univariate and multivariate Extreme Value Theory (EVT) to estimate return levels of the aforementioned variables corresponding to selected return periods. The GEV (Generalized Extreme Value) distribution was implemented to extrapolate the marine variables to levels more extreme than those observed. Joint probability analysis was also conducted for the studied variables. A bivariate extreme value model was used to produce estimates of joint probabilities of extreme wave heights and storm surges. Considering wave period extremes, a conditional distribution on extreme wave height was utilized. The main conclusions of the present work are summarized below:

- The GEV distribution function seems to provide a very good fit to directional wave height, storm surge and directional wave period annual maxima. Diagnostic plots (p-p plots, q-q plots, return level plots and density plots) prove that the GEV is an appropriate model for such extremes.
- The log-transformation of the wave height and storm surge monthly data ensures that the Weibull model is a suitable distribution for the monthly maxima. The log-transformation of monthly maxima seems to reduce the effects of seasonality on extremal analysis, enhancing the accuracy of extreme value predictions. For wave period, the log-transformation does not have significant effects on return level estimates, because seasonal effects are not so prominent in this signal’s extremes. However, it should be noted that in cases where seasonal data are available, a time-dependent GEV model should be applied.
- When the log-transformation is applied to monthly maxima of wave height and storm surge and the fitting of the GEV distribution to the transformed data is satisfactory, the range of the return level confidence interval is significantly reduced, compared to the case of annual maxima.
- The bivariate logistic distribution function is selected as an appropriate model for joint extreme directional wave height and storm surge events, because these extremes are characterized by asymptotic dependence. For wave period, a conditional GEV distribution of the variable on the wave height is utilized.
- Concomitant estimates of wave height and storm surge extremes are significantly reduced compared to the univariate estimates corresponding to similar return periods. For annual maxima, wave period return level estimates do not show significant divergencies between the univariate stationary model and the conditional distribution function, while when monthly maxima are utilized some differences are observed mainly at the northeastern and the eastern direction.

Acknowledgments

The support of the European Commission through FP7.2009-1, Contract 244104 - THESEUS “Innovative technologies for safer European coasts in a changing climate”, is gratefully acknowledged.

References

Athanassoulis, G.A., Skarsoulis, E.K. and Belibassakis, K.A. 1994. Bivariate distributions with given marginals with an application to wave climate description, Applied Ocean Research, 16: 1-17.

Coles, S. 2001. An introduction to statistical modeling of extreme values, Springer Series in Statistics, 1st Edition., London, 210 pp.

Coles, S., Heffernan, J. and Tawn, J. 1999. Dependence measures of extreme value analysis, Extremes, 2(4): 339-365.

Coles, S., Pericchi, L.R. and Sisson, S. 2003. A fully probabilistic approach to extreme rainfall modeling, Journal of Hydrology, 273: 35-50

Coles, S.G. and Casson, E.A. 1999. Spatial regression models for extremes, Extremes, 1: 449-468.

Coles, S.G. and Tawn, J.A. 1991. Modelling extreme multivariate events, Journal of the Royal Statistical Society B, 53: 377-392.

Coles, S.G. and Tawn, J.A. 1994. Statistical methods for multivariate extremes: an application to structural design, Applied Statistics, 43: 1-48.

Coles, S.G. and Tawn, J.A. 1996. Modeling extremes of the areal rainfall process, Journal of the Royal Statistical Society B, 58: 329-347.

Davison, A.C. and Smith, R.L. 1990. Models for exceedances over high thresholds, Journal of the Royal Statistical Society B, 52: 393-442.

De Haan, L. and De Ronde, J. 1998. Sea and wind: Multivariate extremes at work, Extremes, 1: 7-45.

Ferreira, J.A. and Soares, C.G. 2002. Modelling bivariate distributions of significant wave height and mean period, Applied Ocean Research, 24: 31-45.

Galiatsatou, P. 2007. Joint exceedance probabilities of extreme waves and storm surges, XXXIII Congress of IAHR, JFK Competition, Venice, Italy, 780 pp (abstract).

Galiatsatou, P. and Prinos, P. 2006. Analysis of extreme rainfall events using a Poisson process, Proc. of the 10th Hellenic Conference by Hellenic Hydrotechnical Association, Kalithea/Chalkidiki, Greece: 47-54

Galiatsatou, P. and Prinos, P. 2008. Bivariate analysis and joint exceedance probabilities of extreme wave heights and periods, Coastal Engineering 2008, Proc. of the 31st International Conference, Hamburg, Germany: 4121-4133.

Galiatsatou, P., Prinos, P. and Sánchez-Arcilla, A. 2008. Estimation of extremes. Conventional versus Bayesian techniques, Journal of Hydraulic Research, 46(2): 211-223.

Goda, Y. 2000. Random seas and Design of Maritime Structures, In: Liu, P.L.-F. (Ed.), Advanced Series on Ocean Engineering, World Scientific, Singapore, 443 pp.

Günther, H., Hasselmann, S. and Janssen, P.A.E.M. 1992. Wave model Cycle 4, Technical Report No.4, Deutsches Klima Rechen Zentrum, Germany.

Hawkes, P.J., Gouldby, B.P., Tawn, J.A. and Owen, M.W. 2002. The joint probability of waves and water levels in coastal engineering design, Journal of Hydraulic Research, 40(3): 241-251.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. and Joseph, D. 1996. The NCEP/NCAR reanalysis project, Buletin of the American Meteorological Society, 77: 437-471.

Keremedchiev, S., Trifonova, E., Eftimova, P. and Nikolova, A. 2008. Assessment of the geodynamic vulnerabilityof risk coastal areas along Varna coast, Proc. of the 5th International Conference, Global changes Vulnerability, Mitigation and Adaptation. 17-18 April, St. Kilent Ohridski" University Press., Sofia, Bulgaria: 138–144.

Lavrenov, I. 1998. Numerical modeling of wind waves in spatially heterogeneous ocean, Gidrometeoizdat, St. Petersburg, 449 pp. (in Russian).

Morton, I.D. and Bowers, J. 1996. Extreme value analysis in a multivariate offshore environment, Applied Ocean Research, 8: 303-317.

Repko, A., van Gelder, P.H.A.J.M., Voortman, H.G. and Vrijling, J.K. 2004. Bivariate description of offshore wave conditions with physics-based extreme value statistics, Applied Ocean Research, 26: 162-170.

Schlather, M. and Tawn, J.A. 2003. A dependence measure for multivariate and spatial extreme values: Properties and inference, Biometrika, 90(1): 139-156.

Tawn, J.A. 1988. Bivariate extreme value theory: Models and estimation, Biometrika, 75: 397-415.

Tawn, J.A. 1992. Estimating probabilities of extreme sea levels, Applied Statistics, 41: 77-93.

The SWAN Team, 2006. SWAN User Manual, Delft University of Technology, the Netherlands

Valchev, N., Davidan, I., Belberov, Z., Palazov, A. and Valcheva, N. 2008. Estimation of wind wave climate of the Western Black Sea during the last 50 years, Proc. of 9th Int. Conf., Marine Sciences and Technologies - Black Sea'2008, Varna, Bulgaria: 231-239.

van Gelder, P.H.A.J.M. 2000. Statistical Methods for the Risk-based design of civil structures, PhD thesis. Delft: University of Technology, The Netherlands.

Walshaw, D. 1994. Getting the most from your extreme wind data: a step by step guide, Journal of Research of the National Institute of Standards and Technology, 99: 399-411.

Walshaw, D. and Anderson, C.W. 2000. A model for extreme wind gusts, Applied Statistics, 49: 499-508.

Zachary, S., Feld, G., Ward, G. and Wolfram, J. 1998. Multivariate extrapolation in the offshore environment, Applied Ocean Research, 20: 273-295.