# Can the generalized Pareto Distribution be useful towards developing ship stability criteria?

Panayiotis A. Anastopoulos, *Department of Naval Architecture and Marine Engineering, National Technical University of Athens, Greece,* panasto@central.ntua.gr

Kostas J. Spyrou, *Department of Naval Architecture and Marine Engineering, National Technical University of Athens, Greece,* k.spyrou@central.ntua.gr

**ABSTRACT**

The paper investigates the effectiveness of the generalized Pareto Distribution (GPD) for modelling the tail of the distribution of ship rolling motions and particularly, for calculating the probability of capsize in beam seas. To this end, large-scale Monte Carlo numerical experiments were performed for an ocean surveillance ship assumed to operate in two qualitatively different, in terms of the observed frequency of stability failures, sea states; one where capsizes are realized quite often and another where they are extremely rare. For both sea conditions, GPD models were fitted to datasets containing roll exceedances above a pre-defined threshold and their reliability is tested herein against the rough Monte Carlo estimates, obtained by direct counting. The possibility of approximating the tail through several GPDs is discussed and the idea of associating threshold selection with the shape of the GZ curve is proposed for enhancing the accuracy of the approach. To evaluate the rumored "extrapolation" character of the GPD beyond the largest observation used in the fitting procedure, a comparison with the predictions of the "critical wave groups" method is presented for the second (mild) sea state.

**Keywords:** *Probability, Capsize, Generalized Pareto Distribution, Statistical extrapolation, Extreme events, Critical wave groups.*

## 1. INTRODUCTION

Several techniques can be employed for obtaining the distribution of the responses of a dynamical system subjected to random excitation (e.g. Chai et al., 2017). However, their application in the problem of ship capsize is hindered by their large computational requirements and/or deficiencies in dealing with the complexity of ship dynamics at large angles. Brute-force Monte Carlo simulations, despite being very attractive due to their accuracy, can easily turn into a computationally intensive exercise when a large number of extremely rare events, like capsizing, must be produced.

One possibility to alleviate the problem could be the tools provided by Extreme Value Theory (EVT), a branch of statistics focused on making inferences about the extreme values in a random process. Specifically, the second extreme value theorem (Balkema and de Haan, 1974; Pickands, 1975) states that, under certain conditions, the generalized Pareto distribution (GPD) is a limiting distribution for excesses over thresholds. This has motivated the development of a number of threshold-based methods seeking a solution to the problem of rarity of extreme ship responses through fitting the GPD to data obtained from pertinent time-domain simulations (e.g. Belenky et al., 2016; Campbell et al., 2016). Nonetheless, it is the strong data-driven character of such methods that may eventually deteriorate their effectiveness and therefore, their application for direct ship stability assessment remains an open question.

As is well known, the main issue, arising rather naturally in practical implementations of the theorem, is the selection of an appropriate threshold for fitting the GPD. Despite the model being mathematically exact at infinitely high levels, it is believed that it could still be reliable if determined with respect to a sufficiently high threshold. This, runs the danger, on the one hand, of idly expending computational resources if an exceptionally high threshold is set, resulting in datasets with only few (if not any at all) extremes. On the other hand, a lower threshold may not be able to produce reliably the tail. In practical ship stability, normally we do not need very large roll angles for judging safety since, beyond some moderate to high angle, the

flooding of closed spaces is inevitable. Hence, a question is raised whether the GPD could be meaningfully applied towards developing a stability criterion. Much of effort has been put over the last years in efficiently fitting the GPD using reasonably-sized datasets generated by fast, yet qualitatively realistic, hydrodynamic codes (e.g. Weems et al., 2016).

In our current work, the possibility of analyzing the tail structure through successive GPD fits is discussed for the problem of ship rolling in beam seas. At the same time, an attempt is made to associate threshold selection with the shape of the GZ curve of a vessel. The idea is to identify regimes where response exhibits different probabilistic qualities and then, utilize the limits of these regimes for thresholding. The performance of the approach for calculating the probability of capsize in severe sea conditions is tested against the rough Monte Carlo estimates, obtained by direct counting. Finally, to evaluate the reliability of the GPD for "statistical extrapolation" (i.e. for predicting events beyond the largest observation used in the fitting procedure), a comparison with the results of the "critical wave groups" method (Anastopoulos and Spyrou, 2018) is presented for a sea state characterized by very rare extremes.

## 2. MATHEMATICAL BACKGROUND

In this section, the second extreme value theorem is formulated and the basic properties of the GPD are outlined. The potential of the model for treating the problem of rarity, described in the above, is discussed in the context of a more general framework, commonly known as the "principle of separation" (e.g. Belenky et al., 2012; Mohamad and Sapsis, 2016).

### *The principle of separation*

The term is often utilized to express the idea of decomposing the ship response problem into sub-problems with the aim of analyzing the rare extremes separately from a background state, mostly associated with conventional non-rare outcomes. Thence, the "non-rare" part deals with the distribution of the conditions that can lead to the occurrence of extreme events, while the "rare" one targets the conditional probability of extremes, given that specific conditions are met:

$$Pr(X > x) = \underbrace{Pr(X > x|X > u^*)}_{rare} \times \underbrace{Pr(X > u^*)}_{non-rare} \quad (1)$$

where $X$ is the response process, $x$ is the associated state variable and $u^*$ is a threshold introduced for distinguishing extreme from non-extreme regimes.

As realized, ship motions have, thus far, been classified with respect to their relative frequency of occurrence (rare/non-rare), rather than according to the corresponding level of nonlinearity governing the dynamics of each sub-problem. In the "rare" part, however, one is confronted with phenomena that are not only very unlikely, but also strongly nonlinear. On the contrary, a "non-rare" event is not essentially linear; neither nonlinearity itself is sufficient to infer rarity. To explicitly account for the effect of nonlinearity also on the solution of the "non-rare" part, the last term in Eq. (1) is further decomposed as:

$$Pr(X > u^*) = \underbrace{Pr(X > u^*|X > u_L)}_{nonlinear} \times \underbrace{Pr(X > u_L)}_{linear} \quad (2)$$

where $u_L$ is an intermediate threshold indicating the limit between linear and nonlinear ship responses within the "non-rare" sub-region. Definitely, through this concept, one could go even deeper by disassembling both the "rare" and "non-rare" sub-problems of Eq. (1) in more parts; yet this would require a rational procedure for selecting those additional intermediate thresholds $u_i, i = 1, \dots, n$ that would separate regimes with different levels of nonlinearity.

In this setting, it is straightforward to calculate the last term in Eq. (2) using a Gaussian distribution. Mathematical justification for the solution of the "rare" sub-problem will be provided by the second extreme value theorem, presented in the following section. As for the probability of non-rare and nonlinear events, there are numerous statistical models to try. In this study, however, the GPD is employed once again knowing that it embodies a large class of distribution functions covering a continuous range of possible shapes. This allows for

the data to decide the most suitable amongst the models integrated into the GPD.

***The generalized Pareto Distribution (GPD)***

Generally, the GPD is specified by three parameters ($u$, $\sigma$, $\xi$) and below it is expressed in terms of its complementary distribution function $F_X(x) = 1 - \bar{F}_X(x)$:

$$\bar{F}_X(x) = \begin{cases} \left(1 + \frac{\xi(x-u)}{\sigma}\right)^{-1/\xi} & , if\ \xi \neq 0 \\ exp\left(-\frac{x-u}{\sigma}\right) & , if\ \xi = 0 \end{cases} \quad (3)$$

where $x \geq u$, if $\xi \geq 0$ and $u \leq x \leq u - \sigma/\xi$, if $\xi < 0$. In Eq. (3), $u$ is the location parameter of the distribution representing the minimum value that the associated random variable $X$ can attain. Whenever the GPD is employed for modelling the tail of another distribution, $u$ is basically the point where the two distributions merge. The scale parameter $\sigma$ is the "spread" factor, controlling the dispersion of $X$ above $u$. Finally, $\xi$ affects the shape of the distribution in a more qualitative way. For distributions with exponentially decreasing tails, such as the Normal, the GPD leads to $\xi = 0$. For heavy-tailed distributions, often encountered in the case of unbounded systems, $\xi > 0$. The opposite ($\xi < 0$) implies a light-tailed distribution and thus, the existence of an upper bound at $x = u - \sigma/\xi$.

The theoretical importance of Eq. (3) was proved by Balkema and de Haan (1974) and Pickands (1975) who showed that the distribution of independent and identically-distributed (i.i.d.) excesses over $u$ asymptotically tends towards the GPD, as $u \to \infty$. The statement holds if and only if the parent distribution belongs to the so called "domain of attraction" of one of the extreme value distributions (i.e. Gumbel, Fréchet and reverse Weibull), all incorporated into a single model, known as the generalized Extreme Value distribution (GEV). Moreover, it can be verified that if times until exceedance constitute a Poisson random process with GPD excesses, then the GEV is obtained as the distribution of the corresponding extremes. Another interesting property of the GPD is "threshold stability", meaning that if $X$ is a GP-distributed random variable for some $u^* > 0$, then it is also generalized Pareto for any $u > u^*$ retaining the same shape parameter. It is worth noting that the GPD is uniquely characterized through the last two properties since no other family of distributions exhibits such qualities (Davison and Smith, 1990).

***Threshold selection***

On these terms, it is rather natural to assume that local stabilization of the shape parameter could be the key for detecting the minimum threshold value above which the distribution of excesses has practically converged to the GPD. The idea has been discussed in several studies, often in comparison with alternative identification procedures, such as those described in e.g. Campbell et al. (2016). Yet, the threshold stability property itself could be the source of inherent limitations in pinning down the threshold. If a dataset obeys the GPD at one threshold, then, the model, in order to preserve its validity at all higher thresholds, should be free to adapt through its only left unconstrained parameter, i.e. the scale parameter. Equally, restricting the threshold to a fixed value in an attempt to extrapolate a trend into the tail region could entail the possibility of overfitting.

The invariance of the model form at high levels was an additional motivation for investigating the tail structure by employing successive GPDs in Eqs. (1) and (2). Even though consistency with the theorem may not be fulfilled for $u_L$, being essentially the angle up to where ship motions are relatively small, the GPD, due to its very flexibility, will probably succeed in fitting data within the intermediate range $[u_L, u^*]$. The crucial step, however, is the selection of $u^*$ so as to reflect a lower bound for the occurrence of extremes. From a ship design perspective, the angle $\varphi_{max}$ corresponding to the maximum of the GZ curve could be tried since rolling beyond this limit is quite likely to result in capsize or, at least, in an extreme dynamic event.

## 3. RESULTS AND DISCUSSION

Massive Monte Carlo simulations were performed for an ocean surveillance ship, with main parameters listed in Table 1, to evaluate the accuracy of the GPD-based approach presented in the above. The concept of separation, as expressed through Eqs. (1) and (2), is illustrated in Figure 1, where the GZ curve of the vessel is divided into three sub-regions with limits indicated by vertical lines:

I. $\varphi \in [0, u_L]$, with $u_L = 20\text{deg}$

II. $\varphi \in [u_L, u^*]$, with $u^* = 37\text{deg}$

III. $\varphi > u^*$

**Table 1: Main parameters of the vessel.**

| Parameter | Dimensional value | Dimensions |
|---|---|---|
| $I + A_{44}$ | $5.540 \times 10^7$ | $kg \cdot m^2$ |
| $\Delta$ | $2.056 \times 10^6$ | $kg$ |
| $B_1$ | $5.263 \times 10^6$ | $kg \cdot m^2/s$ |
| $B_2$ | $2.875 \times 10^6$ | $kg \cdot m^2$ |
| $C_1$ | 3.167 | $m$ |
| $C_3$ | $-2.513$ | $m$ |

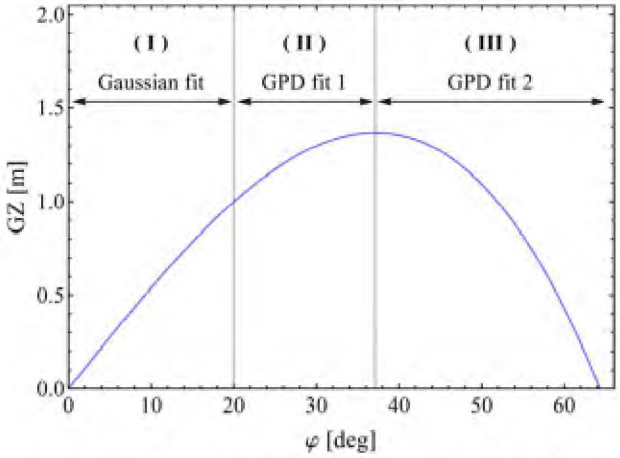


**Figure 1: The restoring arm of the vessel divided into sub-regions: (I) non-rare/linear, (II) mildly rare/nonlinear and (III) rare/nonlinear.**

The ship is assumed to operate in sea conditions described by the Bretschneider spectrum (Ochi, 1998):

$$S_{\eta\eta}(\omega) = \frac{1.25}{4}\frac{\omega_p^4}{\omega^5}H_s^2 exp\left[-\frac{5}{4}\cdot\left(\frac{\omega_p}{\omega}\right)^4\right] \quad (4)$$

where $H_s$ is the significant wave height and $\omega_p = 2\pi/T_p$ is the modal frequency of the spectrum. Two sea states of slightly different severity were duly selected for demonstrating certain capabilities and limitations of the proposed method. Their characteristics are given in Table 2.

**Table 2: Sea state characteristics.**

| | $H_s$ | $T_p$ |
|---|---|---|
| Sea state *A* | 4m | 11s |
| Sea state *B* | 3m | 11s |

Time-histories of roll motion $\varphi(t)$ were generated using a simple 1DOF roll equation:

$$(I + A_{44})\ddot{\varphi} + B_1\dot{\varphi} + B_2\dot{\varphi}|\dot{\varphi}| + \\ +g\Delta(C_1\varphi + C_3\varphi^3) \\ = M(t) \quad (5)$$

with $I+A_{44}$ being the total roll moment of inertia (including the added mass effect), $g$ is the gravitational acceleration, $\Delta$ is the ship displacement and $B_1$, $B_2$ and $C_1$, $C_3$ are the damping and restoring coefficients, respectively. The wave-induced moment was modelled using the standard spectral representation method (St. Denis and Pierson, 1953):

$$M(t) = \\ \sum_n \sqrt{2S_{\eta\eta}(\omega_n)F_{roll}(\omega_n)\delta\omega_n}\cos\theta_n(t) \quad (6)$$

where $\theta_n(t) = \omega_n t + \varepsilon_n$. In Eq. (6), $\varepsilon_n$ are random variables uniformly distributed over $[0,2\pi)$, $d\omega$ is the frequency resolution, $A_n$ are the amplitudes of the wave components and $\omega_n$ are the associated frequencies. Details for the roll moment amplitude operator $F_{roll}$ of the vessel can be found in Su (2012).

Eventually, statistics of roll motion were obtained without assuming the ergodic property for the response. Consequently, the analysis was made on a set of 6,000,000 short-duration realizations, sampled at a fixed time instant $t_s = 150\text{s}$. The great benefit from this approach is that collected roll data are statistically independent, as required by the second extreme value theorem. Roll data were partitioned in 15 datasets and for each dataset, the GPD parameters (shape and scale) were calculated using the Maximum Likelihood Estimation method (MLE). The mean values of the 15 pairs of parameters were selected as the most representative values for fitting the whole dataset (6,000,000 samples).

***Sea state scenario A***

In this case study the objective is to evaluate the reliability of the GPD for calculating the probability of capsize when data are available in the entire range of stability $[0, \varphi_v]$, where $\varphi_v = 64\text{deg}$ is the angle of vanishing stability of the vessel. The selection of the capsize limit was based on the well-known feature of Eq. (5) concerning the time-depending shifting of the unstable equilibrium in the presence of wave excitation (e.g. Falzarano et al., 1992). In this regard, response trajectories that exceeded (in absolute

sense) the limiting value $1.2\varphi_v$ before reaching $t_s =$ 150s were marked as corresponding to capsize, resulting in a total number of 393 capsizes for the specific sea state. No doubt, considering stability failure at an exceptionally high roll angle is unreasonable since flooding is very likely to occur at lower angles. As a matter of fact, it is sufficient to confirm accuracy in GPD predictions only up to intermediate roll angles representing practical capsize limits (e.g. 40deg-50deg). For scientific curiosity reasons however, and since a similar model could be the subject of investigation in a different (non-marine) context, the tail region $[\varphi_{max}, \varphi_v]$ is examined in its entirety just for highlighting particular features of the ship rolling process that may not be so evident at lower levels.

Next, results are first presented for the case of "bounded" ship motions, meaning that desired statistics were computed after filtering out the 393 capsize cases. As realized, eliminating the possibility of capsize may conceal valuable information for our analysis. It is, nevertheless, interesting to investigate the effectiveness of traditional techniques of EVT, such as the POT/EPOT (peaks or envelope peaks over threshold) methods, which rely solely on the peak excesses of a random process for fitting the GPD. Since a "peak" by definition implies the return of a response trajectory towards the upright state, it is clear that these methods deal with a qualitatively different problem where the underlying system remains always bounded. On the contrary, in our approach the GPD is fitted to all the exceedances recorded at the selected sampling instant $t_s$, regardless of being peaks.

Figure 2 shows the probability of exceedance $P_e$ of rolling angles $\varphi \in [0, \varphi_v]$ derived from the Monte Carlo (MC) simulations through direct counting (circles) for the bounded system. The solution of the linear "non-rare" sub-problem, being the Gaussian fit curve (solid line), is extended up to region (III) for comparison purposes. Dashed lines indicate the solution of the combined nonlinear sub-problem ("non-rare" + "rare"), obtained by two individual GPD fits; one in region (II) and one in region (III). The shaded area illustrates the corresponding 95% confidence interval (CI). In analogy to Figure 1, vertical lines denoting the limits of regions (I-III) are included. Details for the estimated GPD parameters are provided in Table 3.

**Table 3: GPD fitting results (bounded system).**

| region II: [ 20deg, φmax ] | | | |
|---|---|---|---|
| scale parameter | | shape parameter | |
| mean value | 95% CI | mean value | 95% CI |
| 3.471 | [3.443, 3.499] | -0.024 | [-0.029, -0.018] |
| **region III: [ φmax, 64deg ]** | | | |
| scale parameter | | shape parameter | |
| mean value | 95% CI | mean value | 95% CI |
| 3.550 | [3.240, 3.860] | 0.071 | [-0.019, 0.161] |

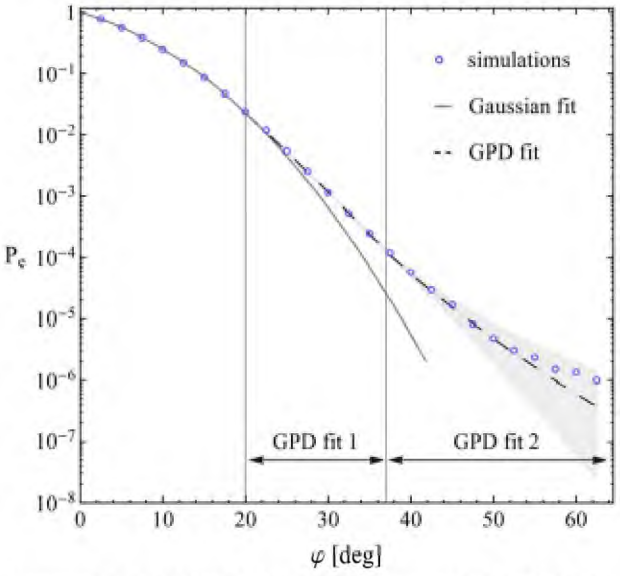


**Figure 2: GPD fits (dashed lines) vs. rough Monte Carlo estimates (circles) for the bounded system.**

As observed, there is good coincidence between the proposed calculation scheme and the MC results in the entire range of stability of the vessel. Moreover, the negative shape parameter in region (II) confirms the existence of a right boundary, as anticipated. Despite that, a heavy tail is eventually obtained since in region (III) the shape parameter becomes positive, yet with the associated confidence interval containing also negative values. The fact that the method fails to maintain the light tail trend in region (III) is, therefore, explained by the uncertainties arising in the estimation of the shape parameter at higher levels, where data are naturally fewer. Finally, it is remarkable that there is less discrepancy in the computation of the scale parameter, given that its value is practically the same in both regimes.

Below, the assumption of bounded motions is removed to assess the validity of the treatment presented so far. To this end, statistics were derived separately for threshold exceedances that led to capsize ("escapes") and for short-time exceedances that remained bounded in the long run ("returns"). In

Figure 3, $P_c$ is the conditional probability of a return/escape, given that a roll angle threshold $\varphi$, displayed on the horizontal axis, has already been exceeded. Since $P_c$ is, in fact, the ratio of observed escapes/returns to the total number of exceedances over a threshold $\varphi \in [0, \varphi_v]$, this plot essentially reflects the contribution of each outcome to the overall probability of exceedance $P_e$. A circle has been placed on the curve of the escapes at $u^* = \varphi_{max}$ to highlight that in region (III) extremes are, at least, 34% underpredicted with respect to their "true" values that would be obtained if capsizes had been included in the calculations. This demonstrates the necessity of developing methods free of POT/EPOT-like assumptions, often introduced in the light of "strict-sense stationarity" of ship response (Kougioumtzoglou and Spanos, 2014).

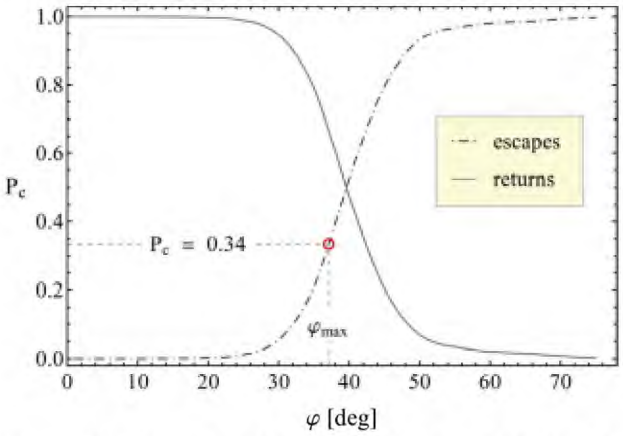


**Figure 3: Contribution of escapes and returns to the total probability of a threshold exceedance.**

At the same time, Figure 3 reveals new locations for potential thresholding, other than those used in this study. Even more, one may be tempted to analyze individual sub-problems into more parts than those proposed here. For instance, one could perform the decomposition over both/either the point where the two curves intersect (e.g. at 40deg) and/or the angle where the maximum curvature on the escapes curve is observed (e.g. at 50deg). However, proceeding to exhaustive decompositions is not recommended because information could be lost due to the separation principle itself. The concept assumes that threshold exceedances have negligible dependence on the statistics below the threshold. In this sense, it may be more difficult to capture the whole picture when approximating the solution through a large number of conditionals, considering that extremes may not eventually be consistent with the mechanism that generates the main body of the data.

In Figure 4, the probability of exceedance $P_e$ was derived by analyzing the entire sample, including the 393 capsize cases. The notation is the same as in Figure 2. The results of the corresponding GPD fitting procedure are summarized in Table 4. As noticed, the MC trend (circles) implies a heavy tail that in region (III) becomes almost parallel to the $\varphi$-axis. However, this cannot be inferred from the GPD model of the current method (dashed lines). Evidence for the tail structure has already been given in Figure 3 where it is shown that above 40deg exceedance probabilities are mostly determined by the escaping trajectories. With returns gradually vanishing in the very extreme region ($\varphi > 50$deg), the probability of exceedance $P_e$ naturally tends to the probability of capsize (393 capsizes / 6,000,000 samples). In Figure 4, this resulted in almost two orders of magnitude greater probabilities than those presented in Figure 2 for the bounded system.

**Table 4: GPD fitting results (unbounded system).**

| **region II: [ 20deg, φmax ]** | | | |
|---|---|---|---|
| scale parameter | | shape parameter | |
| mean value | 95% CI | mean value | 95% CI |
| 3.221 | [3.200, 3.243] | 0.085 | [0.080, 0.090] |
| **region III: [ φmax, 64deg ]** | | | |
| scale parameter | | shape parameter | |
| mean value | 95% CI | mean value | 95% CI |
| 7.038 | [5.526, 8.551] | 0.781 | [0.657, 0.904] |

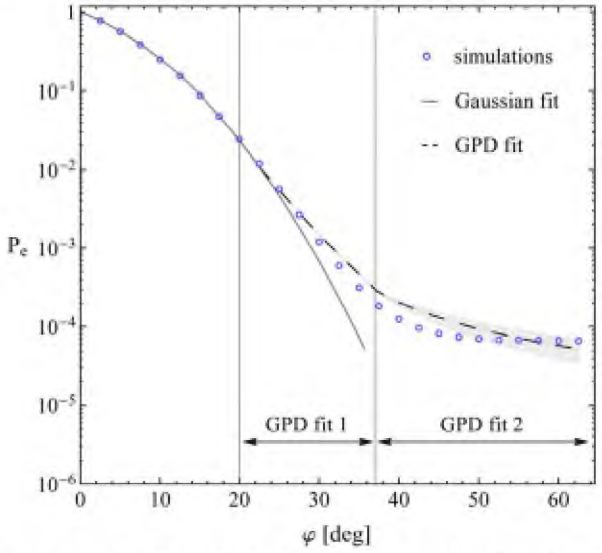


**Figure 4: GPD fits (dashed lines) vs. rough Monte Carlo estimates (circles) for the unbounded system.**

The poor performance of the method could stem, at least partially, from the very special shape of the roll response distribution, provided that validity of the GPD is asserted only if the underlying distribution belongs to the domain of attraction of the GEV distribution. However, there is no proof on whether ship rolling truly fulfills this requirement. It is remarkable, though, that in region (III) the produced GPD is characterized by a large positive shape parameter, indicating that the model realizes the qualitative changes induced by the now existing possibility of system escape. Again, large uncertainties are observed in region (III), despite the presence of quantitatively more extremes than in the case of bounded motions.

Although having a "rich", in terms of capsize occurrences, sample enhances the reliability of the MC estimates, the coexistence of states with distinct dynamics (escapes-returns) entails technical difficulties in their joint statistical description. Specifically, to calculate exceedance probabilities from a sample containing aggregated data of escapes and returns, one has to define the "capsize state" (here noted as $\varphi_\infty$). In our MC setup, if a response trajectory exceeded the capsize limit $1.2\varphi_v$ at some time instant $t < t_s$, the integration of Eq. (5) was terminated and a fixed value $\varphi_\infty = 1.2\varphi_v$ was eventually kept for further analysis (e.g. Figure 4). On the other hand, assigning the exact same value $\varphi_\infty$ to all capsized cases will inevitably result in artificial "mass" concentration in the corresponding probability density function (PDF), as illustrated in Figure 5. Due to the data-driven nature of fitting procedures, the location of this mass is expected to affect the calculation of the GPD parameters.

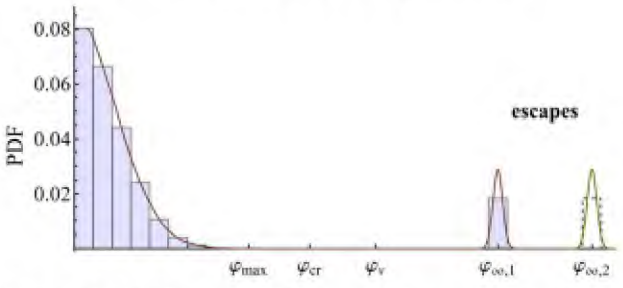


**Figure 5: Sensitivity of GPD estimates to the statistical description of escapes.**

The sensitivity of the current method to the selection of $\varphi_\infty$ in calculating the probability of exceeding 50deg was investigated for four $\varphi_\infty$ scenarios ($1.2\varphi_v$, 80deg, 90deg and 180deg). The results confirmed that by changing the relative distance between $\varphi_\infty$ and the main probability mass the GPD monotonically overestimates (from 1.2 up to 2.4 times) the corresponding probability obtained from the MC simulations in Figure 4. Hence, in Figure 4, the GPD was shown in its utmost performance since setting $\varphi_\infty > 1.2\varphi_v$ would certainly deteriorate its accuracy. This is because the sample variance in region (III) varies through $\varphi_\infty$ and thus, the GPD adapts, although not very successfully, to the data. This sensitivity justifies why the scale parameter is larger in Table 4 than in Table 3, where the $\varphi_\infty$ parameter is not involved.

***Sea state scenario B***

Lowering $H_S$ by only 1m leads to substantial changes in ship behavior, given that for the specific sea conditions all the collected observations were below $\varphi_{max}$ (no capsizes recorded). Therefore, the interest here lies in utilizing the GPD for predicting events that are considerably more extreme than those found in the available simulation data. However, evaluating the "extrapolation" quality of the model having only few nonlinear/extreme data is a non-trivial task. One idea could be to compute the percentage of datasets (out of the 15 partitions) with associated GPD estimates containing within their confidence band the "true" probability of stability failure (obtained by analyzing the entire sample). Although straightforward, the approach would still suffer from uncertainty issues due to the calculation of the GPD parameters from essentially small subsets (Weems et al., 2016).

To avoid such problems, in this study the GPD trends are compared with the predictions of the "critical wave groups" method (Themelis and Spyrou, 2007). Unlike the method presented here, the "critical wave groups" scheme does not make any assumptions regarding the shape of the distribution of extreme responses. Instead, it quantifies instability tendency implicitly, through the probability of encountering any wave group that could provoke the instability using distributions describing statistical properties of the wave field. Recently, the potential of the method for handling the rarity of extremes was demonstrated by Anastopoulos and Spyrou (2018).

Considering that the largest (in absolute sense) observed roll angle was only 35deg, the GPD was first fitted to exceedances over $u_L = 20$deg. Then, it

was extrapolated to region (III). In Figure 6, the results (dashed line) are tested against the MC (circles). As before, the shaded area refers to the 95% confidence intervals (CI) of the associated GPD parameters, given in Table 5. Information for the "critical wave groups" (CWG) probabilities is directly available from the work of Anastopoulos and Spyrou (2018) who applied the method to the vessel examined here and for the same sea conditions. Although their results cover the entire nonlinear part $[u_L, \varphi_v]$, in this investigation, staying below 40deg seems to be sufficient for reaching conclusions since from very early (25-30deg) the GPD and the "critical wave groups" curves exhibit different trends. Besides, due to lack of data in region (III), the GPD confidence band will become excessively large. Finally, in this plot, the Gaussian fit (solid line) is not extended beyond region (I) because it was noticed that the result would be almost identical with the GPD curve. This is explained by the large confidence interval of the shape parameter in Table 5, indicating that nonlinear data are very few (only 2‰ of the sample size) because the ship spends most of the time below $u_L$. The GPD captures this feature but without further guidance it cannot do more than to extrapolate Gaussianity also in region (II).

Despite the unambiguous linear character of the GPD in Figure 6, the negative shape parameter in Table 5 suggests that the model eventually turns into a light tail. Since the probability of capsize is, in general, non-zero (even for this seemingly innocuous sea state), a heavy tail should be expected. Here, though, it is masked by the problem of rarity, leading the GPD to assume the existence of a physical boundary. This is the reason for the deviation between the GPD and the "critical wave groups" curves. The latter succeeds in tracing the unobserved heavy tail and because of its consistency with the MC values from lower levels (20-25deg) one could argue that it is more appropriate for extrapolation in the specific sea conditions. However, in view of the inherent uncertainties in the interpretation of direct counting estimates when data in the range of interest are very few, comparing results obtained from techniques originating from different principles would, at least, contribute towards their mutual development, if not achieving the ultimate validation goal.

**Table 5: GPD fitting results (no escapes observed).**

| region II: [ 20deg, φmax ] | | | |
|---|---|---|---|
| scale parameter | | shape parameter | |
| mean value | 95% CI | mean value | 95% CI |
| 1.893 | [1.836, 1.951] | -0.051 | [-0.071, -0.032] |

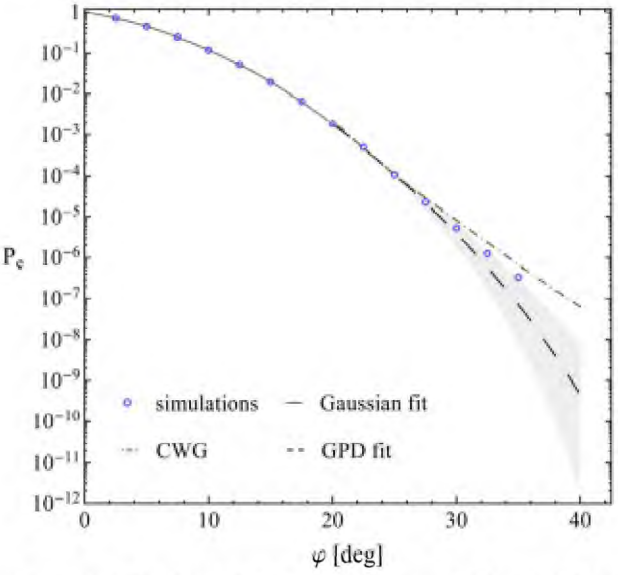


**Figure 6: GPD fit (dashed line) vs. rough Monte Carlo estimates (circles) and comparison with the "critical wave groups" (CWG) predictions (dot-dashed line).**

## 4. SUMMARY AND CONCLUSIONS

A method based on Extreme Value Theory (EVT) was proposed for calculating the probability of exceeding exceptionally high roll angles in beam seas. The method analyses the ship response problem into three parts (sub-problems), each associated with a different level of rarity and/or nonlinearity. For the first part, targeting statistical description of small-amplitude motions, a Gaussian distribution was utilized. In the nonlinear part, the solution was composed by fitting the generalized Pareto Distribution (GPD) to roll exceedances over two levels: a) a "non-rare" intermediate threshold and b) a "rare" extreme threshold. The selection of these thresholds was based on the shape of the GZ curve which provides indications for the limits of regimes where the response distribution exhibits qualitatively different probabilistic characteristics.

The performance of the approach was tested against the rough estimates of Monte Carlo simulations, obtained by direct counting. Several aspects regarding the implementation of the approach were discussed and particular attention was given to the problem of capsize. The results reveal that, given "sufficient" data, the method can accurately determine the probability of extreme

dynamic events, yet if the possibility of system escape is practically zero. However, information is essentially lost due to this assumption since escaping induces qualitative changes in the shape of the response distribution. In the case of unbounded motions, though, the GPD-tail produced by our method could not fit the data successfully. In the light of this finding, one could speculate that ship capsize is not within the range of applicability of classical EVT tools.

Finally, the "statistical extrapolation" character of the approach was evaluated through a comparison with the predictions of the "critical wave groups" method. In this context, preliminary evidence suggests that, for the examined sea conditions, the latter may be more suitable for making inferences beyond the largest observation. However, further investigation is definitely needed for reaching more general conclusions. Towards this direction, assessing methods with different backgrounds against each other seems the only option for their mutual validation in regimes where extremes cannot be directly "seen" through straightforward Monte Carlo procedures. It should be noted, though, that the desire of controlling the probability of stability failures directly through the ship design parameters requires, in fact, knowledge of the GPD form at a time when simulation data are often not available. Therefore, even if the effectiveness of the GPD idea is eventually proven, this very desire will presumably remain unsatisfied due to the data-driven nature of the concept itself.

## ACKNOWLEDGEMENTS

The work of Mr. Anastopoulos was partly supported by NTUA's Special Account for Research Grants (ELKE). The idea of performing multiple GPD fits came up during a discussion with Dr. Vadim Belenky. Both Dr. Vadim Belenky and Mr. Kenneth Weems (David Taylor Model Basin, NSWCCD) are thanked for providing valuable feedback during this study.

## REFERENCES

Anastopoulos, P.A. and Spyrou, K.J., 2018, "Evaluation of an Improved Critical Wave Groups Method for the Prediction of Extreme Roll Motions", Proceedings of the 13th International Conference on the Stability of Ships and Ocean Vehicles (STAB2018), Kobe, Japan, pp. 565-574.

Balkema, A. and de Haan, L., 1974, "Residual Life Time at Great Age", *Annals of Probability* 2, pp. 792-804.

Belenky, V., Weems, K.M., Bassler, C.C., Dipper, M.J., Campbell, B.L. and Spyrou, K.J., 2012, "Approaches to Rare Events in Stochastic Dynamics of Ships", *Probabilistic Engineering Mechanics* 28, pp. 30-38.

Belenky, V., Weems, K. and Lin, W.M., 2016, "Split-Time Method for Estimation of Probability of Capsizing Caused by Pure Loss of Stability", *Ocean Engineering* 122, pp. 333-343.

Campbell, B., Belenky, V. and Pipiras, V., 2016, "Application of the Envelope Peaks Over Threshold (EPOT) Method for Probabilistic Assessment of Dynamic Stability", *Ocean Engineering* 120, pp. 298-304.

Chai, W., Dostal, L., Naess, A. and Leira, B.J., 2017, "Comparative Study of the Path Integration Method and the Stochastic Averaging Method for Nonlinear Roll Motion in Random Beam Seas", *Procedia Engineering* 199, pp. 1110-1121.

Davison, A.C. and Smith, R.L., 1990, "Models for Exceedances over High Thresholds", *Journal of the Royal Statistical Society. Series B (Methodological)* 52, pp. 393-442.

Falzarano, J.M., Shaw, S.W. and Troesch, A.W., 1992, "Application of Global Methods for Analyzing Dynamical Systems to Ship Rolling Motion and Capsizing", *International Journal of Bifurcation and Chaos* 2, pp. 101-115.

Mohamad, M.A. and Sapsis, T.P., 2016, "Probabilistic Response and Rare Events in Mathieu's Equation under Correlated Parametric Excitation", *Ocean Engineering* 120, pp. 289-297.

Ochi, M., 1998, "Ocean Waves: The Stochastic Approach", Cambridge University Press, Cambridge, England, ISBN: 978-0-521-01767-1.

Pickands, J., 1975, "Statistical Inference Using Extreme Order Statistics", *Annals of Statistics* 3, pp. 119-131.

Kougioumtzoglou, I.A. and Spanos, P.D, 2014, "Stochastic Response Analysis of the Softening Duffing Oscillator and Ship Capsizing Probability Determination via a Numerical Path Integral Approach", *Probabilistic Engineering Mechanics* 35, pp. 67-74.

St. Denis, M. and Pierson, W.J., 1953, "On the Motions of Ships in Confused Seas", *SNAME Transactions* 61, pp. 280-332.

Su, Z., 2012, "Nonlinear Response and Stability Analysis of Vessel Rolling Motion in Random Waves Using Stochastic Dynamical Systems", PhD thesis, Texas A&M University, United States.

Themelis, N. and Spyrou, K.J., 2007, "Probabilistic Assessment of Ship Stability", *SNAME Transactions* 115, pp. 181-206.

Weems, K., Belenky, V. and Campbell, B., 2016, "Validation of Split-Time Method with Volume-Based Numerical Simulation", Proceedings of the 15th International Ship Stability Workshop (ISSW2016), Stockholm, Sweeden, pp. 103-108.