32nd Symposium on Naval Hydrodynamics
Hamburg, Germany, 5-10 August 2018

Tail Structure of Roll and Metric of Capsizing in Irregular Waves

Vadim Belenky[1], Kenneth Weems[1], Vladas Pipiras[2], Dylan Glotzer[2] and Themis Sapsis[3]

([1]David Taylor Model Basin / NSWCCD, Maryland, USA,
[2]University of North Carolina at Chapel Hill, North Carolina, USA,
[3]Massachusetts Institute of Technology, Massachusetts, USA)

ABSTRACT

Extreme values of ship motions and loads in irregular waves are of considerable interest for both the analysis of prospective designs and the development of operational guidance. However, the direct evaluation of these extremes via Monte-Carlo methods is cost-prohibitive, as they are too rare to be predicted by a model test or numerical simulation of sufficient fidelity in realistic sea conditions. This paper describes recent developments in extreme value assessment procedures that are based on characterizing the extreme response from a finite set of numerical simulations. The specific focus of the present work is to use physical considerations to determine the type of tail of a distribution for a random variable that characterizes the extreme motion. This essentially adds physical information into the statistical model and may decrease the statistical uncertainty associated with extreme value prediction. The paper considers tails of two random variables: peaks of roll motions and values of the capsizing metric from the split-time method.

A single degree-of-freedom dynamical system with piecewise linear restoring is used for theoretical considerations. Such a system, with certain assumptions, allows for closed-form solutions of the distributions of large roll angles and the capsizing metric. A piecewise linear system with triangular stiffness is a qualitative model of a dynamical system with softening stiffness (reduced restoring at larger roll angles), so these closed-form solutions reveal the principle form of the tails of the considered variables. Distributions of the roll peaks must have heavy tails, while the distribution of the capsizing metric is expected to have an exponential tail.

Large-volume samples of ship motion in random irregular seas were generated using rapid volume-based numerical simulation. Small subsets of these samples were then used for extreme value predictions that incorporate theoretical types of the tails. Systematic comparison was made to confirm the structure of the tails of roll peaks and the capsizing metric.

INTRODUCTION

The need to use extrapolation in order to obtain extreme motions and loads from numerical calculations for a ship in irregular waves is a result of the transition to time-domain calculations. These codes incorporate nonlinearities that are not possible to account for in the frequency domain and lead to a non-Gaussian response even as the waves themselves remain Gaussian. The most significant nonlinearities associated with motions and loads of a ship in waves are related to the changing geometry of the submerged portion of the hull. As the influence of these nonlinearities is strongest in the largest motions, it is the tail of the distribution that is of practical interest.

The problem of the assessment of extreme values is typically formulated as an estimation of the probability of exceedance of a given large value. Less commonly, it may be formulated as the largest value that would be observed in a given time duration. Both of these formulations require an approximation of the tail of the distribution of the quantity of interest.

The approximation of the tail has to be performed using the results of time-domain numerical simulations, as that is the only available information of the ship response. As relevant large values of the response may not be observed during simulations of practical length, the estimation of the probability of exceedance of a large value is essentially an extrapolation task. Belenky, *et al.* (2012) reviews three possible methods of extrapolation applicable for a ship's dynamical response:

- Split-time method
- Peak-over threshold (POT)
- Critical wave group method

The first two methods are based on the theory of extreme values, which establishes a limit distribution of the largest value in a sample (the first extreme value theorem). This distribution is commonly referred to as the Generalized Extreme Value (GEV) distribution.

Using GEV, an approximation of a distribution above a "large-enough" threshold (the second extreme value theorem) is derived, which is known as the Generalized Pareto distribution (GPD). The second extreme value theorem is the mathematical background of the Peaks-over-Threshold method (Pickands, 1975).

Both theorems are applicable without limitations only for independent samples. As ship response is correlated (due to inertia and correlated excitation), independent points need to be extracted from the sample with a procedure usually referred to as "de-clustering." One convenient way is to use an envelope. Peaks of the envelope of a ship response are far enough apart to be independent (Fig. 1). This particular application sometimes is referred as EPOT – Envelop Peaks over Threshold (Campbell, *et al.* 2016). EPOT has been applied to large roll and pitch angles and excessive transversal and vertical accelerations.

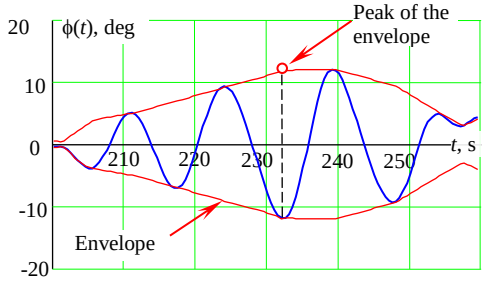

Figure 1: De-clustering with an envelope

The split-time method is intended for estimating the probability of complex and rare physical phenomena in which the physics of the problem changes with the extreme response, such as capsizing and broaching-to. The method is being developed under the US Office of Naval Research (ONR) funded project "A Probabilistic Procedure for Evaluating the Dynamic Stability and Capsizing of Naval Vessels" (Belenky, *et al.* 2016). As capsizing cannot be observed during a simulation or set of simulations of reasonable length, a special metric of capsizing likelihood is introduced. It is computed at the instant of upcrossing of an intermediate roll threshold. The roll rate is perturbed until capsizing is observed (Fig. 2). The difference between the roll rate at upcrossing and the roll rate when capsizing is observed is the metric, as it indicates "the distance to trouble."

$$y_i = 1 + \dot{\phi}_{Ui} - \dot{\phi}_{Ci} \; ; \quad i = 1,\ldots, N_U \qquad (1)$$

$\dot{\phi}_{Ci}$ is the critical roll rate calculated for the $\underline{i}$-th upcrossing, and $\dot{\phi}_{Ui}$ is the roll rate observed at the i-th upcrossing, N_U number of observed upcrossings.

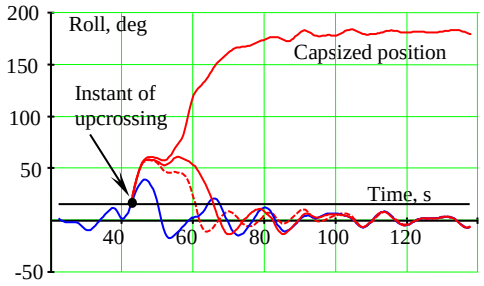

Figure 2: Motion perturbation for computing the capsizing metric

Computation of the metric values (1) over a number of upcrossings creates a sample that can be used for extrapolation with GPD. For 3 DOF simulations, the de-clustering technique can be based on convergence time, which is when the perturbed solution converges back to the original unperturbed solution. This approach does not work for 6 DOF system, for which a technique based on an averaged decorrelation time is used instead (Weems and Belenky 2017).

The critical wave group method is based on a completely different approach, though the principle remains the same: consider a complex nonlinear dynamical phenomenon and extreme rarity separately. The idea of the wave group method is to extract groups of large waves and consider them separately, assuming that the phenomena of interest are caused by rare groups of large waves separated by benign waves, see Fig 4.

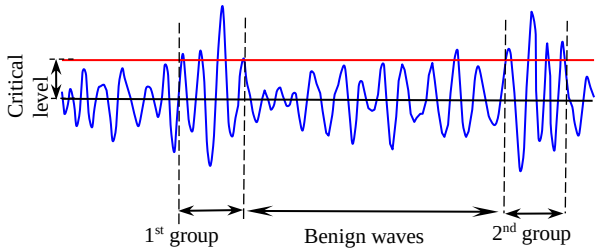

Figure 3: On critical wave group approach

The first complete application of the wave group method for the probabilistic assessment of dynamic stability was carried out by Themelis and Spyrou (2007). Several theoretical descriptions of wave groups are available in the literature. One such description is Quasi-Determinism (QD), which is reviewed in Boccotti (2014). QD defines the wave group profile around a large wave, thus significantly decreasing the number of parameters needed to characterize the wave groups. Descriptions of recent developments, as well as an examination of the application of critical wave group method to model tests in a wave basin, can be found in Anastopoulos, *et al.* (2015) and Romolo, *et al.* (2016).

However attractive from a theoretical perspective, the practical application of these extrapolation methods require robust validation. What does it mean in relation to an extrapolation method and to rare events? The most current description can be found in Smith and Zuzick (2015). The validation of extrapolation requires very fast, but at least qualitatively correct simulation code, capable of reproducing data sets comprising millions of hours of irregular sea response, making even rare events observable. Small portions of this simulation date set are then used with the extrapolation method being tested. The extrapolated result is compared with a true value observed from the full data set. The most recent description of such code is available from Weems, *et al.* (2018). This reference, along with Smith and Zuzick (2015), give a complete picture of the validation attempts of the EPOT and split-time methods.

Basic testing of the critical wave group method against a Mont-Carlo result was carried out by Malara, *et al.* (2014) using the Duffing equation. As the latter can be solved very quickly, it can be used as a source of rare events and extreme values associated with them. Validation using fast ship motion simulations remains for the future work.

One of the outcomes if the validation of extrapolation with EPOT and split-time was the realization that the confidence interval of the extrapolated estimate is generally not small. While the lower boundary has mostly academic value, the upper boundary of the confidence interval of the extrapolated estimate is actually the value that constitutes the correct answer to use in any risk-based analysis. Some cases have shown this boundary to be two orders of magnitude or more above the true value. While some applications may accept an answer at this level of uncertainty, it would be highly desirable to decrease the uncertainty without increasing the volume of the sample. Can this be done?

The answer is actually yes. Fitting GPD to a sample of peaks or capsizing likelihood metric is a data-driven procedure. Including some physical information into the statistical model may decrease uncertainty. GPD is a two-parameter distribution. Each parameter is a result of fitting procedure (usually through maximum likelihood), which is a sort of averaging procedure performed on a finite data sample. Thus, each estimated parameter carries statistical uncertainty. However, the parameters are dependent as they were fit with the same data.

Glotzer, *et al.* (2017) introduced a ratio between the parameters for extreme pitch motion, assuming a physical limit for a pitch of ONR tumblehome configuration. One can justify the assumption by the existence of a rather long flat portion of the longitudinal GZ curve, where dynamical system is no longer capable of accepting any energy from the excitation since its instantaneous natural frequency becomes zero. This assumption resulted in a significant decrease of the statistical uncertainty.

The GPD approximates a "generic" tail for any data set. Including some physical information into the statistic model may decrease the uncertainty and increase the reliability of the extrapolation. This paper focuses on how it can be done for large roll angles and capsizing metrics by understanding the nature of the tails of their distributions.

STRUCTURE OF A DISTRIBUTION TAIL

As noted in the previous section, GPD is derived from GEV as an approximation for a tail above a sufficiently large threshold. More details can be found in Coles (2001), where the logic of the derivation of GPD from GEV is outlined. Formally, GPD has two parameters: shape ξ and scale σ. Additionally, there is a threshold u; the distribution is only applicable for the random values $y>u$.

$$pdf(y) = \begin{cases} \frac{1}{\sigma}\left(1+\xi\frac{y-u}{\sigma}\right)^{-\left(1+\frac{1}{\xi}\right)}; & if \quad \xi \neq 0 \\ \frac{1}{\sigma}\exp\left(-\frac{y-u}{\sigma}\right); & if \quad \xi = 0 \end{cases} \tag{2}$$

The scale parameter σ is positive, while the shape parameter ξ can be both positive and negative. For the case $\xi=0$, the GPD becomes an exponential distribution for the variable y-u with the parameters $1/\sigma$. A negative shape parameter imposes a limitation on the expression in parenthesis:

$$u < y < u - \frac{\sigma}{\xi}, \quad \xi < 0 \tag{3}$$

This limitation formally introduces a right bound for the negative values of shape parameter:

$$pdf(y) = 0, \quad if \begin{cases} y > u - \frac{\sigma}{\xi} \\ \xi < 0 \end{cases} \tag{4}$$

That right bound was used by Glotzer, *et al.* (2017) to introduce physical information into the statistical model, which led to significant uncertainty decrease.

Shape parameter defines the type of tail: heavy, exponential or light, as shown in Fig. 4. The exponential tail describes the extreme values of a normal distribution. The heavy tail is above the exponential tail, while light tail is below. As the exponential tail is the smallest infinite tail, the light tail has a limit, which is its right bound. The heavy tail is unbounded. So the most basic question to ask is the type of the tail.

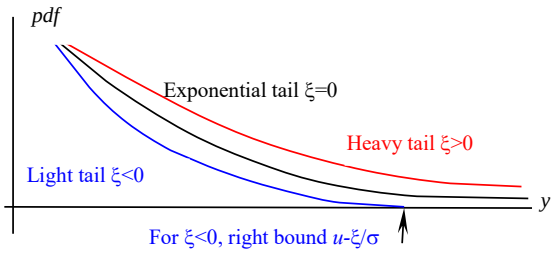

Figure 4: Types of tails

PIECEWISE LINEAR (PWL) MODEL

PWL Model as a Nonlinear Dynamical System

A single degree-of-freedom dynamical system with piecewise linear restoring function is a convenient object for theoretical study as it offers closed-form solutions for many interesting problems. Consider the differential equation:

$$\ddot{\phi} + 2\delta\dot{\phi} + \omega_0^2 f_L(\phi) = f_{E\phi}(t) \qquad (5)$$

where δ is a linear damping coefficient and $f_{E\phi}$ is a stationary stochastic process of roll excitation, while the roll restoring f_L is shown in fig.5.

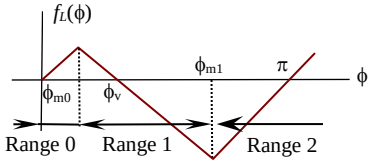

Figure 5: Piecewise linear restoring term

The differential equation (5) has a closed form solution within each range. As the equation (5) is linear within each range, the general solution is presented as a sum of a general solution of the autonomous equation and a particular solution of heteronomous equation:

$$\phi = \begin{cases} \phi_a e^{-\delta t}\sin(\omega_{d0}t + \varepsilon) + p_0(t) \\ \qquad if \quad -\phi_{m0} \le \phi \le \phi_{m0} \\ Ae^{\lambda_1 t} + Be^{\lambda_2 t} + p_1(t) + \phi_V \\ \qquad if \quad \phi_{m0} < \phi \le \phi_{m1} \\ \phi_{a2} e^{-\delta t}\sin(\omega_{d2}t + \varepsilon_2) + p_2(t) + \pi \\ \qquad if \quad \phi > \phi_{m0} \end{cases} \qquad (6)$$

ϕ_a, ε, A, B, ϕ_{a2} and ε_2 are arbitrary constants that are dependent on the initial conditions at the "switching" of the ranges; ω_{d0} and ω_{d2} are frequencies of the damped oscillation in ranges 0 and 2, respectively; λ_1 and λ_2 are eigenvalues for the solution in Range 1. p_0, p_1 and p_3 are particular solutions, similar in shape to the excitation $f_{E\phi}$, *i.e.* stationary stochastic processes.

The system (5) has all the properties of a nonlinear system; Belenky (2000) describes its nonlinear properties, typical for the softening nonlinearity including:

- Loss of isochronism, *i.e.* dependence of natural frequency on initial amplitude
- Fold bifurcation: coexistence of low- and high-amplitude stable response to mono-periodic excitation of the same frequency, observed for excitation frequencies slightly lower than natural frequency
- Flip bifurcation: response period doubling sequence, leading to deterministic chaos under mono-periodic excitation with frequency slightly higher than natural frequency
- Erosion of the safe basin (set of initial conditions in the phase plane that do not lead to capsizing) for large amplitude mono-periodic excitation, with the frequency close to natural frequency

These phenomena are considered to be essential nonlinear behaviors observed for a nonlinear dynamical system with softening nonlinearity – e.g. Duffing equation (Belenky and Sevastianov 2007). Equation (5) can therefore be considered as a qualitative representation of a nonlinear dynamical system with softening nonlinearity. The latter is also considered as the simplest qualitative mathematical model of nonlinear roll motions.

Large Roll in PWL Model

As the dynamical system (5) may be considered as a qualitative mathematical model of nonlinear roll motions, consider a distribution of local maxima of the solution (6) in the absence of capsizing and if these maxima exceed the level of ϕ_{m0}.

Note that a resonance phenomenon is not possible in Range 1, as the general solution of autonomous equation does not contain any oscillatory function. As a result:

$$\mathrm{var}(p_0) >> \mathrm{var}(p_1) << \mathrm{var}(p_2) \qquad (7)$$

here var(..) is a variance operator. More details on this argument can be found in Belenky, *et al.* (2016). The eigenvalues for the solution at the Range 2 are:

$$\lambda_{1,2} = -\delta \pm \sqrt{\omega_0^2 k_{f1} + \delta^2} \qquad (8)$$

where k_{f1} is an absolute value of the slope of the stiffness function at Range 1. For the sake of simplicity of further derivation, also assume that damping is absent in Range 1, so

$$\lambda_1 = -\lambda_2 = \omega_0\sqrt{k_{f1}} = \omega_{d1} \qquad (8)$$

Then the solution (6) in Range 1 can be expressed as:

$$\phi = H\cosh(\omega_{d1}t + \varepsilon) + \phi_v \qquad \phi_{m0} < \phi \le \phi_{m1} \tag{10}$$

Arbitrary constants are expressed as:

$$H = -\frac{1}{\omega_{d1}}\sqrt{\omega_{d1}^2(\phi_{m0} - \phi_v)^2 - \dot{\phi}_U^2} \tag{11}$$

$$\varepsilon = \tanh^{-1}\left(\frac{\dot{\phi}_U}{\omega_{d1}(\phi_{m0} - \phi_v)}\right) \tag{12}$$

where $\dot{\phi}_U$ is a roll rate at upcrossing. The value of the peak is expressed as

$$\phi_{\max}(\dot{\phi}_U) = H + \phi_v \qquad 0 < \dot{\phi}_U \le \dot{\phi}_C \tag{13}$$

$\dot{\phi}_{Ci}$ is critical roll rate corresponding to capsizing conditions (see the next subsection).

Within the accepted assumptions, all the quantities in (13) are constant with exception of the roll rate at upcrossing $\dot{\phi}_U$. It is a random variable, as stochastic excitation has been kept for range 0. Assuming that upcrossing are rare, so the general solution of the autonomous equation at Range 0 has enough time to die out between upcrossings, the distribution of the roll rate at upcrossing follows Rayleigh (Leadbetter, *et al.* 1983, p. 201). The distribution needs to be normalized for the condition of the absence of capsizing (if capsizing happens, the ship is not coming back, and there will be no roll peak):

$$pdf(\dot{\phi}_U) = \left(1 - \exp\left(-\frac{\dot{\phi}_C^2}{2\sigma_d^2}\right)\right)^{-1} \frac{\dot{\phi}_U}{\sigma_{\dot{\phi}}^2} \exp\left(-\frac{\dot{\phi}_U^2}{2\sigma_{\dot{\phi}}^2}\right) \tag{14}$$
$$0 < \dot{\phi}_U < \dot{\phi}_C$$

The distribution (14) can be used even if the upcrossings are clustered, see Belenky, *et al.* (2018) for details. Equation (13) is a deterministic function of a single random variable with known distribution and this function is monotonic, thus the distribution the roll peaks / local maxima of the roll angles can be expressed as:

$$pdf(\phi_{\max}) = pdf\left(G^{-1}(\phi_{\max})\right) \cdot \left|\frac{dG^{-1}(\phi_{\max})}{d\phi_{\max}}\right| \tag{15}$$
$$\phi_{m0} < \phi_{\max} < \phi_v$$

where $G^{-1}$is an inverse of the function defined in equation (13):

$$G^{-1}(\phi_{\max}) = \omega_{d1}\sqrt{(\phi_v - \phi_{m0})^2 - (\phi_v - \phi_{\max})^2} \tag{16}$$
$$\phi_{m0} < \phi_{\max} < \phi_v$$

As a result, the distribution density of the roll peaks is expressed in the following form:

$$pdf(\phi_{\max}) = \frac{\phi_v - \phi_{\max}}{C}\exp\left(-\frac{\omega_{d1}^2(\phi_v - \phi_{\max})^2}{2\sigma_{\dot{\phi}}^2}\right) \tag{17}$$

where:

$$C = \frac{\sigma_{\dot{\phi}}^2}{\omega_{d1}^2}\left(\exp\left(-\frac{\omega_{d1}^2(\phi_v - \phi_{m0})^2}{2\sigma_{\dot{\phi}}^2}\right) - 1\right)$$
$$\phi_{m0} < \phi_{\max} < \phi_v$$

As Equation (17) presents the entire distribution of roll peaks for the PWL system, it contains an answer as to how the PWL response differs from the linear response. The distribution of the peaks of a linear response can be approximated by truncated Rayleigh distribution (see, *e.g.* Belenky and Campbell, 2012):

$$pdf(\phi_{\max}) = C_L \frac{\phi_{\max}}{\sigma_x^2}\exp\left(-\frac{\phi_{\max}^2}{2\sigma_x^2}\right)$$
$$C_L = \exp\left(\frac{\phi_{m0}^2}{2\sigma_x^2}\right) \quad ; \quad \phi_{\max} > \phi_{m0} \tag{18}$$

Fig. 6 shows both PDFs. The peaks of the linear response are expected to follow an exponential tail, thus the PWL response peaks do have a heavy tail. The tail remains heavy until some very large value in the vicinity of the angle of vanishing stability. Then the tail turns light making the inflection point. It is caused by the non-capsizing condition: closer to the angle of vanishing stability means more and more trajectories lead to capsizing. As the inflection point is close to the angle of vanishing stability, the tail of PWL peaks is heavy for most of Range 1. This is the sort of information expected from the present analysis.

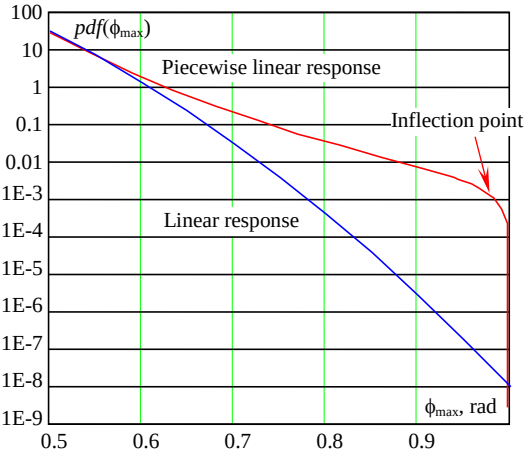

Figure 6: PDFs of peaks of linear response and PWL response (Belenky *et al*., 2016)

Belenky, *et al.* (2018) describes a solution without assuming zero damping in Range 1. The complete PDF cannot be expressed as a closed-form function but its behavior at limit, when the roll angle approaches the angle of vanishing stability, points to a heavy tail.

A completely different approach described in Belenky, *et al.* (2016, 2018) assumes that the excitation is white noise and obtains the PDF from Fokker-Plank-

Kolmogorov (FPK) equation. While realistic excitation is, of course, far from white noise, the shape of the PDF of a nonlinear response seems to be the same for correlated or uncorrelated excitation (Maki, 2016). For the qualitative study described in this section, the assumption of white noise excitation seems to be quite appropriate.

The FPK approach not only confirms that the tail is heavy, but helps to explain why. This is a result of the stretching of the phase plane, caused by softening nonlinearity.

Another way to illustrate this stretching is to compare short portions of the time histories of the piecewise linear and linear systems, starting with some positive roll rate from the angle ϕ_{m0}, see Figure 7. The response of the linear system is described by a trigonometric function, sine or cosine, with appropriate phase shift. The PWL response in the absence of capsizing is described by a hyperbolic cosine, see equation (10). It is easy to see that hyperbolic cosine always stays above the trigonometric cosine for the same initial conditions.

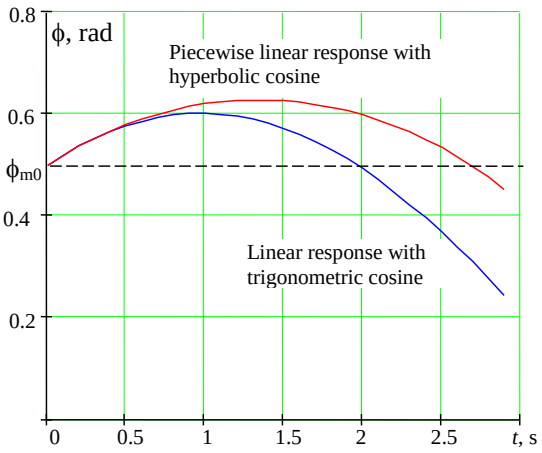

Figure 7: Piecewise linear response above the knuckle point vs. linear response (Belenky, *et al.* 2016)

The PWL system spends more time in Range 1 than the linear system under the same initial conditions. As a result, the probability of finding the PWL system in Range 1 is higher and the tail of the response is heavier than the linear one. As both responses start from the same initial roll rate, the maximum of the PWL response must be larger than the linear one. Thus, the tail of peaks of the PWL response is heavier than the linear one, as can be seen in Figure 6.

Capsizing Metric in PWL Model

Capsizing conditions can be derived from the general solution over Range 1. Whether a capsizing is imminent is defined by the sign of the arbitrary constant A that corresponds to a positive eigenvalue λ_1. Since the exponential term with the positive index is unlimited, it will lead the solution either back to Range 0 if $A<0$ or to Range 2 if $A<0$. The latter is capsizing as the solution in Range 2 will lead to roll around the stable equilibrium at π, i.e. "mast down." There is a sort of "grey area" when A is close to zero and where other terms in the solution (6) may influence the outcome; however this "grey area" is too small to be statistically significant. The capsizing condition can therefore found through $A=0$, leading to the following expression for the critical roll rate of PWL system:

$$\dot{\phi}_C \approx -\lambda_2 \phi_V = \phi_V (\delta + \sqrt{\omega_0^2 k_{f1} + \delta^2}) \qquad (19)$$

The approximately equal sign reflects an assumption that the particular solution in Range 2 is small in comparison with other terms. The metric of likelihood of capsizing is then expressed as:

$$y = 1 + \dot{\phi}_U - \phi_V (\delta + \sqrt{\omega_0^2 k_{f1} + \delta^2}); \qquad (20)$$

The metric of likelihood of capsizing is a deterministic function of a single random argument, the roll rate at the instant of upcrossing. Its distribution following Rayleigh:

$$pdf(\dot{\phi}_U) = \frac{\dot{\phi}_U}{\sigma_{\dot{\phi}}^2} \exp\left(-\frac{\dot{\phi}_U^2}{2\sigma_{\dot{\phi}}^2}\right); \quad \dot{\phi}_U > 0 \qquad (21)$$

The tail of the Rayleigh distribution is exponential. De Haan and Ferreira (2007) contains a proof that a normal distribution that has a tail similar to the Rayleigh tends to exponential tail – see the example 1.1.7 in the cited reference. Thus, the metric of capsizing likelihood is expected to have an exponential tail.

EPOT WITH HEAVY TAILS

Modeling Heavy Tails

Let Y be a positive variable of interest. Its distribution is said to have a heavy (power-law, Pareto) tail if

$$P(Y > y) = C(y) y^{-1/\xi}, \; y > 0, \qquad (22)$$

where $\xi > 0$ is a parameter and $C(y)$ is the so-called "slowly varying" function at infinity. The simplest example of such a slowly varying function is any positive function satisfying $C(y) \sim C$ with a positive constant C as y approaches infinity. Another example would be for $C(y) \sim \log(y)$ as y increases. Following the results from the previous section, the distribution of envelope peaks of roll motion is suggested to be heavy-tailed.

A number of ways to estimate the parameter ξ have been proposed in the literature (*e.g.* Beirlant, *et al.* 2004). If $Y_1, \ldots, Y_n$ is a sample of independent copies of Y, let $Y_{n,n} \leq \cdots \leq Y_{1,n}$ be its order statistic. The most

commonly used estimator of ξ is the so-called Hill estimator defined as

$$\hat{\xi}_k = \frac{1}{k}\sum_{j=1}^{k}\big(\log(Y_{j,n}) - \log(Y_{k+1,n})\big), \qquad (23)$$

where the index k refers to the number of upper order statistics used in estimation. It is known that, under suitable assumptions on large k,

$$\sqrt{k}(\hat{\xi}_k - \xi) \approx \mathcal{N}(0, \xi^2), \qquad (24)$$

where $\mathcal{N}(\mu, \sigma^2)$ refers to a normal distribution with mean μ and variance σ^2, which can be used to set confidence intervals in a standard way.

A number of methods are also available to decide on the index k or threshold $u = Y_{k+1,n}$, above which the distribution is considered to be exactly power-law or Pareto, that is $P(Y > y) = Cy^{-1/\xi}$, $y > u$. One of these methods is based on the so-called prediction error criterion (Dupuis and Victoria-Feser, 2006, and Mager, 2015). Let $X_{i,n} = \log(Y_{i,n}) - \log(u)\,, i = 1, \dots, k$. The prediction criterion concerns the so-called mean squared prediction error defined as

$$\Gamma(k) = \frac{1}{k}\sum_{i=1}^{k} E\left(\frac{\hat{X}_{i,k} - EX_{i,k}}{\sigma_i}\right)^2. \qquad (25)$$

Here, ${\sigma_i}^2 = \mathrm{var}(X_{i,k})$ and $\hat{X}_{i,k}$ is the estimated value of $X_{i,k}$ according to the model of interest (here, the exponential; see below). By expressing $\Gamma(k)$ as

$$\frac{1}{k}\sum_{i=1}^{k} E\left(\frac{X_{i,k} - X_{i,k}}{\sigma_i}\right)^2 + \frac{2}{k}\sum_{i=1}^{k}\frac{Cov(\hat{X}_{i,k}, X_{i,k})}{{\sigma_i}^2} - 1 \qquad (26)$$

(e.g. Mager (2015), Lemma 4.2), it is then estimated as

$$\hat{\Gamma}(k) = \frac{1}{k}\sum_{i=1}^{k}\left(\frac{\hat{X}_{i,k} - X_{i,k}}{\hat{\sigma}_i^2}\right)^2 + \frac{2}{k}\sum_{i=1}^{k}\frac{\widehat{Cov}(\hat{X}_{i,k}, X_{i,k})}{\hat{\sigma}_i^2} - 1\,. \qquad (27)$$

where hats indicate the estimators. The index $\hat{k}$ for the threshold selection is chosen as that minimizing $\hat{\Gamma}(k)$ over some range of values k. Mager (2015) suggests using the range $[\min(40, 0.02n)\,, 0.2n]$.

To conclude the description of the method, one still needs to specify the various quantities in the definition of $\hat{\Gamma}(k)$. Under the above Pareto model for $Y > u$, $\log(Y) - \log(u)$ follows an exponential distribution with mean ξ, and $X_{i,n}, i = 1, \dots, k$, are its order statistics. Under the exponential model, the distributions of the order statistics $X_{i,k}$ are well-known (e.g. Embrechts, *et al.* (2013), Example 4.1.5). In particular, with details omitted here, one can show that

$${\sigma_i}^2 = \xi^2 \sum_{j=i}^{k}\frac{1}{j^2}, \qquad (28)$$

where ξ is the mean parameter of the above exponential distribution. The estimator $\hat{X}_{i,k}$ of $X_{i,k}$ is defined as follows. The quantity $X_{i,k}$ can be thought as the $(1 - i/(k+1))th$ quantile of the exponential distribution. But for this distribution, this quantile is $-\xi\log(i/(k+1))$, which suggests using

$$\hat{X}_{i,k} = -\hat{\xi}_k \log\left(\frac{i}{k+1}\right). \qquad (29)$$

One can also show (with details omitted) that $\mathrm{cov}(\hat{X}_{i,k}, X_{i,k}) \approx \frac{\xi^2}{k}\left(\log(\frac{i}{k+1})\right)^2$, suggesting use of

$$\widehat{\mathrm{cov}}(\hat{X}_{i,k}, X_{i,k}) = \frac{{\hat{\xi}_k}^2}{k}\left(\log(\frac{i}{k+1})\right)^2. \qquad (30)$$

Substituting these quantities into the definition of $\hat{\Gamma}(k)$ leads to the expression

$$\hat{\Gamma}(k) = {\hat{\xi}_k}^{-2}\sum_{i=1}^{k}\left(\sum_{j=i}^{k}\frac{1}{j^2}\right)^{-1}\left(X_{i,k} + \hat{\xi}_k\log(\frac{i}{k+1})\right)^2 + \frac{2}{k}\sum_{i=1}^{k}\left(\sum_{j=i}^{k}\frac{1}{j^2}\right)^{-1}\left(\log(\frac{i}{k+1})\right)^2 - 1. \qquad (31)$$

In estimating the exceedance probability $P(Y > y^*)$ for a heavy-tailed distribution and some target value y^*, one writes

$$P(Y > y^*) = P(Y > u)\frac{P(Y > y^*)}{P(Y > u)}. \qquad (32)$$

Here, u is a threshold above which Y follows a Pareto model and which can be selected through the prediction error criterion described above. Then,

$$\frac{P(Y > y^*)}{P(Y > u)} = \frac{C(y^*){y^*}^{-\frac{1}{\xi}}}{C(u)u^{-\frac{1}{\xi}}} \approx \left(\frac{y^*}{u}\right)^{-\frac{1}{\xi}} \qquad (33)$$

and hence

$$P(Y > y^*) = P(Y > u)\left(\frac{y^*}{u}\right)^{-\frac{1}{\xi}}. \qquad (34)$$

In practice, the probability $P(Y > u)$ is estimated as the sample proportion of data above the threshold u and ξ is taken as the Hill estimator associated with that threshold. Confidence intervals for the probability $P(Y > u)$ are set based on the standard binomial calculations, and those for ξ are set based on equation (23).

Testing Heavy Tails

A statistical validation of heavy tails was carried out using the fast volume-based 3 DOF calculations described in Weems, *et al.* (2018). The fast code was used to create a very large sample of data where large roll angles are observable. These observations are used to estimate a "true value" with direct counting. A series of small subsets of this large sample was then used to generate extrapolated estimates to be compared with the "true value." Following the three-tiered methodology of Smith and Zuzick (2015), the tiers are defined as:

- All extrapolations for a single target value;
- All available target values;
- All available environmental conditions.

This separation by tiers depends on the context. When the extrapolation is carried out for large roll angles, multiple target values can be and should be analyzed. However, when capsizing is considered, there is only one target.

The validation value data was computed for the ONR tumblehome configuration (Bishop, *et al*. 2005). All of the calculations were carried for random realizations of a Bretschneider spectrum with a significant wave height of 9 m and modal period of 15 seconds. Speed was set to 6 knots, and KG to 7.5 m, resulting in GM=2.2 m. Other data for the true value calculations are summarized in Table 1.

Table 1. "True value" calculations

Headings Deg.	Total time, hrs	Number of targets	Largest target	Number of exceedances of largest target
15	570,000	5	20	14
22.5	200,000	7	27.5	16
30	200,000	13	45	9
37.5	200,000	15	60	7
45	690,000	15	70	8
60	600,000	15	70	12
90	690,000	9	37.5	12
135	690,000	3	20	6

The tier 1 validation is a set of comparisons of extrapolated estimates with the true value. An example is shown in Fig. 8 for a 45 degree heading (stern quartering seas) and a 45 degrees roll target value. There are 50 extrapolation estimates, each computed from 100 hours of data. The main index of performance is the Passing Rate, indicated the percentage of successful extrapolations. An extrapolation is considered successful if it has an overlay with the confidence interval of the "true value."

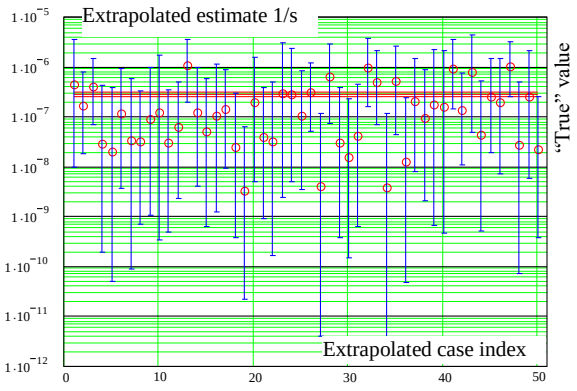

Figure 8: Tier 1 extrapolation validation for a heading of 45 degrees and target value of 45 degrees

Two more performance indicators are considered: conservative distance CD and relative bias RB, defined as:

$$CD = \log_{10}\frac{E(\hat{\lambda}_{Up})}{\hat{\lambda}_T}; \quad RB = \frac{E(\hat{\lambda}_E) - \hat{\lambda}_T}{\hat{\lambda}_T} \qquad (35)$$

Where, $E(\hat{\lambda}_{Up})$ is the upper boundary of the extrapolated estimates averaged over the considered extrapolation data sets, $E(\hat{\lambda}_E)$ is the most probable extrapolated estimate averaged over the considered datasets, and $\hat{\lambda}_T$ is the "true value." Numerical values for the performance indexes are summarized in Table 2.

Table 2. Extrapolation performance example

Heading, deg.	45
Roll angle target value, deg.	45
Passing rate	0.9
True value, 1/s	2.718E-7
Averaged extrapolated estimate $E(\hat{\lambda}_E)$, 1/s	2.247E-7
Relative bias, *RB*	-0.174
Conservative distance, *CD*	0.583
Average time between exceedances, hours	1,022

The extrapolation performance looks quite satisfactory for this combination of heading and target roll angle. The passing rate seems to be reasonable. The averaged extrapolated estimate $E(\hat{\lambda}_E)$ is a measure of validity of the statistics model. It shows a very slight underestimation, indicated by the negative RB value, but the difference between $E(\hat{\lambda}_E)$ and the "true value" is practically non-existent. Overall, the results presented in Fig.8 and Table 2 do not reject the heavy tailed model for roll peaks when considering this target roll angle using simulations from this ship heading.

The second tier of statistical validation considers all available targets. The passing rates are shown in Fig. 9 while conservative distance and relative bias are presented in Fig. 10. An acceptable passing rate for 50 extrapolation data sets is from 0.88 to 1 (Smith, 2018). Such variation of the passing rate can be explained by the natural variability of the statistical estimates. The extrapolations are acceptable for all targets, excluding 50 and 60 degrees, where the passing rates fells to 0.86. The performance for other roll target value is similar to the observed for the 45 degrees target.

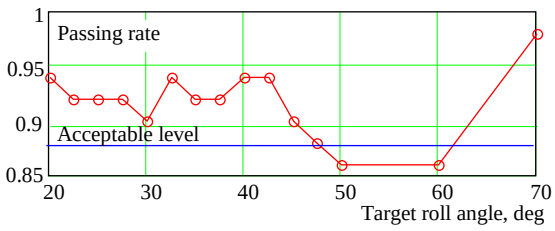

Figure 9: Passing rate for heading of 45 degrees

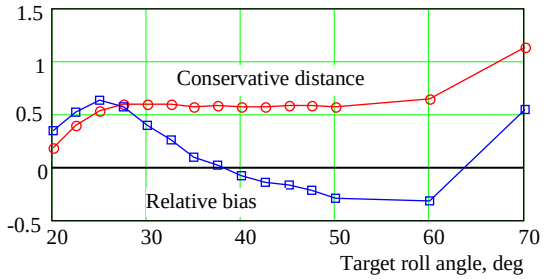

Figure 10: Conservative distance and relative bias for heading of 45 degrees

The third tier assesses the performance over all available conditions. The passing rates are shown in Fig. 11. Two lines are shown. One corresponds to an averaged passing rate over all target values, another the smallest passing rate value encountered among all the target values; for a 45 degree heading it corresponds to a minimum shown in Fig. 9. Obviously, the extrapolation did not work for the heading 135 degrees.

The values of the conservative distance and relative bias, averaged over all target roll angles, are shown in Fig. 12 and 13, respectively. With the exception of the 15 and 135 degree headings, the conservative distance does not exceed one order of magnitude. The relative bias also remains quite small with the exception of the 15 degree heading (2.28) and 135 heading (39.6). The latter number explains the very low passing rates observed for the 135 degree heading; there is a significant overestimation of the probability of exceedance, see Fig 14. Perhaps, the tail is too heavy.

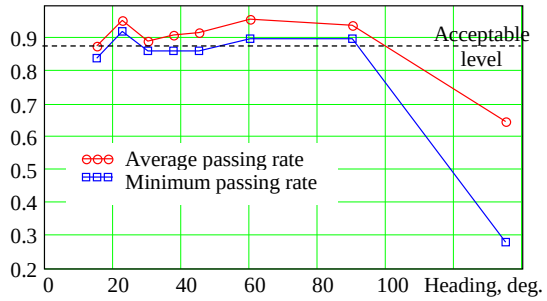

Figure 11: Passing rate for all headings

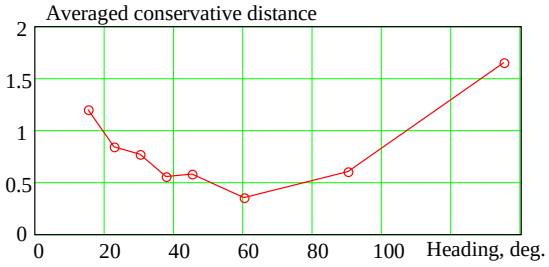

Figure 12: Averaged conservative distance for all headings

Softening nonlinearity was shown to be the reason for peaks of the response to have a heavy tail. Thus, if the tail is not heavy, maybe motions are simply too small for the nonlinearity to have a significant impact. Fig. 15 shows percentiles and largest observed values computed for each extrapolation set and then averaged over all data sets for the same heading angle.

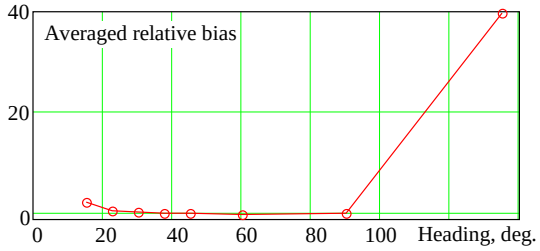

Figure 13: Averaged relative bias distance for all headings

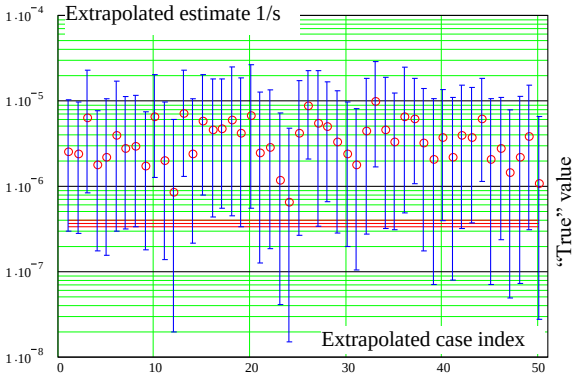

Figure 14: Extrapolations for heading 135 degrees and target value 17.5 degrees

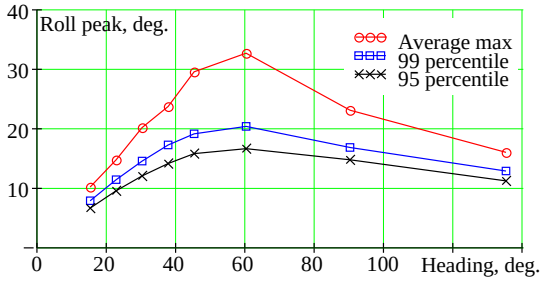

Figure 15: Percentiles and maxima of roll motions

All the curves in Fig. 15 have a maximum at the 60 degree heading, while the excitation is the largest in beam seas. There may be two reasons why the largest motions are observed at the 60 degree heading: synchronous resonance and stability variation. While the latter reason seems to be more plausible as ONR tumblehome configuration is known for her propensity for stability variation, resonance conditions also need to be considered. Fig. 16 shows encounter frequency corresponding the modal period of waves (15 s) for all the considered headings. As one can see from Fig 16, the 60 degree heading is further from synchronous resonance than beam seas, so the increase percentiles around a heading of 60 degrees in Fig. 15 should be attributed to stability variation.

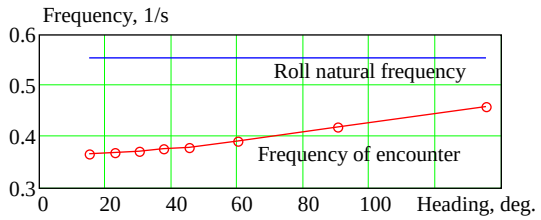

Figure 16: Encounter frequency vs. heading

The influence of the stability variation waves is a known factor for stern quartering and following seas. While transverse stability still changes for other heading as well, it generally happens too quickly for a ship to react on decreased roll stiffness. Stability variations can also invoke parametric roll, which can be significant for ships with certain hull forms and very low roll damping, but the present hull is not particularly susceptible to it. The stability still can therefore be represented with the calm water GZ curve that is close to the averaged GZ curve in waves.

Fig 17 shows the percentiles for 90 (beam) and 135 (bow) headings from Fig 15 placed on the GZ curve normalized by the GM value. As expected, the percentiles for the 90 degree heading are closer to the maximum of the GZ curve. However, the difference between the two headings does not look as dramatic as one could expect. Nevertheless, this difference is sufficient to make the tail heavy. Can this difference be quantified?

Fig. 18 shows values of the derivative of the normalized GZ curve computed at the angles corresponding to averaged maxima and 99 and 95 percentiles for both headings. The following criteria can be deduced from the observation: the normalized derivative should be below 0.6 at the 95, below 0.5 at the 99 percentile, and below 0.4 and at averaged maximum. Unfortunately, at the present time, there are no appropriate data to test these criteria, so they remain in a hypothetical state.

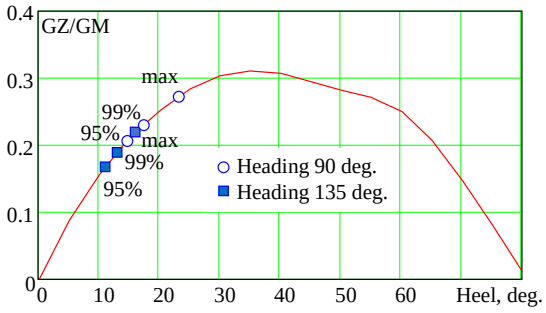

Figure 17: Percentiles shown on normalized GZ curve

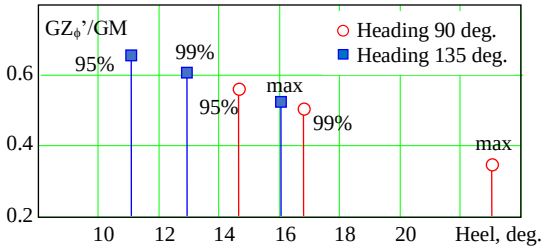

Figure 18: Percentiles and derivatives GZ curve

Data for all other headings cannot be used for this type of analysis as the influence of stability variation is strong, so the calm water GZ curve cannot be used to indicate nonlinearity. Consider the 15 degree heading: the percentiles are smaller than those for the 135 degree heading, see Fig. 15. At the same time, the passing rate is much better; the minimum passing rate is 0.84, which is below the acceptable level of 0.88, but average passing is actually 0.88. That means that there are some target values for which extrapolation was perfect (i.e. deviation of the passing rate from 0.95 can be explained by random reasons). Fig. 19 shows the passing rates for all targets, available for the 15 degree heading. Fig. 20 shows conservative distance and relative bias, while Fig. 21 depicts validation of extrapolation for the target of 17.5 degrees.

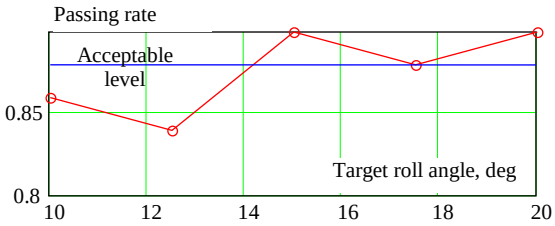

Figure 19: Passing rate for heading of 15 degrees

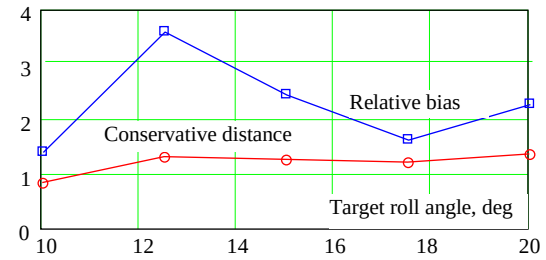

Figure 20: Conservative distance and relative bias for heading of 15 degrees

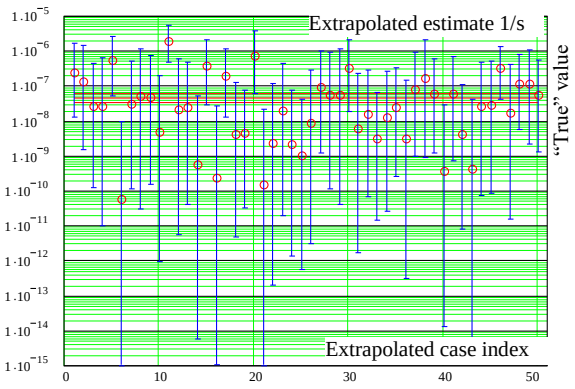

Figure 21: Extrapolations for heading 15 degrees and target value 17.5 degrees

While the extrapolation performance at the 15 degree heading is worse than at the 45 degree heading, it still works and recovers the correct type of tail despite the fact that the observed motions are rather small. It can be explained that large excursions of roll are caused by decreased stability, so despite the small roll angles the influence of the softening nonlinearity remains strong enough to make the tail heavy.

Overall, the performance of the extrapolation for heavy tails, except for the 135 degree heading, can be characterized as "almost there." The passing rate falls short of the required 0.88 for a few cases, but not by much. The conservative distance is well within 1 order of magnitude, excepting the 15 degree heading, where it is slightly higher. Relative bias is largest for the 15 degree heading, which probably explains the larger conservative distance for that heading. Still, the decrease of uncertainty in comparison with two-parameter GPD is substantial and brings EPOT closer to practical application.

TAIL STRUCTURE OF CAPSIZING METRIC

Including physical information into statistical model of large roll angles, as described in the previous section, did help to reduce their uncertainty. While some issues remain unresolved, it is clear that this provides a way forward. Now consider adding physical information into the model of the tail of the distribution of the capsizing metric (1).

The model of the tail of the distribution for the capsizing metric and its testing is considered in detail in Belenky, *et al.* (2018a). The present paper contains only a very brief description of that model, but includes some additional material that was not included in the cited paper.

Numerical Study of Shape Parameter

An early attempt to apply the exponential distribution to the capsizing metric computed with the volume-based method (Belenky, *et al.* 2013) led to a significant overestimation of the actual probability of capsizing. Application of the two-parameter GPD led to successful validation (Weems, *et al.* 2016), but the question remains: is the true tail of the capsizing metric actually exponential? One way to check this is to perform massive simulations of a data set containing actual capsizing, fit the two-parameter GPD and observe its behavior as it approaches unity, which indicates capsizing.

The simulations were performed for a significant wave height of 8 m and modal period of 15 seconds. The KG value was adjusted to make capsizing observable and a large-volume sample was taken to compute the critical roll rate and the capsizing metric. The latter was used to fit GPD and the evaluation of the shape parameter estimate is monitored for different heading and KG values.

Fig. 22 shows the evolution of the shape parameter (with its confidence interval) with the increase of the threshold for beam seas. The volume of the sample was 875 hours (1750 records, 30 minutes each) and contained 3 capsizes, see Table 3. The estimates of the shape parameters start from negative values and show a clear tendency to increase, reaching zero around the threshold value of 0.88 rad/s. The zero value is contained within the confidence interval for the rest of the threshold with exception of u =0.91, 0.914 and 0.915 rad/s. This picture appears to be consistent with the hypothesis of an exponential tail of the capsizing metric.

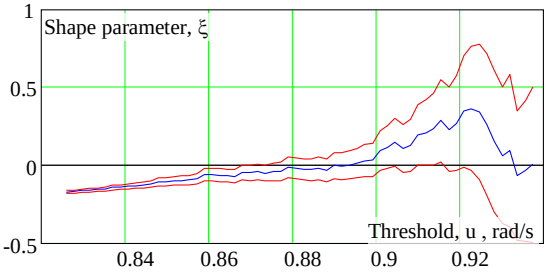

Figure 22: Evolution of shape parameter. ONRTH, beam seas, KG=8.35 m

If one tries to fit GPD to data with known normal or Rayleigh distribution, *i.e.* where the tail is known to be exponential, the general picture will be very similar to the one in Fig. 22. Both normal and Rayleigh distribution contains the square of a variable, taken with a minus sign. The minus square decreases faster than just a line with negative slope. This circumstance can be seen in the GPD as a negative shape parameter for low thresholds. The shape parameter estimate is then expected to stabilize around the zero value. Fig. 22 can be interpreted along these lines.

To see the effect of stability variation, these calculations were repeated for the 45 degree heading. The sample size was increased to 5,000 hours (still a set of records 30 minutes each). The sample contains 206 cases of capsizing. Fig. 23 shows the result, and it is

dramatically different from the one seen in the previous case. The estimate starts near zero and shows two minima, around thresholds 0.65 and 1.01 rad/s.

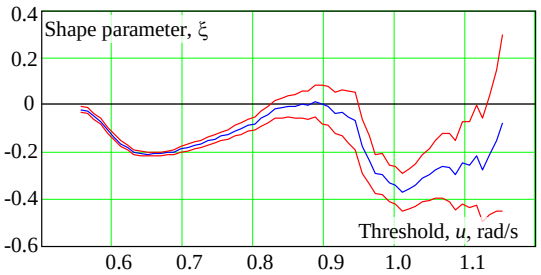

Figure 23: Evolution of shape parameter. ONRTH, heading 45 degrees, KG=7.85 m

To find how stable this behavior is, the calculations were repeated for the 50 degree heading using an even larger sample – 16,000 hours where 1093 capsize cases were observed, see Fig. 24 and Table 3. The behavior of the shape parameter estimate did not change, but the confidence interval has shrunk as the volume of sample has increased.

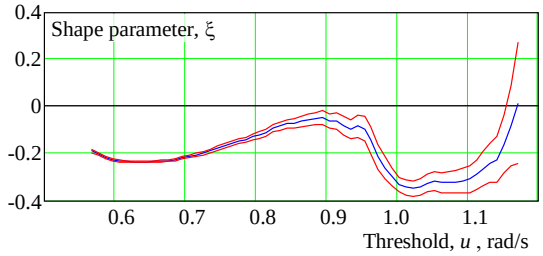

Figure 24: Evolution of shape parameter. ONRTH, heading 50 degrees, KG=7.85 m

Further increase of heading is expected to bring the tail to exponential, as the influence of the stability variation gets weaker. Fig. 25 shows an estimate of the shape parameter's evolution for the 70 degree heading. As in the previous case, the volume of sample was large – 16,000 hours with 1003 capsizing case observed (see Table 3). The behavior of the estimate is very similar to the beam seas result shown in Fig.22, only the stabilization of the shape parameter estimate occurs for a larger threshold value, μ=1.0 rad/s, than the *u*=0.88 rad/s seen in the beam seas case.

Thus, the "double minima" topology of the shape parameter estimated observed for 45 and 50 degree headings can be attributed to stability variation. Can this preliminary conclusion be applicable for other hull forms?

Fig. 26 shows a plot for the shape parameter estimated for ONR flared hull (ONRFL) sailing with 6 knots and a heading of 45 degrees. This hull has the same shape as the ONR tumblehome hull (ONRTH) below the design waterline, but a flared topside similar to conventional destroyers. The sample was also large: 16,000 hours with 53 observed capsizing cases (see Table 3). The behavior of the shape parameter estimate is consistent with the hypothesis of exponential tail. Thus, the "double minima" topology is a result of specific features of stability variation with the ONR tumblehome hull and not necessarily applicable for other hull configurations.

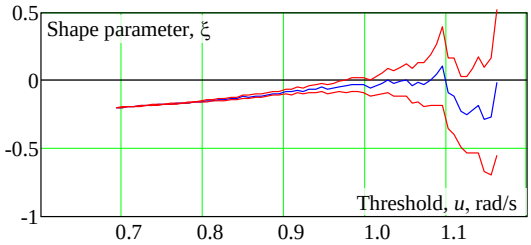

Figure 25: Evolution of shape parameter. ONRTH, heading 70 degrees, KG=8.00 m

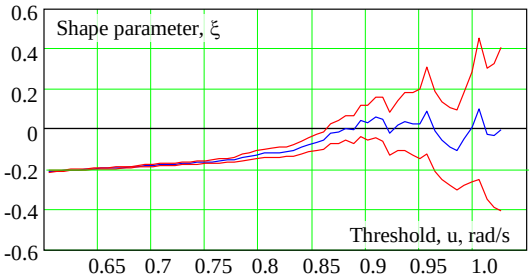

Figure 26: Evolution of shape parameter. ONRFL, heading 45 degrees, KG=8.8 m

Table 3 Conditions for Study of Shape Parameter

Ship	Heading deg.	KG, m	GM	Numb. caps.	Time, hrs.
TH	45	7.85	1.85	206	5,000
TH	50	7.85	1.85	1093	16,000
TH	70	8.00	1.71	1003	16,000
TH	90	8.35	1.36	3	875
FL	45	8.80	0.33	53	16,000

Fitting Exponential Tail

The numerical study partially confirmed the hypothesis that the capsizing metric may have an exponential tail. Leaving the effect of stability variation for future study, consider a more robust method of fitting an exponential tail for the capsizing metric. Below is an abridged description of this study, for which full details can be found in Belenky, *et al.* (2018a).

To fit the exponential tail, a threshold has to be found. If the threshold has been found correctly, the distribution above it should be exponential, *i.e.* passing the goodness-of-fit test. Kolmogorov-Smirnov (KS) is one of those tests that judges the goodness of fit by the largest distances between the statistical and theoretical Cumulative Distribution Functions (CDF). The problem is that the KS test works well when the theoretical distribution is completely defined and all the parameters are known; however the exponential distribution parameter here is not known and needs to be estimated.

The necessary modification to KS test for exponential distribution can be found in Stephens (1974).

The null hypothesis H_0 is that the capsizing metric above the candidate threshold y_i-u follows exponential distribution. The test statistic for the null hypothesis is defined as

$$D^* = \left(D - \frac{0.2}{n}\right)\left(\sqrt{n} + 0.26 + \frac{0.5}{\sqrt{n}}\right) \quad (36)$$

where n number of data points above the threshold u and

$$D = \max\{D^+, D^-\} \quad (37)$$

The values D^+ and D^- are defined as follows:

$$D^+ = \max_{j=1,\ldots,n} \left|\frac{j}{n} - z_j\right| \quad (38)$$

$$D^- = \max_{j=1,\ldots,n} \left|z_j - \frac{j-1}{n}\right| \quad (39)$$

The values z_j are essentially the exponential CDF:

$$z_j = 1 - \exp\left(\hat{\gamma}(\text{sort}(y_j) - u)\right) \quad (40)$$

where $\hat{\gamma}$ is an estimate of exponential parameter:

$$\hat{\gamma} = n\left(\sum_{j=1}^{n}(y_j - u)\right)^{-1} \quad (41)$$

The critical values for D^* are available from Stephens (1974) and are listed in Table 4 for several values of the significance level.

Table 4 Critical value of D^*

Sign. level	0.01	0.025	0.05	0.10	0.15	0.20
D^*	1.308	1.190	1.094	0.990	0.926	0.880
Sign. level	0.25	0.30	0.35	0.40	0.45	0.50
D^*	0.835	0.795	0.766	0.736	0.710	0.685

An application of this goodness-of-fit test is somewhat similar to the KS test. The value of D^* computed for a given threshold should be less than the critical value from Table 4 corresponding to the chosen significance level. The fitting procedure is essentially finding the largest value of the threshold that provides a given level of significance. Selection of the level of significance and other details, as well as other method of fitting exponential tail are discussed in Belenky, *et al.* (2018a).

Testing Exponential Tail

Testing of the exponential tail fit for the capsizing metric was performed using the data set from Weems, *et al.* (2017) for the ONR tumblehome configuration sailing with a speed of 6 knots in long-crested seas with the significant wave height of 9 m and modal period of 14 seconds. The heading angle varied from 35 to 70 degrees.

The 45 and 50 degree headings present the most interesting test cases, recalling the complex behavior of shape parameter estimates observed in Fig. 23 and 24. Attempts to validate the extrapolation of the capsizing metric with an exponential tail fit for these heading brought surprisingly high passing rate, see Fig. 27 and 28.

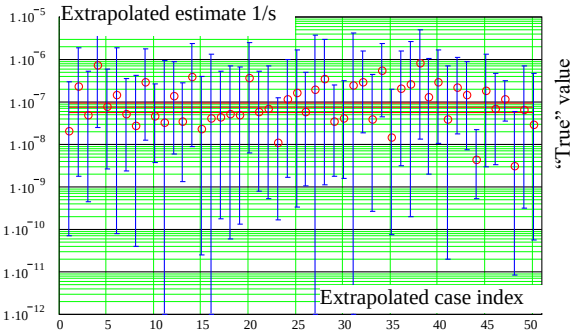

Figure 27: Validation of capsizing metric for the 45 degree heading; passing rate is 0.98, *CD*=1.2, *RB*=1.1

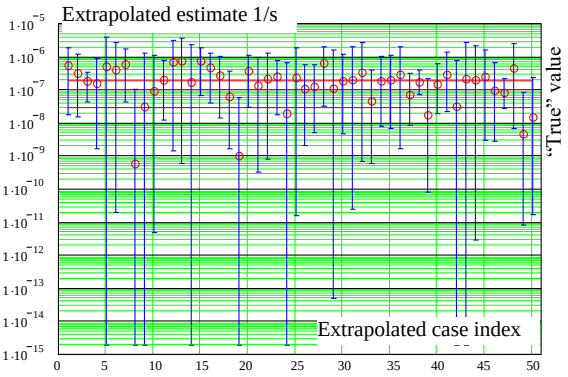

Figure 28: Validation of capsizing metric for the 50 degree heading; passing rate is 1.0, *CD*=1.27, *RB*=1.63

The exponential fit was done with the goodness-of-fit method while the significance level was taken equal to 0.2. The other performance indicators are in the acceptable range, though they could be better. The practical applicability of the exponential fit for the capsizing metric for ONR tumblehome sailing in stern quartering seas is good news indeed. The behavior of shape parameter remains a puzzle however.

The passing rates for all of the headings is plotted in Fig. 29. All the values are within the accepted range. The performance indexes, conservative distance and relative bias, while not perfect, seems to be acceptable. The conservative distance is just above unity and relative bias is positive, but not very large. Belenky, *et al.* (2018a) contains results computed for other significance level as well as results computed with an alternative method of fitting exponential tail using the minimization of the prediction error.

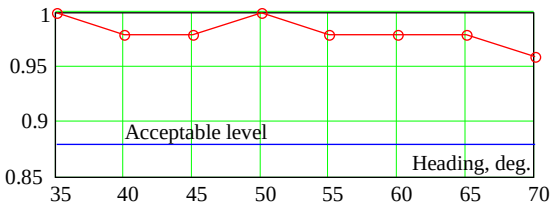

Figure 29: Passing rate for capsizing metric

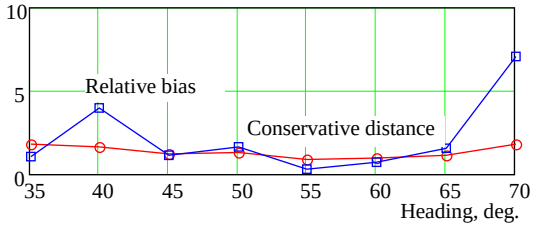

Figure 30: Conservative distance and relative bias for capsizing metric

Relation between Metric and GZ Curve

The shape parameter plots in Figs. 23 and 24 differ from those in Figs. 22, 25 and 26 in two noticeable features. First, the shape estimates in Figs. 23 and 24 decrease initially, before climbing towards zero. Second, these estimates then decrease again around a metric value of 1, before climbing up again for the very largest values of the metric. The following discussion puts forth a possible explanation for the first difference, but any underlying phenomena for the second difference remain to be clarified.

On one hand, the specified different features in the shape parameter plots may simply result from the different characters of the histograms of the metric values underlying Figs. 22-26. For example, Fig. 31 presents the histogram of the values for ONRTH behind Fig. 23, with an inset that zooms into the tail of the histogram around the metric value 1. Fig. 32 provides a histogram of the metric values for ONRFL used in Fig. 26.

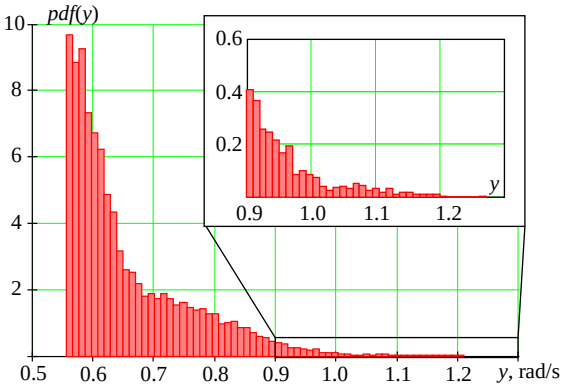

Figure 31: Histogram of the metric values. ONRTH, heading 45 degrees, KG=7.85 m

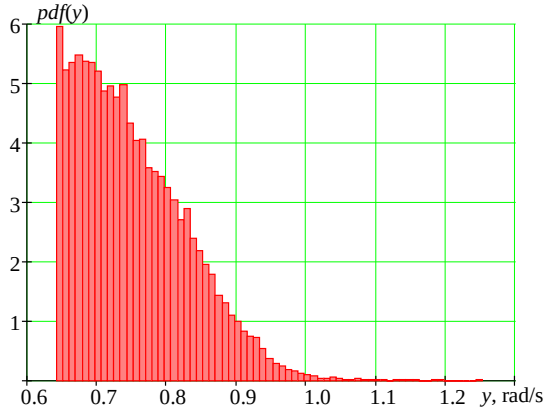

Figure 32: Histogram of the metric values. ONRFL, heading 45 degrees, KG=8.8 m

Note from Fig. 31 that the histogram for ONRTH has certainly a bimodal character (with the second mode around the value 0.7) and perhaps even being trimodal (with the third mode in the zoomed histogram around the value 1.05), which is not the case with the histogram for ONRFL. This multimodal feature also explains the character of the shape parameter plot in Fig. 23. Because of the second mode, the GPD sees a heavier tail for smaller thresholds and thus yields larger shape parameter estimates. When the thresholds are around the second mode (the value 0.7 of the metric), this apparent heaviness of the tail disappears and shape parameter estimates decrease. They start increasing again because of the third mode, before again dropping around the third mode. The very tail of the histogram is consistent with the zero GPD shape parameter.

If multimodality of the histogram is responsible for the character of the shape parameter plot, what might be causing it? The second pronounced mode is likely caused by the bimodality of the restoring force for ONRTH in certain headings. Similarly, the absence of such a second mode for ONRFL may be caused by the lack of bimodal features in the restoring force. Figs. 33 and 34 present the histograms of the angle of maxima of the instantaneous GZ curves for ONRTH and ONRFL, and the bimodality is indeed evident in the former but not in the latter case. The algorithm for calculating the instantaneous GZ curve is described in Belenky and Weems (2008).

Similarly, Fig. 35 presents GZ curves for ONRTH at several positions on a regular wave of length 150 m and height 7.5 m, with the bimodality now clearly seen in the form of the two clusters of the curves. The bimodality in the GZ curves can be thought of as a result of the wave-piercing geometry of the ONRTH bow.

Though not a formal proof, restoring forces without bimodal features thus seem to lead to a unimodal histogram of the resulting metric values. Restoring forces with bimodal features can then be expected to lead to a bimodal histogram, as described above, since the

restoring forces from the two clusters (or modes) are naturally expected to yield different average values and hence two modes of the resulting metric values. As noted in the beginning of the section, any phenomena underlying the apparent third mode in the zoomed histogram in Fig. 31 remain to be clarified.

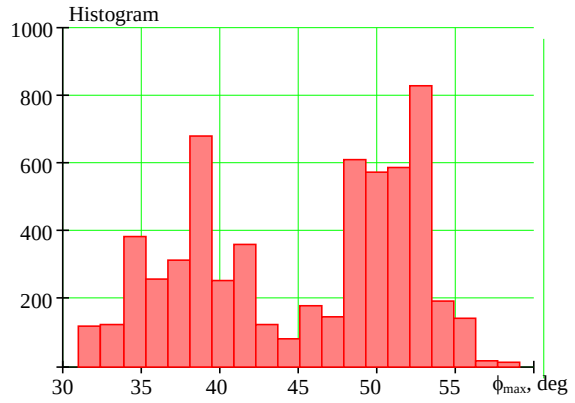

Figure 33: Histogram of the maxima positions of the instantaneous GZ curves. ONRTH, heading 45 degrees, KG=7.85 m

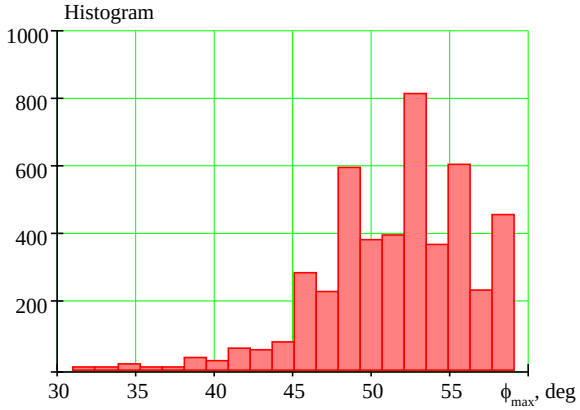

Figure 34: Histogram of the maxima positions of the instantaneous GZ curves. ONRFL, heading 45 degrees, KG=8.8 m

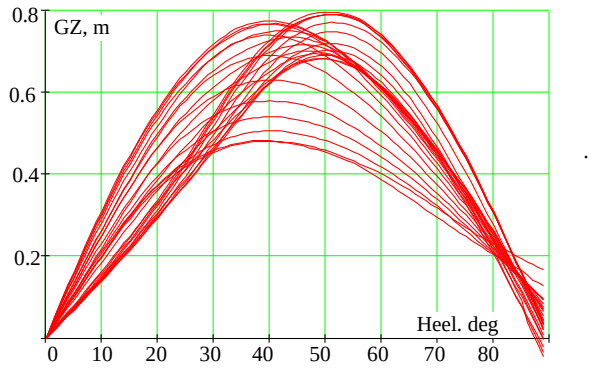

Figure 35: GZ curves for several positions of ONRTH on a regular wave of length 150 m and height 7.5 m

SUMMARY, CONCLUSIONS, AND FUTURE WORK

The objective of the work presented in this paper was seeing how to insert physics into a statistical model of extreme values in order to reduce uncertainty and improve the prediction's reliability. The second extreme value theorem states that Generalized Pareto Distribution can approximate a tail of any distribution above a high-enough threshold. The GPD model is completely data-driven, and its statistical uncertainty reflects the volume of the available sample. For rare events and extreme values, the statistical uncertainty of the prediction may be large. To reduce the uncertainty without having to increase the volume of sample, one can try adding physical information into the statistical model. Of course, such a hybrid model strongly depends on context of the physical problem. This paper describes how physical information can be added to extreme-value models of large roll angles and capsizing metrics.

A piecewise linear dynamical system with triangle-shaped stiffness represents most properties of a dynamical systems with softening nonlinearity, while allowing for a closed-form solution to the distribution of peaks of roll motions. One can easily see that the tail of that distribution is heavy. Thus, a Pareto distribution with heavy tail can be used instead of the more generic GPD. The uncertainty is decreased because the model has only one parameter to fit. The heavy tail was tested for a large sample, containing extreme values, from which extreme-value predictions were made on small subsets and compared with actual observations. The heavy tail was found to work well where there was enough influence of softening nonlinearity, but did not work for small-amplitude motion cases, where the motion was more linear and the assumptions about the physical model were no longer applicable.

The piecewise linear model was also used to examine the tail of the capsizing metric computed in the split-time method. Since this metric is largely a function of the roll rate at the instant of upcrossing, and the roll rate is known to be "a weak nonlinearity," the metric is expected to have an exponential tail. The exponential character was tested in a similar manner – the probability of capsizing events observed in a large-volume sample was compared with predictions computed from small-volume samples. The exponential approximation worked well, despite a very complex behavior of the tail of the capsizing metric for the ONRTH configuration in stern quartering seas. As with the heavy tail assumption for the roll peaks, the exponential assumption of the capsizing metric led to a decrease of uncertainty manifested through the transition from the two-parameter GPD to a one-parameter exponential tail.

Finally, an interesting relationship was observed between the complex tail behavior and multi-

modality of distributions of capsizing metrics and the angles of maxima of the instantaneous GZ curves in waves. This relationship presents its own interest for further study as a link between certain types of hull forms and their dynamical behavior in extreme seas.

ACKNOWLEDGEMENTS

The work described in this paper has been funded by the Office of Naval Research (ONR) under Dr. Woei-Min Lin and by the NSWCCD Independent Applied Research (IAR) program under Dr. Jack Price. The participation of Prof. Sapsis was facilitated by the NSWCCD Summer Faculty Program, while the participation of Prof. Pipiras was facilitated by NSWCCD Summer Faculty and Sabbatical Programs, both of which are also managed by Dr. Jack Price.

The authors are very grateful for the support that made this work possible.

REFERENCES

Anastopoulos, P.A., Spyrou, K.J., Bassler, C.C. and V. Belenky "Towards an Improved Critical Wave Groups Method for the Probabilistic Assessment of Large Ship Motions in Irregular Sea," Probabilistic Engineering Mechanics Vol. 44, 2015, pp 18-27.

Beirlant, J., Goegebeur, Y., Teugels, J. and J. Segers, Statistics of Extremes, Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 2004. Theory and applications, with contributions from Daniel De Waal and Chris Ferro.

Belenky, V. L., "Piecewise linear approach to nonlinear ship dynamics," Contemporary Ideas on Ship Stability, Vassalos, D., Hamamoto, M., Papanikolaou, A. and D. Molyneux, eds., Elsevier, 2000, ISBN 0-08-043652-8, pp.149-160.

Belenky, V.L. and N.B. Sevastianov, Stability and Safety of Ships: Risk of Capsizing, Second edition SNAME, Jersey City, ISBN 0-939773-61-9, 2007.

Belenky, V. and K.M. Weems, "Probabilistic Qualities of Stability Change in Waves," Proc. 10th Inlt. Ship Stability Workshop, 2008, Daejon, Korea, pp. 95-108.

Belenky, V. and Campbell, B., "Statistical Extrapolation for Direct Stability Assessment," Proc. 11th Intl. Conf. on Stability of Ships and Ocean Vehicles (STAB 2012), 2012, Athens, Greece, pp. 243-256

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

Belenky, V., Weems, K., and Pipiras, V., "Split-time Method for Calculation of Probability of Capsizing Due to Pure Loss of Stability," Proceedings of the 13th International Ship Stability Workshop, 2013, Brest, France, pp. 70-78.

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

Belenky, V., Glotzer, D., Pipiras, V. and T. Sapsis, "On the Tail of Nonlinear Roll Motions," Proc. 15th Intl. Ship Stability Workshop, 2016, Stockholm, Sweden, pp. 109-114.

Belenky, V., Glotzer, D., Pipiras, V. and T. Sapsis, "Distribution tail structure and extreme value analysis of constrained piecewise linear oscillators," 2018 (Preprint).

Belenky, V. Weems, K., Pipiras, V., "Extreme-Value Properties of the Split-Time Metric," Proceedings of 13th International Conference on the Stability of Ships and Ocean Vehicles (STAB 2018), 2018, Kobe, Japan. (Submitted).

Bishop, R. C., W. Belknap, C. Turner, B. Simon, and J. H. Kim, Parametric Investigation on the Influence of GM, Roll damping, and Above-Water Form on the Roll Response of Model 5613, Report NSWCCD-50-TR-2005/027, 2005.

Boccotti, P., Wave Mechanics and Wave Loads on Marine Structures, Butterworth-Heinemann, Elsevier, 2014, pp. 1–344.

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

Coles, S., An Introduction to Statistical Modeling of Extreme Values, London, Springer-Verlag, 2001.

De Haan, L. and Ferreira, A., Extreme Value Theory: An Introduction, Springer Science & Business Media, 2007.

Dupuis, D. J. and M.-P. Victoria-Feser, "A robust prediction error criterion for Pareto modelling of upper tails," Canadian Journal of Statistics, Vol. 34, 2006, pp. 639–658.

Glotzer, D., Pipiras, V., Belenky, V., Campbell, B., T. Smith, "Confidence Interval for Exceedance Probabilities with Application to Extreme Ship Motions", REVSTAT Statistical Journal, Vol. 15, No 4, 2017, pp.537-563.

Embrechts, P., Klüppelberg, C., Mikosch, T., Modelling Extremal Events: For Insurance and Finance, Springer Science & Business Media, 2013.

Leadbetter, M. R., Lindgren, G. & Rootzen, H., Extremes and Related Properties of Random Sequences and Processes, Springer Series in Statistics, Springer-Verlag, New York-Berlin, 1983.

Mager, J., "Automatic threshold selection of the peaks over threshold method", Master's Thesis, Technische Universitat Munchen, 2015.

Maki, A., "Estimation Method of the Capsizing Probability in Irregular Beam Seas Using Non-Gaussian Probability Density Function," Journal of Marine Science and Technology, Vol. 22, Issue 2, 2016, pp.351-360.

Malara. G., Spanos, P.D., Arena. F., "Maximum roll angle estimation of a ship in confused sea waves via a quasi-deterministic approach," Probabilistic Eng. Mech., Vol. 35, 2014, pp. 75-81.

Pickands, J., "Statistical Inference Using Extreme Order Statistics," The Annals of Statistics, Vol. 3, No 1, 1975, pp. 119-131.

Romolo, A., Malara, G., Laface, V., Arena, F., "Space-time long-term statistics of ocean storms," Probabilistic Eng. Mech., Vol.44, 2016, pp.150-162.

Smith, T.C., "Validation Approach for Statistical Extrapolation," Contemporary Ideas on Ship Stability, Belenky, V., Neves, M., Spyrou, K., Umeda, N., van Walree, F., eds., Springer, 2018 (Accepted).

Smith, T. C., and Zuzick, A., "Validation of Statistical Extrapolation Methods for Large Motion Prediction," Proc. 12th Intl. Conf. on Stability of Ships and Ocean Vehicles (STAB 2015), 2015, Glasgow, UK.

Stephens, M. A., "Edf statistics for goodness of t and some comparisons," Journal of the American Statistical Association, Vol. 69, No. 347, pp. 730-737, 1974.

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

Weems, K., Belenky, V. and B. Campbell, "Validation of Split-time Method with Volume-Based Numerical Calculations," Proc. 15th Intl. Ship Stability Workshop, 2016, Stockholm, Sweden, pp. 103-108.

Weems, K., Belenky, V., "Including Diffraction and Radiation into Probabilistic Description of Capsizing," Proc. 16th Intl. Ship Stability Workshop, 2017, Belgrade, Serbia, pp. 117-123.

Weems, K., Belenky, V., Spyrou, K., "Numerical Simulations for Validating Models of Extreme Ship Motions in Irregular Waves," Submitted to 32nd Symposium on Naval Hydrodynamics, 2018, Hamburg Germany.