Ocean Engineering 120 (2016) 298–304

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

Application of the envelope peaks over threshold (EPOT) method for probabilistic assessment of dynamic stability

Bradley Campbell [a], Vadim Belenky [a,*], Vladas Pipiras [b]

[a] David Taylor Model Basin - NSWCCD, 9500 Macarthur Blvd., Bethesda, MD, USA
[b] University of North Carolina at Chapel Hill, 305 Hanes Hall, Chapel Hill, NC, USA

ARTICLE INFO

Article history:
Received 22 November 2015
Accepted 1 March 2016
Available online 24 March 2016

Keywords:
Large roll
Statistical extrapolation
Generalized Pareto distribution
EPOT

ABSTRACT

This paper reviews the research and development of the Envelope Peaks over Threshold (EPOT) method that has taken place in the last three years. The EPOT method is intended for the statistical extrapolation of ship motions and accelerations from time-domain numerical simulations, or possibly, from a model test. To model the relationship between probability and time, the large roll angle events must be independent, so Poisson flow can be used. The method uses the envelope of the signal to ensure the independence of large exceedances. The most significant development was application of the generalized Pareto Distribution (GPD) for approximation of the tail, replacing the previously used Weibull distribution. This paper reviews the main aspects of modeling the GPD, including its mathematical justification, fitting the parameters of the distribution, and evaluating the probability of exceedance and its confidence interval.

1. Introduction

The rarity of dynamic stability failures in realistic sea condition makes the problem of extrapolation inevitable. This can be illustrated in the following example. If we assume an hourly stability failure rate of $10^{-6}\,h^{-1}$ (Kobylinski and Kastner, 2003), then we can expect to see (on average) one failure every 1,000,000 h. If we require 10 observations for a reliable statistical estimate; then we need to simulate 10,000,000 h. Even if an advanced hydrodynamic code could run in a real time and a cluster with 1000 processors is dedicated to the task, it would take 10,000 h per condition (combination of seaway, speed and heading) to perform the assessment. The cost of the calculations prohibits direct simulation in this manner.

Additionally stability failure is associated with large-amplitude motions and is expected to be nonlinear. Indeed, capsize is related to the ultimate nonlinearity – transition to another equilibrium. In order to have enough fidelity to model this problem, the hydrodynamic code must be quite sophisticated (see a review by Reed et al., 2014). The probability of capsizing is the topic of a multiyear ONR research project titled "A Probabilistic Procedure for Evaluating the Dynamic Stability and Capsizing of Naval Vessels" (Belenky et al., 2016).

IMO document SLF 54/3/1, Annex 1 (IMO, 2011) defines intact stability failure as a state of inability of a ship to remain within design limits of roll (heel, list) angle combined with high rigid body accelerations. This includes also partial stability failure when a ship is subjected to a large roll angle or excessive accelerations, but does not capsize. Following the same logic one could also include an excessive pitch angle. As this study focuses on partial stability failure, peak over threshold method (POT) was chosen (Pickands, 1975). Introducing a threshold allows considering the data that are more influenced by nonlinearity; this incorporates changing physics into the statistical estimates.

To satisfy the requirement of independent peaks over threshold, the peaks of envelope were used instead of the peaks of the process itself (Campbell and Belenky, 2010). The review of this research effort is available from Belenky and Campbell (2012). That work included consideration of the relationship between probability and time, the probabilistic properties of peaks, application of envelope theory and the extreme value distribution.

The relationship between time and probability is key to the proper treatment of the partial stability failures. It may be modeled with Poisson, which requires the independence of the failure events. In the case of capsizing, the enforcement of Poisson Flow is not required, since capsizing can only occur once per record (the possibility of several capsizings within one record can be safely ignored for practical cases). Belenky and Campbell (2012) also review different ways of statistical characterization of the rate of events, the only parameter of the Poisson flow.

Classical POT methods use the Generalized Pareto Distribution (GPD) to approximate the tail of the distribution above a threshold. However, under certain conditions the GPD may be right

* Corresponding author.

http://dx.doi.org/10.1016/j.oceaneng.2016.03.006
0029-8018/Published by Elsevier Ltd.

bounded, that is, there is some value above which the probability of exceedances is zero. This is not a problem for conventional statistical consideration, when we are interested in the quantiles of the GPD (i.e., the probability is given and the level needs to be found). In ship stability generally the failure level is known and related to down flooding or cargo shifting angles and probability is to be found. The physical meaning of the right bound was not clear at that time (and still is not completely clear). As a result, the Weibull distribution was used for modeling the tail.

Normally distributed wave elevation was the subject of study in Belenky and Campbell (2012). This was a logical first test for these techniques. The study concluded that the distribution of large absolute values of peaks can be approximated by Rayleigh law. The Rayleigh distribution is a particular case of Weibull distribution when the shape parameter equals two. Thus, deviation of this parameter from two may be suitable for representing nonlinearity in a dynamical system.

To investigate the performance of a POT scheme based on the Weibull distribution, a model representing ship motions with realistic stability variation was used (Weems and Wundrow, 2013; Weems and Belenky, 2015). It was found that Weibull distribution does not have enough flexibility to approximate the tail of large-amplitude ship motions and the consideration of the GPD was started again.

Application of the GPD with EPOT produced very reasonable results (Smith and Zuzick, 2015). The techniques used to fit GPD, estimate the probability of exceedance of a given level and evaluate its uncertainty are described in Campbell et al. (2016, 2014) and Glotzer et al. (2016) and briefly reviewed in the rest of this paper.

2. Mathematical background

2.1. Distribution of order statistics

In order to understand why statistical extrapolation is possible when the underlying distribution is unknown, we begin with order statistics.

Consider a set of n independent realizations of random variable z. Assume that the distribution is given in a form of a cumulative distribution function (CDF) and probability density function (PDF). Sorting the observed values from the largest to smallest we have:

$$y_i = \mathrm{sort}(z_i) \quad i = 1, \ldots, n \tag{1}$$

Indeed, for randomly selected values y and z:

$$\mathrm{pdf}(y) = \mathrm{pdf}(z); \quad \mathrm{CDF}(y) = \mathrm{CDF}(z) \tag{2}$$

Consider a value that happens to be k-th in the list ($1 \le k \le n$). It is a random number, because, if one generates another set of realizations of variable z, and sorts them, another value will be the k-th. This random number is referred as k-th order statistic. Like any other random variable, y_k has its own distribution. This distribution is (see, e.g. David and Nagaraja, 2003):

$$\mathrm{pdf}(y|k) = \mathrm{pdf}(y)\frac{n!}{(k-1)!(n-k)!} \cdot (\mathrm{CDF}(y))^{\,k-1}(1-\mathrm{CDF}(y))^{n-k} \tag{3}$$

2.2. Generalized extreme value (GEV) distribution

Consideration of distribution of the largest value ($k=1$) when the number of observations n grows, leads to a limit, known as Generalized Extreme Value (GEV) distribution (see e.g. Coles, 2001):

$$\mathrm{pdf}(x) = \frac{1}{\sigma}\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-\left(1+\frac{1}{\xi}\right)} \cdot \exp\left(-\left(1+\xi\frac{x-\mu}{\sigma}\right)^{\frac{1}{\xi}}\right) \tag{4}$$

ξ is a shape parameter, σ is scale parameter ($\sigma > 0$); μ is a shift parameter, Eq. (4) is non-zero for:

$$\begin{aligned} &x > \mu - \tfrac{\sigma}{\xi} \quad \text{for } \xi > 0 \\ &x < \mu - \tfrac{\sigma}{\xi} \quad \text{for } \xi < 0 \end{aligned} \tag{5}$$

and is zero otherwise. If the shape parameter $\xi = 0$:

$$\mathrm{pdf}(x) = \frac{1}{\sigma}\exp\left(\frac{x-\mu}{\sigma}\right) \cdot \exp\left(-\exp\left(\frac{x-\mu}{\sigma}\right)\right) \tag{6}$$

for any values of x.

It is important that the limit (4)–(6) does not depend on the distribution z. That means that all the extreme values have the same distribution if one considers a sample of sufficient volume. This is the essence of the extreme value theorem, sometimes referred to as the Fisher–Tippet–Gnedenko theorem (see, e.g., Coles, 2001).

Direct application of the extreme value theorem for probabilistic assessment of dynamic stability can be found in McTaggar (2000), and McTaggart and de Kat (2000). However, several issues remained unresolved; including the question how large the sample should be (in terms of record length and number of records) to claim limiting properties of GEV.

2.3. Generalized Pareto distribution (GPD)

The large sample volume needed for direct application of the GEV is partially driven by the fact that only a single value (the largest one from the time window) is used to find the parameters of distribution. The desire to use more data leads to the idea of peaks over threshold methods.

Take μ as a threshold and find the distribution of the data exceeding this threshold, i.e., consider conditional probability. The generalized Pareto distribution is derived from the GEV with the threshold condition applied. The basic logic of this derivation is available in Coles (2001). The GPD is expressed as

$$f(x) = \begin{cases} \frac{1}{\sigma}\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-\left(1+\frac{1}{\xi}\right)}; \\ \text{if} \quad \mu < x, \quad \xi > 0, \quad or \\ \mu < x < \mu - \frac{\sigma}{\xi}, \quad \xi < 0 \\ \frac{1}{\sigma}\exp\left(-\frac{x-\mu}{\sigma}\right) \quad ; \\ \text{if} \quad \mu < x, \quad \xi = 0 \end{cases} \tag{7}$$

where ξ is the shape parameter, σ is the scale parameter ($\sigma > 0$) and μ is the threshold, above which, the GPD is believed to be applicable. Zero-value of the shape parameter is an important particular case, approximating the tail of some practically important distributions, including normal and Rayleigh distributions.

Eq. (7) expresses the second extreme value theorem, referred as Pickands–Balkema–de Haan theorem. It states that the tail of independent random variables can be approximated with the GPD.