Probabilistic methods for estimation of the extreme value statistics of ship ice loads

Wei Chai, Bernt J. Leira, Arvid Naess

PII: S0165-232X(17)30307-5
DOI: doi:10.1016/j.coldregions.2017.11.012
Reference: COLTEC 2489

To appear in: *Cold Regions Science and Technology*

Received date: 3 July 2017
Revised date: 12 October 2017
Accepted date: 20 November 2017

Please cite this article as: Wei Chai, Bernt J. Leira, Arvid Naess , Probabilistic methods for estimation of the extreme value statistics of ship ice loads. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Coltec(2017), doi:10.1016/j.coldregions.2017.11.012

Probabilistic methods for estimation of the extreme value statistics of ship ice loads

Wei Chai[a]*, Bernt J. Leira[a] and Arvid Naess[b, c]

[a]*Department of Marine Technology, Norwegian University of Science and Technology, Trondheim, Norway*

[b]*Centre for Ships and Ocean Structures, Norwegian University of Science and Technology, Trondheim, Norway*

[c]*Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway*

*corresponding author: chai.wei@ntnu.no

Abstract: It is well known that when a ship sails in ice-covered regions, the ship-ice interaction process is complex and the associated ice loads on the hull is a stochastic process. Therefore, statistical models and methods should be applied to describe the ice load process. The aim of this work is to present a novel method for estimating the extreme ice loads which is directly related to the reliability of the vessel. This method, briefly referenced to as the ACER (average conditional exceedance rate) method, can provide a reasonable extreme value prediction of the ice loads by efficiently utilizing the available data, which was collected by an ice load monitoring (ILM) system. The basic idea for the ACER approach lies in the fact that a sequence of nonparametric distribution functions are constructed in order to approximate the extreme value distribution of the collected time history. The main principle of the ACER method is presented in detail. Furthermore, the methods based on the classic extreme value theory are also introduced in order to provide a benchmark study.

Keywords: ice loads; ACER method; extreme value prediction

1. Introduction

Increased activities related to maritime transport and exploiting natural resources (e.g. oil and gas, mineral) in the arctic regions promote the requirement of ice-capable vessels and offshore structures (Hahn et al., 2017). For ships in the arctic, ice loads caused by ship-ice interaction represent the dominant load and the main structural challenge. Therefore, the knowledge of ice loads on the ship hull is essential for reliability-based design and operation of

ice-going vessels in the arctic regions.

However, a full understanding of such knowledge is far from achieved since the ship-ice interaction is a fairly complex process, which depends on the ice conditions (e.g. ice thickness, ice strength, ice age), the ship hull geometry as well as on the relative velocity between the vessel and the ice (Lubbad and Løset, 2011). Generally, for the first-year ice, the ship-ice interaction process is initiated by a localized crushing of the free ice edge, as the ship advances and penetrates the ice features, the contact area and crushing forces increase. The ice eventually deflects and the bending stresses promote a flexural failure at a certain breaking distance from the crushing region (Jordaan, 2001; Valanto, 2001). However, it should be noted that, even though the interaction between ice features and ship hull is governed by deterministic laws of mechanics, ice loads on ship hull is, for all practical purposes, random by nature (e.g. Figure 1).

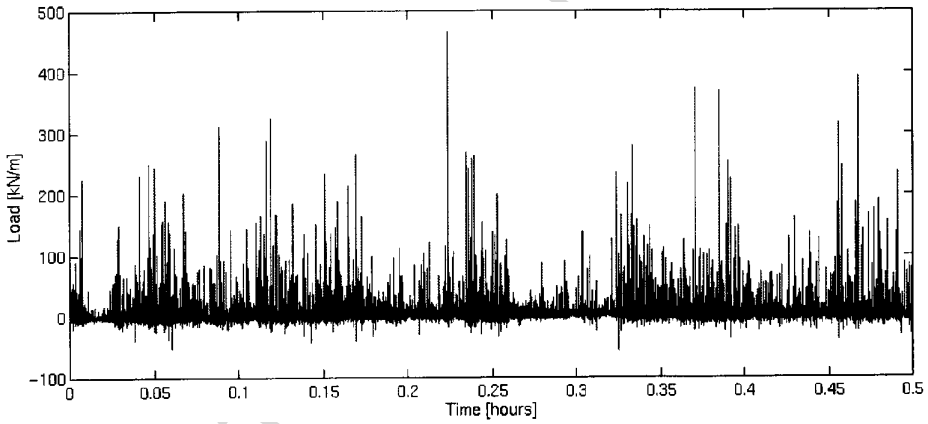

Figure 1. Stochastic nature of the ice loads measured on a frame (Lensu, 2002)

There are two categories of sources causing the randomness of ice-induced loads. One is the variation of the ice conditions in the arctic regions, which includes physical ice properties (e.g. ice thickness, density, porosity, floe size, etc.) and ice mechanical properties (e.g. flexural, tensile, shear, uni- and multi-axial compression strength, etc.). The other is related to the complex ship-ice interaction process in association with the crushing, bending, friction forces as well as the static and hydrodynamics forces (Riska, 1987). In addition, ship operations and movements also affect the ice breaking process. Therefore, probabilistic models and methods should be applied to describe the properties of ice loads. Furthermore, full-scale field experiments, with the information of ice loads, ice conditions, operation situations, and propulsion data etc., enriched by probabilistic methods have been recommended as a reliable basis for studying the ice loads statistics and the responses of the ice-going vessels.

Within the scope of probability theory, standard methods can be applied to analyze the stochastic time series, such as the statistical parameters (e.g. the mean value, standard deviation, skewness, kurtosis), correlation functions, power spectrum, statistical distributions and also extreme value statistics, etc. (Newland, 1993). The extreme values of ice loads are directly related to the reliability of the vessels since the ultimate limit states (ULS) are generally based on extreme load effects (Naess and Moan, 2012). Former studies of the extreme ice loads prediction are mainly based on the classic extreme value theory, which includes the peak amplitude approach (Ralph and Jordaan, 2013) and the asymptotic method (Lensu, 2002; Suominen et al., 2015; Suyuthi et al., 2012). In the former method, statistical models are introduced to describe the parent (initial) distribution of loads based on the measured peak amplitudes of the ice-induced loads. When the parent distribution is known, the extreme value distribution is given in the form of the power of encountered number of events. Kujala and his colleagues suggested that the Weibull probability distribution (Suominen and Kujala, 2010; Suyuthi et al., 2013) or the exponential distribution or lognormal distribution (Kujala and Vuorio, 1986) would give the best fit to the data collected in the Baltic Sea. In addition, (Jordaan et al., 1993) applied the event-maximum method and found that the tail of ice loads collected in the North Chukchi sea can be fitted as an exponential distribution by using probability paper. However, from the full-scale measurements of KV Svalbard in the sea area of Spitsbergen in 2007 and 2008, Suyuthi et al. (2014) mentioned that some sets of peak amplitudes cannot be well modeled by the traditional statistical models, such as the exponential and Weibull model. In their study, a generalized probabilistic model, i.e. a three-parameter exponential model, was proposed in order to provide a better description of the measured data. On the other hand, a proper number of time windows with equal duration are required for the asymptotic approach. The measured time series are then divided and the maxima values in each time window are identified for usage. Based on the available maxima data, an empirical cumulative distribution is estimated from the order statistics. Subsequently, extreme value models, such as the Gumbel distribution, are used for extreme value prediction (Lensu, 2002; Suominen et al., 2015).

The main objective of our work is to provide a reasonable extreme value estimation of ice-induced loads on the ship hull by utilizing available collected data from the full-scale measurements. The principles, feasibilities and weakness of the above two methods, i.e. the peak amplitude method and the asymptotic approach are given in the following sections in detail. In this work, a novel approach, namely the ACER (average conditional exceedance rate) method, is introduced for the extreme value prediction of the collected ice loads. Unlike the

former methods based on the parametric distribution functions, this method estimates the exact extreme value distribution by constructing a sequence of nonparametric distribution functions, i.e. the ACER functions. Then, the empirical ACER functions estimated from the measured data are combined with an optimization procedure for the purpose of predicting long return period design values. This extrapolation method is based on an assumed tail behavior of the extreme value distribution consistent with the form of the asymptotic extreme value distribution, thereby allowing for sub-asymptotic data to be included and still retain asymptotic consistency (Næss and Gaidai, 2009). All the underlying ice load data available points to the Gumbel distribution as the correct asymptotic distribution in the case of stationary data, and, in fact, also for the long term, nonstationary case. This method has been well benchmarked against existing conventional methods and it has been applied for predicting the extreme wind speed (Karpa and Naess, 2013), wave height (Gaidai et al., 2017) , and ship roll response (Gaidai et al., 2016).

The present paper is organized as follows. Section 2 describes the full-scale measurement of the KV Svalbard in the winter of 2007. The collected time series of ice-induced loads and associated information with respect to the expedition route are presented. Principles of different extreme value prediction methods are described in Section 3, and Section 4 shows the relevant results of extreme ice loads prediction. The advantages of the ACER method are shown by relevant examples. The method proposed as well as the results and conclusions obtained in this work could hopefully promote the development of reliability-based design and operation for ice-going vessels.

2. Full-scale measurement for the ice loads

A growing number of full-scale measurement campaigns have been carried out in various ice-covered regions since 1970s and most of the full-scale measurements have been summarized in the Arctic Technology report (ISSC, 2015). The full-scale measurements can be classified as long-term and short-term in general. Long-term measurements are usually taken over several winter seasons and they can be applied to study the relationship between the prevailing ice conditions and the ice-induced loads. However, long-term measurements have been performed only in the Baltic Sea (such as MS Kemira voyages during the winters from 1985 to 1988) and the Antarctic Sea (Kujula, 1994). Basically, most of the full-scale measurements are based on the short-term measurement in which the time period is hours or

minutes and ice loads collected by short-term measurements have been widely used for extreme ice load prediction and fatigue damage evaluation of the vessels sailing in ice-covered regions.

In this work, the short-term full-scale measurement onboard the Norwegian coastguard vessel KV Svalbard is selected for extreme ice load prediction. This Polar-10 icebreaker has been used for scientific expeditions in the Barents Sea and around the Svalbard islands in recent years. During the years 2006-2008, Det Norske Veritas (now DNV·GL) launched an Ice Load Monitoring (ILM) project in the sea area of Svalbard islands (Leira et al., 2009). In this project, an ILM system was installed in order to collect the data of ice conditions along the expedition route, the navigation data (e.g. ship speed, propulsion power) and records of ice-induced loads on the hull, etc. (Nyseth, 2006). As a part of the ILM project, a two-week expedition was performed in the vicinity of Svalbard islands in March 2007 and an ice load time series of six hours was selected from the expedition for current study.

Generally, ice-induced loads can be divided into the local ice load and the global ice load. The former term refers to the loads induced by the ice on the ship's shell structure and acting on the transverse frames during the ice-breaking process and the latter term refers to the loads that causes a global structural response in the hull girder. In the ILM project, the fiber optic strain sensors were instrumented in order to record the local ice loads in the bow region (L1-L4 in Figure 2) and the bow intermediate region (L5-L8) where the loading levels are significantly higher than other regions. The measured shear strain is converted into shear stress and then the total shear force can be obtained by integrating the (pre-assumed) shear stress distribution over the cross section of the frame. Therefore, the local ice load can be taken as the difference between the shear forces at the upper and lower parts of the frame. The load on a transverse frame has a unit [kN], but for practical applications it is usually treated as a line loads by dividing the load by the frame spacing of the transverse frame and then the new unit is [kN/m] (Suyuthi et al., 2012).

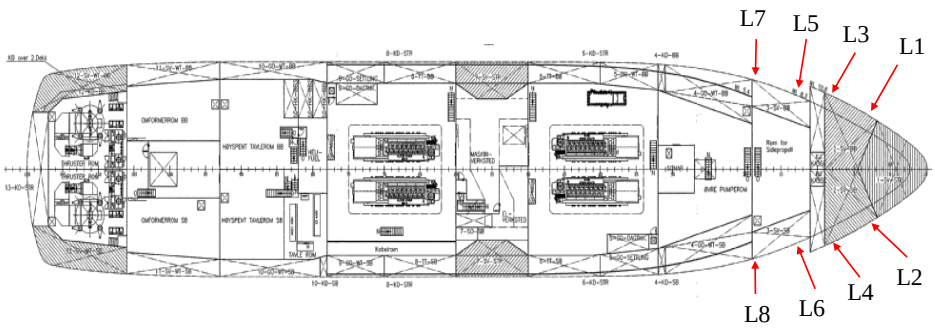

Figure 2. Strain sensor arrangement (L1 represents location number 1)

As mentioned in Section 1, due to the variations in the ice conditions and the complex ship-ice interaction process, the magnitude of ice loads changes significantly even in short-term expeditions. For each location number shown in Figure 2 in the bow and the bow intermediate regions, the time series of ice loads with a duration of 6 hours, recorded on March 28, 2007 from 12:30 to 18:30 pm, are selected for analyse in this work. The corresponding route for the expedition in the vicinity of Svalbard islands is presented in Figure 3.

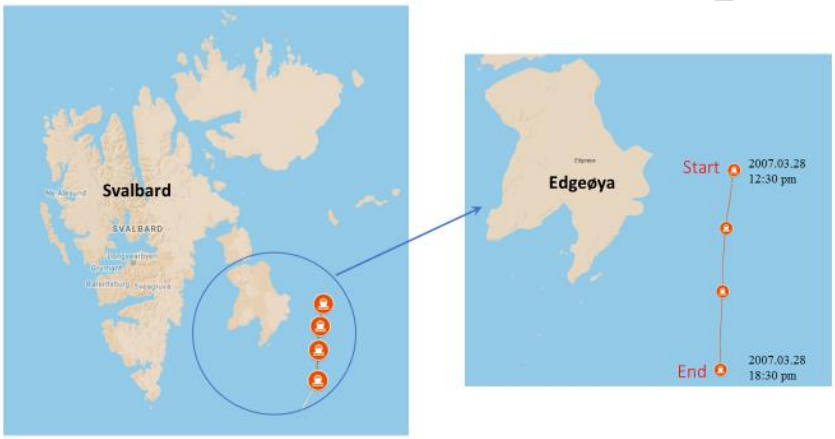

Figure 3. Route of the expedition on March 28, 2007 (6 hours ice loads records started from 12:30 to 18:30 pm)

It has been observed in former studies (Leira et al., 2009; Lensu, 2002; Suominen et al., 2017) that the time series of the measured ice-induced load looks like a sequence of impulses with sharp peaks, e.g. Figure 1. The generation of ice loads which closely relates to the ice-breaking process can explain this phenomenon. Generally, ice-induced loads can be divided into three stages (see Figure 4): approaching stage, crushing stage and the disengaging stage (Kotilainen et al., 2017). The first stage begins when the ice-induced load starts increasing. At this stage, the transvers frame is not in contact with the ice, but the force is the influence from other frames. In the crushing stage, when the hull gets in contact with the ice edge, a crushing failure mechanism takes place first and then the load increases with the ship hull penetration process. When the accumulated force is high enough to initiate bending failure of the ice, a sudden drop of the load can be observed in the disengaging stage. At this stage, the measured load is mainly caused by the reflections from the other frames (Varsta, 1983). Afterwards, such three-stage process will be repeated when the ship hull contacts with subsequent ice edges. In addition, the peak value of the load during the interaction event is referred to as the ice load

peak. By recording the values of ice load peaks during the expedition, the extreme value of ice loads can be estimated.

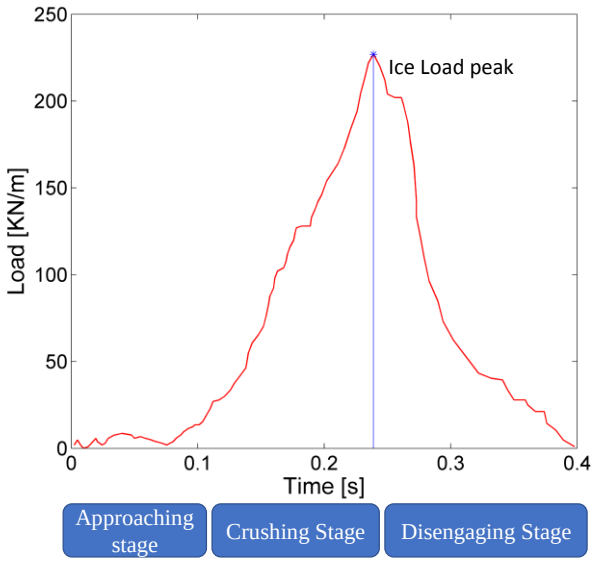

Figure 4. Three stages of ice loads generation process

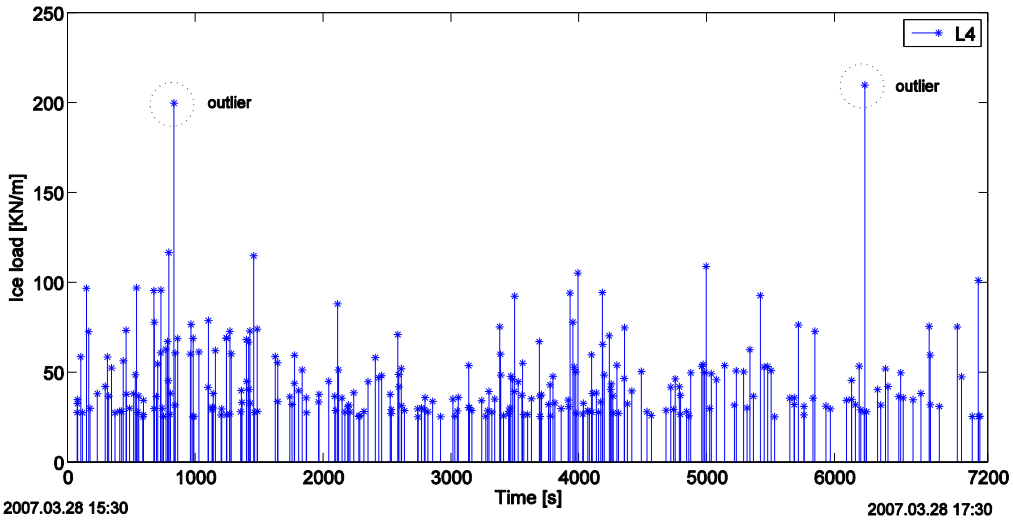

Figure 5. Ice load peaks during the selected period for L4, represented as stem plots

In this work, in order to identify the start and end of the force peak, a lower threshold of 25 kN/m is selected and the method for identification of ice load peaks can be found in Suominen et al. (2017). Figure 5 presents an example of the records for the ice load peaks for L4 during the expedition shown in Figure 3. It is of some importance to note that the sample record from L4 location contains outlying observations which are pointed out in Figure 5. Such ice loads

could be caused by extremely serious ice conditions (such as the ice ridges) encountered by the vessel or by measurement error of the sensors. Such outliers would affect the performance of the classic extreme value theory and their influence should not be ignored. In the subsequent studies, these outlies values are incorporated in the data sets for numerical analysis.

3. Extreme value prediction

3.1 Problem description

In this part, the problem of determining the extreme value distribution is described. Consider a stochastic process $X(t)$, which has been observed over a time interval [0, T]. Assume that values $X_1, \ldots, X_N$, which were derived from the observed process, are allocated to the discrete times $t_1, \ldots, t_N$ in the time interval [0, T]. This could be simply the observed values of $X(t)$ at each t_i, i=1, …, N, or it could be average value or peak values over small time intervals centred at the t_i's (e.g. the ice load peaks shown in Figure 5). The extreme value among the N outcomes of the stochastic process $X(t)$ is defined as: $M_N = \max\{X_1,\ldots, X_N\}$ and the extreme value distribution for large values of η is expressed as:

$$P(\eta) = \mathrm{Prob}(M_N \le \eta) = \mathrm{Prob}(X_1 \le \eta,\ldots, X_N \le \eta) \tag{1}$$

This paper is focused on accurately determining the extreme values from observed ice load peaks during the expedition. The values $X_1, \ldots, X_N$ in the following parts are taken as the recorded ice load peaks and the principles of the peak amplitude approach, the asymptotic method and the ACER method are introduced in this Section.

3.2 Peak amplitude approach

This method requires that the random variables $X_1, \ldots, X_N$ are independent and identically distributed with common distribution $F_X(x)$. If the initial distribution function, $F_X(x)$ is known, the extreme distribution can be evaluated directly and accurately through the classic extreme value theory (Ochi, 1990):

$$P(\eta) = \mathrm{Prob}(X_1 \le \eta,\ldots, X_N \le \eta) = \prod_{i=1}^{N} \mathrm{Prob}(X_i \le \eta) = [F_X(\eta)]^N \tag{2}$$

where N is the number of ice load peak events for a specific duration under consideration.

Figure 6 presents the relationship between the initial distribution and the extreme value distribution, which are given in the form of the probability density function (PDF). For a certain small level of exceedance probability level λ, the corresponding extreme value is given as:

$$\eta = F_X^{-1}\left[(1-\lambda)^{\frac{1}{N}}\right] \tag{3}$$

Statistical analysis is required to obtain the initial distribution and relevant former studies are mentioned in Section 1. However, there is one critical point limiting the application of the peak amplitude approach: the requirement of stationarity for the measured ice load peaks. Specifically, the ice loads are influenced by the random ice conditions to a large extent. However, nonstationary cases of the ice condition and the ice loads are very common during the full-scale measurements, even for short-term cases. The statistical models for the measured data require some form of stationarity, which can only be satisfied for some ideal cases with respect to prevailing ice conditions and ship operation characteristics (Suyuthi et al., 2013). Furthermore, small discrepancies in the estimates of the initial distribution $F_X(x)$, such as the influence from the outliers in the time series, can lead to substantial discrepancies of the extreme value distribution since $N >> 1$.

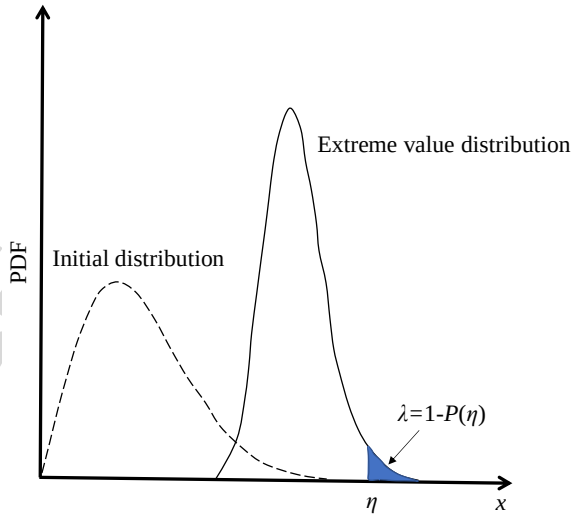

Figure 6. Illustration of the relationship between the initial distribution and the extreme value distribution

3.3 Asymptotic method

In this method, the time series of the ice load peaks is divided into a certain number, K, of intervals with equal duration t (e.g. 1 minute, 5 minutes) and the maxima value in each interval Y_j (j=1, …, K=T/t) are identified. Based on the collected maxima data, an empirical cumulative distribution is estimated and then the type I asymptotic extreme value distribution, i.e. the Gumbel distribution, is applied to estimate the extreme value by fitting the empirical cumulative distribution (Lensu, 2002). Therefore, the extreme distribution is given as:

$$P(\eta) = G_Y(\eta) \tag{4}$$

in which $G_Y(y)$ is the Gumbel distribution with the following expression:

$$G_Y(y) = \exp\left\{-\exp\left[-\left(\frac{y-\beta}{\alpha}\right)\right]\right\} \tag{5}$$

where α and β are the parameters for the Gumbel distribution and they can be estimated by ordinary fitting of the empirical cumulative distribution, such as least square fitting in a probability paper, the method of moment or maximum likelihood method.

This method is also known as the Gumbel method and the Gumbel probability paper is applied in this work to estimate the parameters. For the small level of exceedance probability λ, the corresponding extreme value is given as:

$$\eta = G_Y^{-1}(1-\lambda/K) \tag{6}$$

and correspondingly, the extreme value η can be estimated by extending the line to the value of the cumulative distribution function $G_Y(\eta) = (1-\lambda/K)$.

3.4 ACER method

The above two methods are based on the parametric distribution functions, while the ACER method estimates the extreme value distribution by constructing different orders of the ACER functions, which are available for both the stationary and non-stationary data sets. The principle and development of the ACER functions are presented in Appendix.

With the time series of the ice load peaks, the extreme value can be expressed in the following manner (Gaidai et al., 2016):

$$\begin{aligned} P(\eta) &\approx P_k(\eta) \approx \exp\left(-(N-k+1)\cdot \mathrm{ACER}_k(\eta)\right) \\ &\approx \exp\left(-(N-k+1)\cdot \hat{\varepsilon}_k(\eta)\right) \end{aligned} \tag{7}$$

where k is the order of the ACER function, which is presented as Equation A.11. $\hat{\varepsilon}_k(\eta)$ is the empirical ACER function of order k, which can be obtained by applying the existed time series. As the order k increases, the accuracy of Equation 7 improves, but the amount of data for calculating $\hat{\varepsilon}_k(\eta)$ reduces.

Next, an efficient extrapolation scheme is introduced in order to provide a reasonable estimation of the extreme value. The extrapolation approach for the purpose of extreme response prediction derives from the fact that for ships and marine structures being considered, the mean upcrossing rate and the ACER functions of response level η are in general highly regular in the tail region (Chai et al., 2016; Naess and Moan, 2012). Specifically, in the tail region (e.g. $\eta \geq \eta_0$), the ACER function behaves similarly to $\exp\{-a(\eta-b)^c\}$, where $a > 0$, $b \leq \eta_0$, and $c > 0$ are suitable constants. Therefore, it may be assumed:

$$\varepsilon_k(\eta) \approx q_k \exp\{-a_k(\eta - b_k)^{c_k}\}, \eta \geq \eta_0 \tag{8}$$

where a_k, b_k, c_k and q_k are suitable constants, that in general will be dependent on the order k. Note that any form of the ACER function of the type given by Eq. (8) will provide an extreme value distribution that is asymptotically of the Gumbel type, but yet avoids the limitation of assuming a global Gumbel distribution for the measured extremes, which in reality are rarely asymptotic.

In order to find the optimal values for the parameters a_k, b_k, c_k and q_k, an optimized fitting on the log level is selected. These parameters can be determined by minimizing the following mean square error function:

$$F(q_k, a_k, b_k, c_k) = \sum_{i=1}^{M} \rho_j \left| \ln \hat{\varepsilon}_k(\eta_i) - \ln q + a(\eta_i - b)^c \right|^2 \tag{9}$$

where η_i, $i = 1, \ldots, M$ are levels at which the ACER functions have been empirically estimated, ρ_j denotes a weight factor that puts more emphasis on the more reliable data points. This is in agreement with the principle of best linear unbiased estimator (BLUE) for uncertain data using weighted linear regression (Montgomery et al., 2012). In this paper we shall use a modified version that fits the purpose here, viz. $\rho_i = (\ln CI^+(\eta_i) - \ln CI^-(\eta_i))^{-2}$, where CI^+ and CI^- are the bounds of the 95% confidence interval (CI) determined by Equation A.19.

In the following part, a simplified and transparent two-parameter optimization method based on the Levenberg-Marquardt method is introduced. This is realized by considering Equation 9 when the values of b_k and c_k are obtained, the optimization problem reduces to a standard weighted linear regression problem. That is, with given values of b_k and c_k, the optimal

values of a_k and $\ln q_k$ are found using closed form weighted linear regression formulas in terms of ρ_i, $y_i = \ln\hat{\varepsilon}_k(\eta_i)$ and $x_j=(\eta_j-b)^c$. The optimal values of a_k and q_k are given by the relations:

$$a_k^*(b_k, c_k) = -\frac{\sum_{i=1}^{M} \rho_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{M} \rho_i (x_i - \bar{x})^2} \tag{10}$$

and

$$\ln q_k^*(b_k, c_k) = \bar{y} + a_k^*(b_k, c_k)\bar{x} \tag{11}$$

where $\bar{x} = \sum_{i=1}^{M} \rho_i x_i \Big/ \sum_{i=1}^{M} \rho_i$ and $\bar{y} = \sum_{i=1}^{M} \rho_i y_i \Big/ \sum_{i=1}^{N} \rho_i$.

In order to calculate the final optimal set of parameters, the Levenberg-Marquardt method can now be applied to the function in order to find the optimal values of b_k^* and c_k^*, and then the corresponding q_k^* and a_k^* can be calculated from Equations 10 and 11. For estimation of a confidence interval for a predicted value of the ACER function provided by the optimal curve, the empirical confidence band is reanchored to the optimal curve. The fitted curves to the reanchored confidence band boundaries can be used to determine an optimized confidence interval of the predicted value. As a final point, the predicted value is not very sensitive to the choice of the tail maker η_0, provided it is chosen with some care. With the assistance of the efficient extrapolation scheme, which is based on the assumption of regularity of the ACER functions in the tail regions, the empirical estimations of the ACER functions with respect to the far tail region can be achieved with reliable accuracy for most practical prediction purposes.

4. Numerical results and discussions

4.1 Prediction by the ACER method

In this part, time series of ice load peaks recorded at locations L4 and L7 during the 6-hour expedition are selected for analysis, where Figure 5 shows an example of ice load peaks measured at L4. Figures 7 and 8 present the empirical ACER functions, $\hat{\varepsilon}_k(\eta)$, for different orders of k and $\hat{\varepsilon}_k(\eta)$ is plotted versus the ice load amplitudes with different values of k for these two locations.

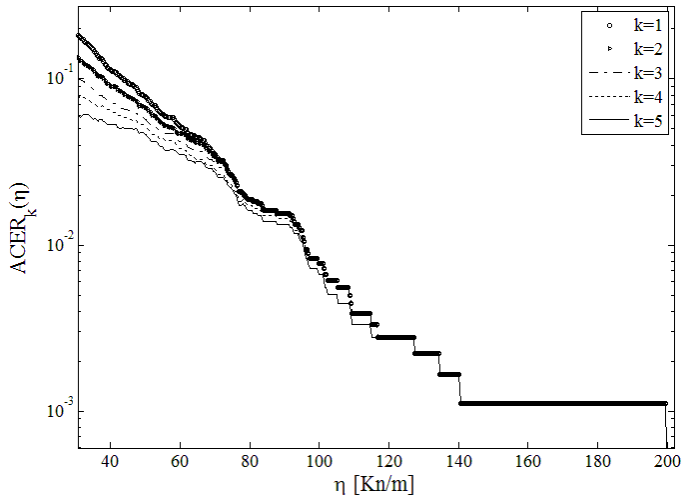

Figure 7. ACER functions with different order *k* at location L4

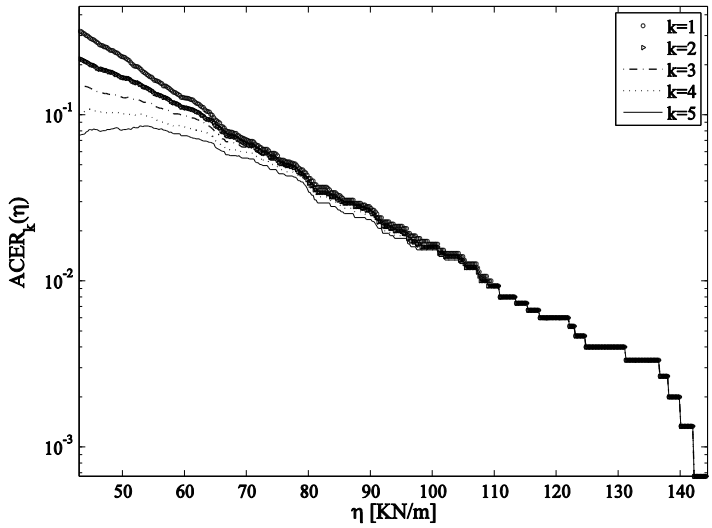

Figure 8. ACER functions with different order *k* at location L7

From Figures 7 and 8, it may be seen that for the lower range of ice load values there is an effect of dependence between the data points. However, this effect does vanish in the tail, indicating that all the ACER functions will coalesce in the deep tail. This allows the use of the first ACER function $\hat{\varepsilon}_1(\eta)$ for extrapolation purposes. This is advantageous since the first ACER function is the one which is most accurately estimated because more data are available for its estimation, thereby enhancing the precision. This illustrates the usefulness of the ACER function plot as a diagnostic tool for investigating the influence of possible dependence between data in the ice load time series for these two cases. Another important observation to be drawn from the ACER plots in Fig. 8, is that the empirical ACER functions are all close to linear in

the tail. This would indicate that the assumption of a Gumbel distribution to describe the extreme value data would work fairly well for the ice load peaks collected at L7, because a Gumbel distribution would correspond to a straight line in the ACER plot. However, it should be noted that one of the requirements for adopting it is violated: the ice load time series is clearly not stationary for the current case.

The target exceedance probability during the expedition is selected to be 0.1, which means 1-Prob ($M_N \leq \eta$) = 0.1. Correspondingly, the 90% fractile value can be predicted by applying Equation 7. Tables 1 and 2 present the estimated extreme ice loads based on different orders of ACER functions and different values of tail maker η_0 for L4 and L7, respectively. The tail maker η_0 represents the start point of the tail region and the extrapolation scheme, a sensitive study should be executed in order to determine its value. It is seen in Table 1 that for the empirical $ACER_1$ function, the estimated 90% fractile values present a clear manifestation of convergence when $\eta_0 \geq 45$ kN/m. Similarly, as given by Table 2, the predicted 90% fractile values based on the empirical $ACER_1$ function tend to converge when $\eta_0 \geq 50$ kN/m. Subsequently, Figures 9 and 10 present the estimations for extreme ice loads based on the $ACER_1$ functions for L4 and L7, respectively. The fitted $ACER_1$ functions in the far tail regions and the corresponding fitted 95% confidence intervals are also presented in these two figures.

Table 1. Predicted extreme ice loads (kN/m) for L4 by the ACER method with different order *k* and different values of tail maker η_0 (kN/m)

Order k	η_0=45	η_0=50	η_0=55	η_0=60	η_0=65
1	185.2	186.9	186.3	185.3	188.4
2	178.7	182.5	185.8	184.7	184.1
3	167.9	172.5	176.6	180.9	178.7
4	162.7	164.8	167.2	170.9	174.7
5	154.6	157.5	160.3	163.1	167.4

Table 2. Predicted extreme ice loads (kN/m) for L7 by the ACER method with different order *k* and different values of tail maker η_0 (kN/m)

Order k	η_0=45	η_0=50	η_0=55	η_0=60	η_0=65
1	225.7	219.3	219.3	206.6	210.7
2	202.1	207.7	216.3	217.6	219.7
3	192.7	199.2	205.1	212.3	209.4
4	189.1	195.2	200.9	208.5	208.6
5	191.0	194.6	203.5	210.3	213.0

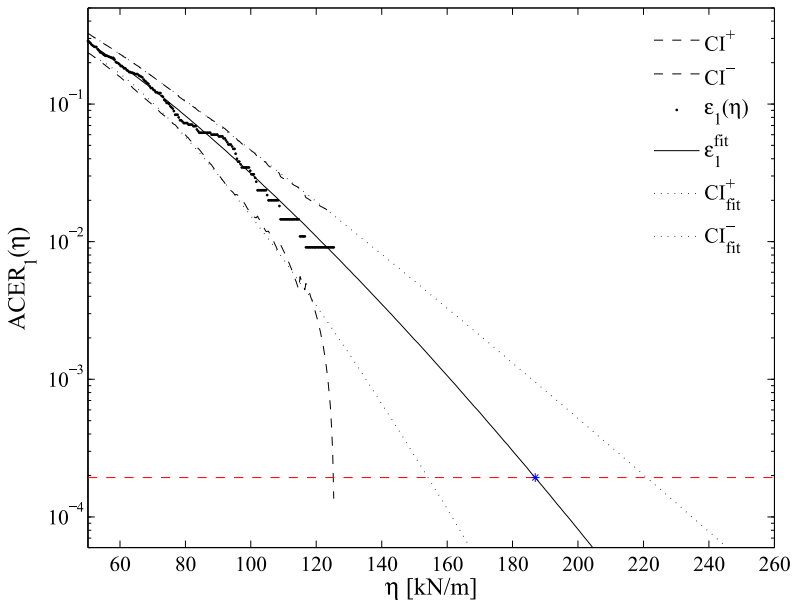

Figure 9. Extreme ice load prediction based on $\hat{\varepsilon}_1(\eta)$ for L4, with the starting point, η_0=50.0, q_1 = 0.6336, a_1 = 0.0011, b_1 =19.99, c_1 = 1.346, $(CI^+\text{-}CI^-)/\eta_{ACER}$ = 35.7%.

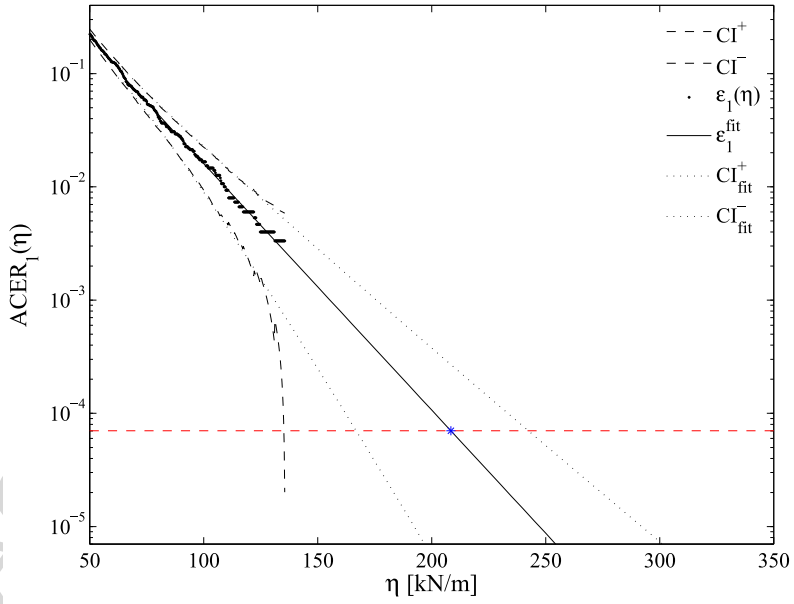

Figure 10. Extreme ice load prediction based on $\hat{\varepsilon}_1(\eta)$ for L7, with the starting point, η_0=62.5, q_1 = 0.3097, a_1 = 0.0501, b_1 =40.9, c_1 = 1. 0, $(CI^+\text{-}CI^-)/\eta_{ACER}$ = 36.3%.

It is shown in Figures 9 and 10 that the ACER method for prediction of the extreme value is implemented by expressing the extreme value distribution in terms of the average conditional exceedance rate and the tail behavior of the empirical extreme value can be accurately captured

by fitting a parametric function to the empirical exceedance rate. The fitting procedure puts more weight on the empirical estimates when they are more accurate. This implies that the ACER method is not sensitive to the outliers in the time series of ice load peaks and relevant example is given by the extreme value prediction for L4 (see Figures 5 and 9).

4.2 Prediction by the Gumbel method

The performance of the asymptotic method, i.e. the Gumbel method, even if it is strictly not applicable, is studied in this part. As a justification, we refer to the observation mentioned in Section 4.1, that the tail of the empirical nonparametric extreme value distribution provided by the ACER method, is linear in the tail in Figure 10, which would correspond to a Gumbel distribution shape in the tail. Since the Gumbel method is based on the maximum value in each time window with equal duration, the influence of the duration of the time window *t*, on the extreme value prediction is investigated at first.

The target 90% fractile value during the expedition can be obtained by Equation 6 in which the number of time windows *K* is equal to the ratio between the exposure period *T* and the duration of time window *t*. The probability paper with linear regression is applied to estimate the extreme values based on different time windows. Figure 11 shows the results of extreme value predictions by the Gumbel method for different values of *t* based on the time series of ice load peaks collected at L7. It is seen in Figures 10 and 11, for the ice load peaks collected at L7, the Gumbel method provides close estimations of the extreme value to the results provides by the ACER method and the length of the time window *t* can influence the estimated value to some extent. Also, it is seen in Figure 11 that the slopes appear to be quite sensitive to the fitting ranges, while the ACER method shown in Figure 10 has satisfactory performance due to the robustness of the extrapolation scheme.

In the meantime, the uncertainties associated with the Gumbel method should be considered. Similar to the ACER method, the 95% confidence interval of the estimated extreme value is introduced in order to qualify the uncertainties and an estimate of the confidence interval can be obtained by applying a parametric bootstrapping method (Davison and Hinkley, 1997). Table 3 shows the 95% confidence intervals for the extreme values provided by the Gumbel method with different length of time intervals based on 100,000 bootstrap samples and Figure 12 presents the empirical PDF of the predicted extreme value based on the samples with *t* =5 minutes.

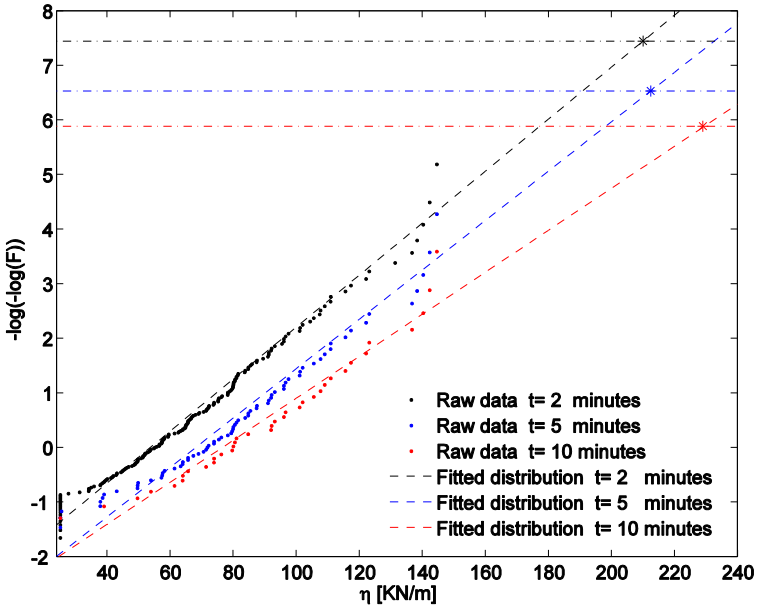

Figure 11. Gumbel method for the extreme value prediction based on different time windows for L7

Table 3. Predicted extreme ice loads for L7 by the Gumbel method with different time windows

Time window *t*	η_{Gumbel} (KN/m)	95 % CI (KN/m)	$(Cl^{+}\text{-}Cl^{-})/\ \eta_{Gumbel}$
2 minutes	210.1	(187.5, 235.1)	22.65 %
5 minutes	212.5	(179.5, 249.4)	32.89 %
10 minutes	228.3	(180.2, 283.5)	45.30 %

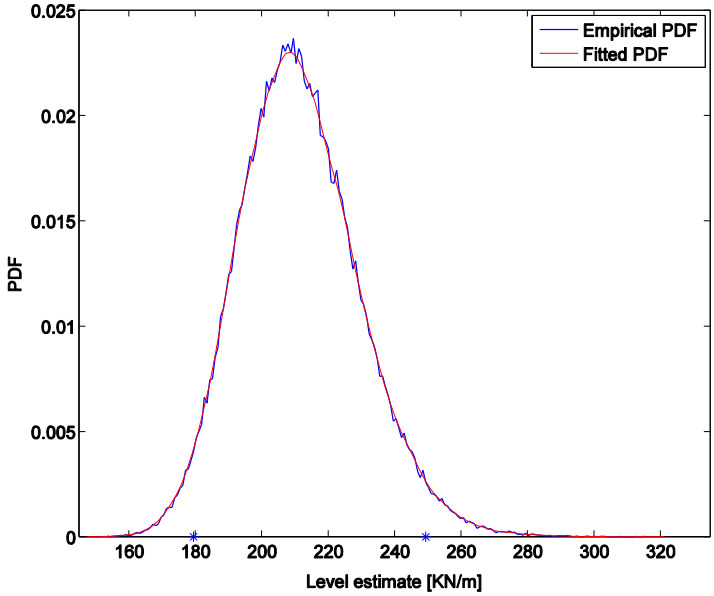

Figure 12. The empirical PDF of the predicted extreme value based on the samples with 5 minutes of time windows for L7, the * indicates the limit of 95% confidence interval.

It is seen in Table 3 that the uncertainty of the Gumbel method increases with the time window *t*. Uncertainties, i.e. the ratio between the confidence interval width and the estimated value are directly related to the size of each sample. From this aspect, the value of t = 2 minutes provides the best performance with respect to the uncertainties. On the other hand, it is seen in Figure 11, decreased length of the time window corresponds to increased number of lower loads in the probability paper and the performance of the linear fitting in the upper tail will be influenced. Therefore, based on the above considerations, the length of time window t is recommended to be 5 minutes for the following study. In addition, it is also seen in Table 3 and Figure 10, the ratios between the confidence interval width to the estimated extreme value provided by the ACER method (k =1) and the Gumbel method with t =5 minutes two methods are very close. However, it should be realized that the uncertainty estimation for the Gumbel method by parametric bootstrapping is based on the rather unrealistic assumption of a global Gumbel distribution for the observed data. This would typically tend to underestimate the real uncertainty.

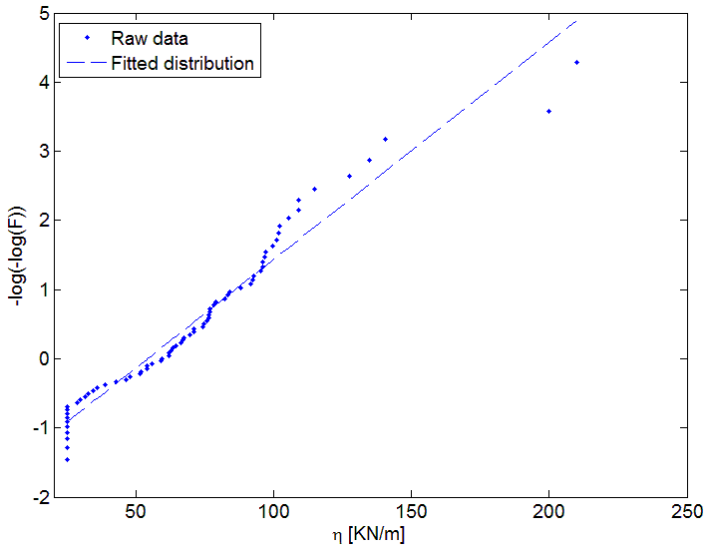

Figure 13. Gumbel method for the extreme value prediction based on the samples with 5 minutes of time windows for L4

Furthermore, the Gumbel method with t = 5 minutes is applied to predict the extreme value of ice load peaks collected at L4 and relevant probability paper is presented in Figure 13. In this case, the 90% fractile value is 244.5 KN/m which is significant higher than 188.4 kN/m

provided by the ACER method. There are two reasons to explain the difference. Firstly, it is seen in Figure 13 that the Gumbel method is affected by the large outliers in the upper tail region to some extent and its influence will be strengthened as the size of the sample decreased. Also, it is found that for such case shown in Figure 13, the overall performance of linear regression for the sample data in the probability paper is not satisfactory. The major weakness of the asymptotic approach is that the asymptotic extreme value theory itself cannot be used in practice to decide to what extent it is applicable for the observed data (Næss and Gaidai, 2009).

4.3 Extreme ice loads on different locations

As mentioned before, the ship-ice interaction is a complex process which relates to multiple factors, thus extreme ice load prediction at different locations could provide some valuable information to understand the ice-breaking process as well as for the aspects relevant to the reliability of the vessel. The ice-breaking process is highly related to the three-stage ice loads generation process described in Section 2 and the process can also be divided into three phases. The initial breaking phase starts when the ship movement brings the hull in contact with the ice features. Then the ice features are break, pushed aside and slide along the hull. The last phase is associated with the cleaning of the ice features. As a supplement to the description of the ice-breaking process, two examples of the ice-breaking by the KV Svalbard in the ice-covered regions are shown in Figure 14.

(a) (b)

Figure 14. KV Svalbard in ice-covered seas

Based on the time series of ice load peaks collected at different locations, the extreme values of the ice loads estimated by the ACER method and the Gumbel method are presented in Figure 15 and Table 4. In addition, for the ACER method, k = 1 is selected and the time window t = 5

minutes is applied in the Gumbel method. It is shown in Figure 15 and Table 4 that the Gumbel method is able to provide satisfactory estimation of the extreme value for some cases but its performance depends on the observed data at different locations. The discrepancies of the estimated extreme values (e.g. L1 and L4) are caused by different principles behind the ACER method and the Gumbel method.

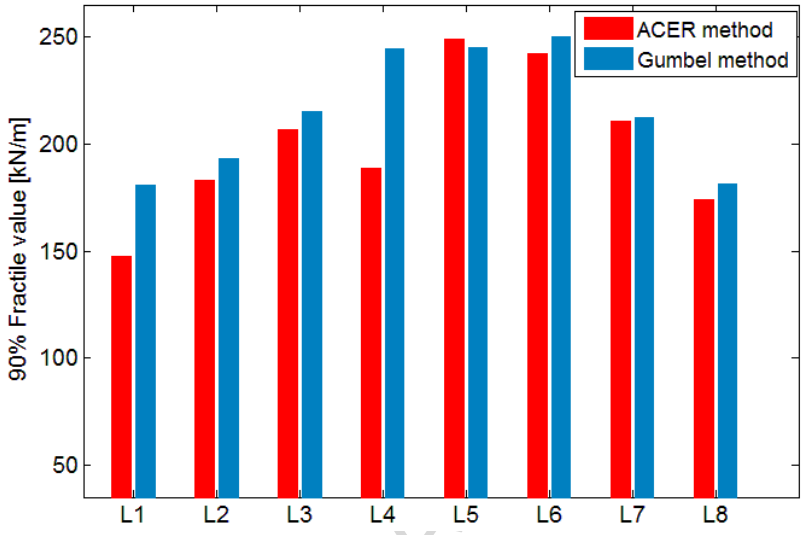

Figure 15. 90% fractile values of the ice load peaks collected at different locations during the 6-hour expedition

Table 4. Numbers of ice load peaks recorded (N) at different locations and the 90% fractile values of the ice load peaks estimated by the ACER method and the Gumbel method

Location	N	η_{ACER} (kN/m)	η_{Gumbel} (kN/m)
L1	565	147.5	180.8
L2	456	182.8	193.4
L3	539	206.7	214.9
L4	545	188.4	244.5
L5	1210	249.1	245.2
L6	2148	242.0	249.8
L7	1501	210.7	212.4
L8	708	173.7	181.1

From the results obtained by the ACER method, it is found that there exist certain differences of the estimated extreme values on the locations L1 and L2, L7 and L8, even though these locations are symmetrical with respect to the centerline of the ship. An explanation for

this phenomenon could be that the prevailing ice conditions in the transverse direction of the vessel are different and vary significantly during the expedition.

Also, it is interesting to notice in Figure 15 and Table 4, for the current vessel, the extreme ice loads in the bow intermediate region (especially at L5 and L6) are much severer than in the bow area. The bow region is the first region to contact and break the ice features during the ice-breaking process. However, it is observed in Table 4, the ship-ice interaction process occurs more frequently in the bow intermediate area. In reality, there are many possible explanations for the difference of the extreme ice loads in different regions, such as the bow hull shape, e.g. the vertical frame angles can change rapidly along the hull and the capacity of icebreaking in the bow intermediate area is inferior to the bow region which has well-designed structure type (see Figure 14) to ensure the efficiency and capacity of icebreaking. Also, this difference could be explained by the fact that some broken ice features do not slip along the ship hull and then the accumulated broken ice features disturb the ice-breaking process in this region. However, the real reason for such difference cannot be easily explained with the current short-term data.

5. Conclusions

In this work, the ACER method was introduced to predict the extreme ice loads by utilizing the records of ice load peaks collected by a full-scale measurement in the vicinity of Svalbard islands. The parametric distribution function-based Gumbel method were introduced for comparison.

The ACER method is implemented by expressing the extreme value distribution in terms of a sequence of nonparametric ACER functions, and then an efficient extrapolation scheme is applied to capture the tail behaviour of ACER functions for extreme value prediction. Based on the numerical studies and discussions above, it has been found that the ACER method has the following advantages when compared with the peak amplitude method and the Gumbel method.

Firstly, for the ACER method, both the stationary and non-stationary data sets can be analysed, while for the peak amplitude approach, stationary is required for the ice load peaks. Moreover, the ACER method with different order k can serve as an effective diagnostic tool to reveal the dependence of the neighbouring maxima, which could also influence the performance of the peak amplitude approach. Since relevant predictions are based on the assumption of stationarity of the ice loads process, the peak amplitude approach is not recommended for extreme ice loads prediction.

Secondly, the Gumbel method is based on the asymptotic assumption but its performance depends on the observed data. For some cases, the overall performance of linear regression in the probability paper is not satisfactory and its performance could also be affected by the large outliers in the upper tail. On the other hand, for the ACER method, the extreme value prediction is realized by fitting a parametric function to the empirical exceedance rate in the tail region and the tail behavior can be accurately captured by applying a fitting procedure which puts more weight on the empirical estimates the more accurately they are. The ACER method can provide satisfactory results for the extreme ice loads prediction without sensitive to the outliers which derives from the use of a modified version of the BLUE (best linear unbiased estimator) principle.

Thirdly, based on the time series of the ice load peaks at different locations, it is found that for current ship, the extreme ice loads in the bow intermediate region are much severer than in the bow area. Even though such phenomenon cannot be explained thoroughly with current data sets, it should be considered in reliability-based design and operation of the ice-going vessels.

Furthermore, within the scope of reliability for ice-going vessel, the relationship between the extreme ice loads, fatigue damage due to the ice loads actions and the prevailing conditions, such as the ice thickness can be studied by means of probabilistic methods, which will be a future work beyond current study.

Acknowledgment

This work has been carried out within the MARTEC research project PRICE (Prediction of ice-ship-interaction for icebreaking vessels). Financial support from the Research Council of Norway (RCN) are acknowledged. Internal discussions with the project partners: Hamburger University of Technology (TUHH-M8 and TUHH-M10), Pella Sietas GmbH in Germany and Roll-Royce Marine AS in Norway are gratefully acknowledged. In addition, special thanks are given to DNV·GL for making the measurements from KV Svalbard available.

References

Chai, W., Naess, A., Leira, B.J., Bulian, G., 2016. Efficient Monte Carlo simulation and Grim effective wave model for predicting the extreme response of a vessel rolling in random head seas. Ocean Engineering 123, 191-203.

Davison, A.C., Hinkley, D.V., 1997. Bootstrap methods and their application. Cambridge university press.

Gaidai, O., Ji, C., Kalogeri, C., Gao, J., 2017. SEM-REV energy site extreme wave prediction. Renewable Energy 101, 894-899.

Gaidai, O., Storhaug, G., Naess, A., 2016. Extreme Value Statistics of Large Container Ship Roll. Journal of ship research 60 (2), 92-100.

Hahn, M., Dankowski, H., Ehlers, S., Erceg, S., Rung, T., Huisman, M., Sjöblom, H., Leira, B.J., Chai, W., 2017. Numerical Prediction of Ship-Ice Interaction: A Project Presentation, ASME 2017 36th International Conference on Ocean, Offshore and Arctic Engineering. American Society of Mechanical Engineers, pp. V008T007A002-V008T007A002.

ISSC, 2015. Arctic Technology, International Ship and offshore Structures Congress (ISSC), Committee V.6 report. , Cascais, Portugal.

Jordaan, I.J., 2001. Mechanics of ice–structure interaction. Engineering Fracture Mechanics 68 (17), 1923-1960.

Jordaan, I.J., Maes, M.A., Brown, P.W., Hermans, I.P., 1993. Probabilistic analysis of local ice pressures. Transactions-American Scociety of Mechanical Engineers Journal of Offshore Mechnics and Arctic Engineering 115, 83-83.

Karpa, O., Naess, A., 2013. Extreme value statistics of wind speed data by the ACER method. Journal of Wind Engineering and Industrial Aerodynamics 112, 1-10.

Kotilainen, M., Vanhatalo, J., Suominen, M., Kujala, P., 2017. Predicting ice-induced load amplitudes on ship bow conditional on ice thickness and ship speed in the Baltic Sea. Cold Regions Science and Technology 135, 116-126.

Kujala, P., Vuorio, J., 1986. Results and statistical analysis of ice load measurements on board icebreaker Sisu in winters 1979 to 1985. Sjöfartsverket, Winter Navigation Research Board, Helsinki, Finland.

Kujula, P., 1994. On the statistics of ice loads on ship hull in the Baltic. Helsinki University of Technology.

Leira, B., Børsheim, L., Espeland, Ø., Amdahl, J., 2009. Ice-load estimation for a ship hull based on continuous response monitoring. Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment 223 (4), 529-540.

Lensu, M., 2002. Short Term Prediction of Ice Loads Experienced by Ice Going Ships. Helsinki University of Technology.

Lubbad, R., Løset, S., 2011. A numerical model for real-time simulation of ship–ice interaction. Cold Regions Science and Technology 65 (2), 111-127.

Montgomery, D.C., Peck, E.A., Vining, G.G., 2012. Introduction to linear regression analysis. John Wiley & Sons.

Næss, A., Gaidai, O., 2009. Estimation of extreme values from sampled time series. Structural Safety 31 (4), 325-334.

Naess, A., Moan, T., 2012. Stochastic dynamics of marine structures. Cambridge University Press, Newyork.

Newland, D.E., 1993. An introduction to random vibrations, spectral & wavelet analysis. Courier Corporation.

Nyseth, H., 2006. Strain measurements on board KV Svalbard with respect to ice loading, Technical Report. Det Norske Veritas (DNV) Oslo, Norway.

Ochi, M.K., 1990. Applied probability and stochastic processes: In Engineering and Physical Sciences. Wiley-Interscience.

Ralph, F., Jordaan, I., 2013. Probabilistic methodology for design of arctic ships, ASME 2013 32nd International Conference on Ocean, Offshore and Arctic Engineering. American Society of Mechanical Engineers, pp. V006T007A010-V006T007A010.

Riska, K., 1987. On the mechanics of the ramming interaction between a ship and a massive ice floe. Publications Valtion Teknillnen Tutkimuskeskus (43).

Suominen, M., Kujala, P., 2010. Analysis of short-term ice load measurements on board MS Kemira during the winters 1987 and 1988, Series AM. Aalto University, School of Science and Technology, Aalto University, School of Science and Technology, Espoo, Finland.

Suominen, M., Kujala, P., Kotilainen, M., 2015. The Encountered Extreme Events and Predicted Maximum Ice-induced Loads on the Ship Hull in the Southern Ocean, ASME 2015 34th International Conference on Ocean, Offshore and Arctic Engineering. American Society of Mechanical Engineers, pp. V008T007A040-V008T007A040.

Suominen, M., Kujala, P., Romanoff, J., Remes, H., 2017. Influence of load length on short-term ice load statistics in full-scale. Marine Structures 52, 153-172.

Suyuthi, A., Leira, B., Riska, K., 2012. Short term extreme statistics of local ice loads on ship hulls. Cold Regions Science and Technology 82, 130-143.

Suyuthi, A., Leira, B., Riska, K., 2013. Statistics of local ice load peaks on ship hulls. Structural Safety 40, 1-10.

Suyuthi, A., Leira, B., Riska, K., 2014. A generalized probabilistic model of ice load peaks on ship hulls in broken-ice fields. Cold Regions Science and Technology 97, 7-20.

Valanto, P., 2001. The resistance of ships in level ice. Transactions-Society of Naval Architects and Marine Engineers 109, 53-83.
Varsta, P., 1983. On the mechanics of ice load on ships in level ice in the Baltic Sea. Technical Research Center of Finland, Technical Research Center of Finland, Espoo, Finland.

Appendix: the ACER functions

According to Section 3.1, the extreme value among the N outcomes of the stochastic process $X(t)$ is defined as: $M_N = \max\{X_1,\ldots,X_N\}$ and the extreme value distribution is expressed as:

$$\begin{aligned} P(\eta) &= \mathrm{Prob}(M_N \le \eta) = \mathrm{Prob}(X_1 \le \eta,\ldots,X_N \le \eta) \\ &= \mathrm{Prob}(X_N \le \eta \,|\, X_{N-1} \le \eta,\ldots,X_1 \le \eta)\cdot \mathrm{Prob}(X_{N-1} \le \eta,\ldots,X_1 \le \eta) \\ &= \prod_{j=2}^{N} \mathrm{Prob}(X_j \le \eta \,|\, X_{j-1} \le \eta,\ldots,X_1 \le \eta)\cdot \mathrm{Prob}(X_1 \le \eta) \end{aligned} \tag{A.1}$$

Then, the cascade of conditioning approximations, i.e., a sequence of nonparametric distribution functions, is introduced to approximate the exact value distribution $P(\eta)$. The first approximation of the cascade is obtained by assuming that all variables X_j are statistically independent and then:

$$P(\eta) \approx \prod_{j=1}^{N} \mathrm{Prob}(X_j \le \eta) = \prod_{j=1}^{N} (1-\alpha_{1j}(\eta)) \tag{A.2}$$

where

$$\alpha_{1j}(\eta) = \mathrm{Prob}(X_j > \eta) \tag{A.3}$$

and

$$P(\eta) \approx P_1(\eta) = \exp\left(-\sum_{j=1}^{N} \alpha_{1j}(\eta)\right) \tag{A.4}$$

where $P_1(\eta)$ is defined by the last equality in Equation A.4. Note that the approximation $\exp(-\zeta) \approx 1-\zeta$ is accurate to within 0.5% for values of $|\zeta|$ as high as 0.1 and the accuracy is rapidly increasing for decreasing values of $|\zeta|$.

In general, the variables X_j are statistically dependent. In such case, a one-step memory approximation given by Equation A.5 is introduced:

$$\mathrm{Prob}(X_j \le \eta \,|\, X_{j-1} \le \eta,\ldots,X_1 \le \eta) \approx \mathrm{Prob}(X_j \le \eta \,|\, X_{j-1} \le \eta),\ j = 2,\ldots,N \tag{A.5}$$

Therefore, by this approximation:

$$P(\eta) \approx \prod_{j=2}^{N} \mathrm{Prob}\left(X_j \le \eta \middle| X_{j-1} \le \eta\right) \cdot \mathrm{Prob}\left(X_1 \le \eta\right) = \prod_{j=2}^{N} (1-\alpha_{2j}(\eta)) \cdot (1-\alpha_{11}(\eta))$$
$$\approx P_2(\eta) = \exp\left(-\sum_{j=2}^{N} \alpha_{2j}(\eta) - \alpha_{11}(\eta)\right) \quad \text{(A.6)}$$

where $\alpha_{2j}(\eta) = \mathrm{Prob}\left(X_j > \eta \middle| X_{j-1} \le \eta\right)$ for $j=2, \ldots, N$.

In order to convey a clear picture of the cascade of conditioning approximations, the third level of approximation is achieved by assuming that:

$$\mathrm{Prob}\left(X_j \le \eta \middle| X_{j-1} \le \eta, \ldots, X_1 \le \eta\right) \approx \mathrm{Prob}\left(X_j \le \eta \middle| X_{j-1} \le \eta, X_{j-2} \le \eta\right), j = 3, ..., N \quad \text{(A.7)}$$

By introducing the approximation A.7, it is obtained that:

$$P(\eta) \approx \prod_{j=3}^{N} \mathrm{Prob}\left(X_j \le \eta \middle| X_{j-1} \le \eta, X_{j-1} \le \eta\right) \cdot \mathrm{Prob}\left(X_2 \le \eta \middle| X_1 \le \eta\right) \cdot \mathrm{Prob}\left(X_1 \le \eta\right)$$
$$= \prod_{j=3}^{N} (1-\alpha_{3j}(\eta)) \cdot (1-\alpha_{22}(\eta)) \cdot (1-\alpha_{11}(\eta)) \quad \text{(A.8)}$$
$$\approx P_3(\eta) = \exp\left(-\sum_{j=3}^{N} \alpha_{3j}(\eta) - \alpha_{22}(\eta) - \alpha_{11}(\eta)\right)$$

in which the notation $\alpha_{3j}(\eta) = \mathrm{Prob}\left(X_j > \eta \middle| X_{j-1} \le \eta, X_{j-2} \le \eta\right)$ for $j=3, \ldots, N$ is introduced.

It is realized that, a general kth approximation will be obtained by continuing the above conditioning process. By introducing the notation $\alpha_{kj}(\eta)=\mathrm{Prob}(X_j > \eta | X_{j-1} \le \eta, \ldots, X_{j-1+k} \le \eta)$, $1 \le k \le j \le N$, we can obtain that:

$$P(\eta) \approx P_k(\eta) = \exp\left(-\sum_{j=k}^{N} \alpha_{kj}(\eta) - \sum_{j=1}^{k-1} \alpha_{jj}(\eta)\right), k \ge 2 \quad \text{(A.9)}$$

where $\alpha_{kj}(\eta)$ represents the probability of exceedance of X_j conditioned on k-1 immediately preceding non-exceedances. Thereby, the cascade $\{P_k(\eta)\}_{k=1}^{N}$ of conditional probability of conditional probability distributions has been constructed to approximate the exact value distribution $P(\eta)$.

For most practical applications, $N \gg k$, and $\sum_{j=1}^{k-1} \alpha_{jj}(\eta)$ is effectively negligible compared to $\sum_{j=k}^{N} \alpha_{kj}(\eta)$. Therefore,

$$P_k(\eta) \approx \exp\left(-\sum_{j=k}^{N} \alpha_{kj}(\eta)\right), k \geq 1 \tag{A.10}$$

then, estimation of the extreme value distribution by the described conditioning approach reduces to estimation of the set of $\alpha_{kj}(\eta)$ functions (Næss and Gaidai, 2009).

The next step is to introduce the concept of average conditional exceedance rate of order k (i.e. the ACER function) as follows:

$$\varepsilon_k(\eta) = \frac{1}{N-k+1} \sum_{j=k}^{N} \alpha_{kj}(\eta), k = 1, 2, \ldots \tag{A.11}$$

and in general, this ACER function also depends on the number of data points N.

The numerical estimation of the ACER functions is based on counting the requisite events, and we start by introducing the following random functions:

$$\begin{aligned} A_{kj}(\eta) &= \mathbf{1}\left\{X_j > \eta, X_{j-1} \leq \eta, \ldots, X_{j-k+1} \leq \eta\right\}, \; j = k, ..., N; k = 2, 3, \ldots \\ B_{kj}(\eta) &= \mathbf{1}\left\{X_{j-1} \leq \eta, \ldots, X_{j-k+1} \leq \eta\right\}, \; j = k, ..., N; k = 2, 3, \ldots \end{aligned} \tag{A.12}$$

where $\mathbf{1}(\Theta) = 1$ if the event Θ is true, otherwise the indicator function will be zero. Then

$$\alpha_{kj}(\eta) = \frac{\mathrm{E}\left[A_{kj}(\eta)\right]}{\mathrm{E}\left[B_{kj}(\eta)\right]}, \; j = k, \ldots, N; k = 2, 3, \ldots, \tag{A.13}$$

where $\mathrm{E}[\cdot]$ denotes the expectation operator.

Assuming that the random process $X(t)$ is ergodic, then obviously, $\varepsilon_k(\eta) = \alpha_{kk}(\eta) = \ldots = \alpha_{kN}(\eta)$ and it may be assumed that for the time series at hand,

$$\varepsilon_k(\eta) = \lim_{N \to \infty} \frac{\sum_{j=k}^{N} a_{kj}(\eta)}{\sum_{j=k}^{N} b_{kj}(\eta)} \tag{A.14}$$

in which $a_{kj}(\eta)$ and $b_{kj}(\eta)$ are realized values of $A_{kj}(\eta)$ and $B_{kj}(\eta)$, respectively, for the observed time series.

It is clear that $\lim_{\eta \to \infty} \sum_{j=k}^{N} b_{jk}(\eta) = N - k + 1 \approx N$. Hence $\lim_{\eta \to \infty} \tilde{\varepsilon}_k(\eta) / \varepsilon_k(\eta) = 1$, where

$$\tilde{\varepsilon}_k(\eta) = \lim_{N \to \infty} \frac{\sum_{j=k}^{N} a_{kj}(\eta)}{N-k+1} \tag{A.15}$$

The advantage of using the modified ACER function $\tilde{\varepsilon}_k(\eta)$ for $k \geq 2$ is that it is easier to use for non-stationary or long-term statistics than $\varepsilon_k(\eta)$. Because our focus is on the values of the ACER functions at the extreme levels, we may use any function that provides correct estimates of the ACER function at the extreme levels.

To see why Equation A.15 may be applicable for non-stationary time series, it is recognized that

$$P_k(\eta) \approx \exp\left(-\sum_{j=k}^{N} \alpha_{kj}(\eta)\right) = \exp\left(-\sum_{j=k}^{N} \frac{\mathrm{E}\left[A_{kj}(\eta)\right]}{\mathrm{E}\left[B_{kj}(\eta)\right]}\right)$$
$$\approx \exp\left(-\sum_{j=k}^{N} \mathrm{E}\left[A_{kj}(\eta)\right]\right), \eta \to \infty. \tag{A.16}$$

If the time series can be segmented into L blocks such that $\mathrm{E}[A_{kj}(\eta)]$ remains approximately constant within each block and such that $\sum_{j\in C_l} \mathrm{E}[A_{kj}(\eta)] \approx \sum_{j\in C_l} a_{kj}(\eta)$ for a sufficient range of η-values, where C_l denotes the set of indices for the data in block number l, l = 1, …, L, then $\sum_{j=k}^{N} \mathrm{E}[A_{kj}(\eta)] \approx \sum_{j=k}^{N} a_{kj}(\eta)$. Hence,

$$P_k(\eta) \approx \exp\left(-(N-k+1)\hat{\varepsilon}_k(\eta)\right) \tag{A.17}$$

where

$$\hat{\varepsilon}_k(\eta) = \frac{1}{N-k+1} \sum_{j=k}^{N} a_{kj}(\eta) \tag{A.18}$$

in which $\hat{\varepsilon}_k(\eta)$ is the empirical ACER function of order k, within $k \leq N$ and N is the total number of local peaks in time series. It should be noted that for a given k and a specified level η, it follows from Equation A.18 that it is only required to run through the whole time series and check for each $j = k, \ldots, N$ whether $a_{kj}(\eta)$ is zero or one.

As for the problem of estimating a confidence interval (CI) for $\tilde{\varepsilon}_k(\eta)$, given a general non-stationary process, it is consistent to assume that the stream of conditional exceedances over a threshold constitute a Poisson process, possibly non-homogeneous. Hence, 95% CI is given by the following expression (Gaidai et al., 2016; Næss and Gaidai, 2009):

$$CI_{\pm}(\eta) = \hat{\varepsilon}_k(\eta)\left(1 \pm \frac{1.96}{\sqrt{(N-k+1)\hat{\varepsilon}_k(\eta)}}\right) \tag{A.19}$$

Research highlights

- A novel method, i.e. the ACER method, is presented for estimating the extreme ice loads
- Main principle of the ACER method is introduced
- Methods based on the classic extreme value theory is applied for benchmark study
- Extreme ice loads on different locations are studied