Contents lists available at ScienceDirect

Ocean Engineering

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

Highlights

Assessment of methods for short-term analysis of riser collision probability

Ocean Engineering xxx (xxxx) xxx

Ping Fu*, Bernt J. Leira, Dag Myrhaug, Wei Chai

- Provide a comprehensive assessment of various methods for flexible riser collision probability.
- Consider the wake effect, which is generated by the upstream riser and acts on the downstream riser.
- The importance of the threshold value is emphasized for better modeling the tail of the distribution and the estimation of the extreme value.

Contents lists available at ScienceDirect

Ocean Engineering

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

Assessment of methods for short-term analysis of riser collision probability

Ping Fu *, Bernt J. Leira, Dag Myrhaug, Wei Chai

Department of Marine Technology, Norwegian University of Science and Technology, NO-7491, Trondheim, Norway

ARTICLE INFO

Keywords:
Riser collision probability
Riser clusters
Extreme value analysis
Current and waves
Wake effect

ABSTRACT

Reliability- and performance-based design of a riser cluster requires an accurate estimation of the extreme values of the stochastic relative distance between the risers during a given time duration. This study presents a comprehensive assessment of six methods for estimating the collision probability between two risers with limited simulation length. A pair of steep-wave risers in tandem arrangement subjected to combined current and wave loads is modeled. The wake effect generated by the upstream riser acting on the downstream riser is considered.

Firstly, the critical locations at which collisions during a short duration are likely to occur are identified. After obtaining nonlinear riser responses, 3-hour short-term extreme relative motions are calculated, which are used for the collision probability estimation. Several methods for the extreme value analysis, including the Gumbel probability paper method, the general extreme value method and the average conditional exceedance rate method are presented. This paper also proposes two methods which offer reliable and satisfactory results for the extreme value analysis of highly skewed processes.

1. Introduction

As the offshore industry moves to deeper water, risers are commonly arranged as clusters with small spacing due to limited size of the platform as well as cost considerations. The dynamic response of riser clusters, induced by waves, currents and platform motions, becomes complex, and this complexity is increased by the arrangement of risers. The influence of neighboring risers may lead to hydrodynamic loads on the individual risers, which are significantly different from the loads they would experience in isolation. This hydrodynamic interference may cause large relative motions between risers, resulting in potential collisions (Rustad et al., 2008). For instance, when a downstream cylinder is placed within the wake field generated by an upstream cylinder, the flow around the downstream cylinder will be changed due to the presence of the upstream one. This change will reduce the drag force acting on the downstream cylinder, and it also might induce an additional lift force if the cylinders are in a staggered arrangement. This wake effect, in turn, reduces the clearance between the cylinders. A literature review on the topic of wake induced oscillation can be found in Fu et al. (2015). Furthermore, the differences in excitation force on neighboring risers may cause relative motions to be large, resulting in the small clearance between risers. Additionally, the floater motions also affect the motions of the individual risers. All these factors may lead to the possibility of riser collision, especially when risers are subjected to a severe sea state. However, relatively few studies have been carried out regarding riser collision considering the wake effect.

Due to the high degree of non-linearity, time domain analysis is required for the collision assessment problem. However, it is impractical to simulate collision events because of the limitations of the simulation time and the complication of the contact problem. Firstly, collision is a rare event in practice, so for a well-designed riser system extremely long simulation periods are usually required for such an event to occur. Secondly, when a clash between risers does occur, the motions of individual risers will both be changed, which are different from when ignoring the presence of the other riser. However, simulating the contact and the motions after contact is still a challenge for most of the available software (and in general for global dynamic analysis computer algorithms). Thus, the riser collision process, represented by the normalized minimum distance between risers is formulated as the main quantity of interest. Since this is a random process, application of extreme value analysis is required. Many techniques for estimation of the extreme value distribution have been proposed including the Gumbel probability paper (GPP) method (Gumbel, 1958), the generalized extreme value (GEV) method, the average conditional exceedance rate (ACER) method (Næss and Gaidai, 2009; Naess and Karpa, 2013; Karpa and Naess, 2013) and the method based on the translation process model (Winterstein, 1987, 1988; Ding and Chen, 2014). The accuracy

* Corresponding author.
E-mail address: ping.fu@ntnu.no (P. Fu).

https://doi.org/10.1016/j.oceaneng.2021.109221
Received 9 June 2020; Received in revised form 10 May 2021; Accepted 22 May 2021

and uncertainty associated with these methods are affected by the properties of the process and the modeling adequacy of these properties based on limited data.

The generalized extreme value method requires a number of time series of equal length, and that the maximum value from each time history is extracted. With the assumption that these maxima are stochastically independent and identically distributed, the extreme value converges to one of three general distribution functions, i.e. Gumbel, Weibull or Frechet distributions (Gumbel, 1958). This method requires a large number of simulations in order to obtain a result with acceptable accuracy. Therefore, an alternative method for the estimation of the extreme value distribution based on a limited number of simulations is introduced. The ACER method expresses the mean level up-crossing rate at the upper tail as a function of the level η. This makes it possible to avoid the commonly adopted assumption that the extreme value data follow an appropriate asymptotic extreme value distribution (Næss and Gaidai, 2009). The translation process method deals with non-Gaussian processes. It translates non-Gaussian process to standard Gaussian process through a monotonic translation function, so that the traditional extreme theory for the Gaussian process can be applied. The translation function can be determined using the statistical moments of the non-Gaussian process, i.e. the moment-based Hermite method (Duggal and Niedzwecki, 1994). This method was used as the probabilistic collision model for a pair of top-tension riser for a given location (Duggal and Niedzwecki, 1993). This paper introduces another approach to calculate the translation function, by mapping the CDF of the non-Gaussian process data to a prescribed parent distribution model. However, in the literature, few studies have been carried out for the collision probabilities (He and Low, 2013; Leira et al., 2002), but most of them do not focus on the extreme value problem.

The purpose of this paper is to evaluate the performances of the Gumbel probability paper method, the generalized extreme value method, the ACER method and two proposed methods in the application of the riser collision probability estimation. Particular attention is given to a pair of flexible wave-risers in tandem arrangement subjected to a combined current and waves flow. The wake effect generated by the upstream riser is considered. The shortest relative distance between the risers is searched for according to the riser motion for a given time duration. The collision probability is then estimated with short-term time history samples for the different extreme value analysis methods.

2. Extreme value analysis

2.1. Generalized extreme value distribution

The largest maximum for each time history is the highest value among all the local maxima,

$$X_e = \max\{X_{m_1}, X_{m_2}, \dots, X_{m_n}\}. \tag{1}$$

where X_e and X_{m_i}, $i = 1...n$ represent the largest maximum and the individual local maxima from a given time series, respectively. By assuming that all the individual maxima are stochastically independent and identically distributed with a common distribution function $F_{X_m}(x)$, the distribution of X_e is given as (Bury, 1975):

$$F_{X_e}(x) = P\{X_e \le x\} = [F_{X_m}(x)]^n \tag{2}$$

It is proven that this equation will normally converge towards one of three types of extreme value distributions as $n \to \infty$, i.e. Gumbel, Weibull or Frechet (Bury, 1975). These three types of extreme value distributions have a common form, i.e. the generalized extreme value (GEV) distribution, given as:

$$F_{Xe}(x) = \exp\left\{-1 + \gamma\left(\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\gamma}}\right\} \tag{3}$$

where μ is the location parameter, σ is the scale parameter, and γ is the shape parameter. The shape parameter strictly affects the shape of the distribution, and governs the tail of each distribution. The shape parameter is derived from the skewness, as it represents where the majority of the data lies. Here $\gamma > 0$ corresponds to the Frechet distribution and $\gamma < 0$ corresponds to the reversed Weibull distribution. Note that the reversed Weibull distribution is the only type of the extreme value distribution with a finite upper limit. Furthermore, $\gamma = 0$ is the limiting case when $\gamma \to 0$, which leads to the Gumbel distribution. In general, a distribution with a larger number of fitting parameters will be able to model the input data more accurately than a distribution with a smaller number of parameters. However, the Gumbel distribution is efficient for small sample sizes. If the size is greater than 50, GEV shows a better overall performance (Cunnane, 1989).

2.2. Gumbel probability paper method

The Gumbel model is the most commonly used distribution model for marine structures. The cumulative distribution function (CDF) of the Gumbel model is given as

$$F_{X_e}(x) = \exp\{-\exp\{-\alpha(x-\mu)\}\} \tag{4}$$

where α is the scale parameter, and μ is the location parameter. The parameters can be determined by fitting a straight line to the data in a Gumbel probability paper (GPP). Details can be found by Naess and Moan (2012).

By taking the logarithm of both the left and right hand side of this equation twice, the following equation is obtained:

$$-\ln[-\ln(F_{X_e}(x))] = \alpha(x-\mu) \tag{5}$$

Further, by introducing $y = -\ln[-\ln(F_{X_e}(x))]$ a linear function, i.e. $y = \alpha(x-\mu)$ is obtained, which implies that in a x–y axis system, the cumulative distribution becomes a straight line. Then, the parameters α and μ can be estimated by least-square fitting of the samples to the straight line.

2.3. Average conditional exceedance rate

Unlike the above mentioned methods based on the parametric distribution functions, the ACER method estimates the extreme value distribution by constructing a sequence of non-parametric distribution functions, i.e. the ACER functions (Naess and Gaidai, 2008; Næss and Gaidai, 2009; Karpa and Naess, 2013; Gaidai et al., 2016, 2018; Chai et al., 2018). Here the basic idea of the ACER method is summarized as following.

Based on the time series of the individual maxima, the extreme value can be expressed as:

$$F_{X_e}(x) \approx P_k(x) \approx \exp\{-(n-k+1)\hat{\epsilon}(x)\} \tag{6}$$

where n is the counted number of maxima during a given time duration T_0; k is the order of the ACER function; P_k is the approximation of the extreme value distribution based on the kth order ACER function, and $\hat{\epsilon}(x)$ is the empirical ACER function of order k, which can be determined by applying the existing time series. As k increases (more complex model), the bias of Eq. (6) will generally reduce, but variance will increase.

In order to predict the extreme value distribution in the upper tail region, an extrapolation scheme is applied. Specifically, in the upper tail region (e.g. for $x \ge x_{m0}$, where $x \ge x_{m0}$ is an appropriately chosen tail marker), the ACER function behaves similarly to the function $\exp\{-a(x-b)^c\}$, where $a > 0$, $b \le x_{m0}$ and $c > 0$ are suitable constants. The empirical ACER function is assumed to have the form:

$$\hat{\epsilon}_k(x) = q_k \exp\left\{-a_k(x-b_k)^{c_k}\right\}; \quad x \ge x_{m0} \tag{7}$$

where a_k, b_k, c_k and q_k are suitable constants which depend on the order k. It should be noted that Eq. (7) is applicable in the upper tail region, i.e., $x \ge x_{m0}$. By comparing the empirical $\hat{\epsilon}_k(x)$ for different values

of k, an appropriate value of k is selected to capture the dependence structure of the time series.

The constants a_k, b_k, c_k and q_k can be determined by minimizing the following mean-square-error function:

$$F(q_k, a_k, b_k, c_k) = \sum_{i=1}^{N} \rho_j |\ln \hat{\epsilon}_k(x_{mi}) - \ln q + a(x_{mi} - b)^c|^2 \tag{8}$$

where x_{mi}, $i = 1, \ldots, n$ are levels at which the ACER functions have been empirically estimated. The weight factor ρ_j is given by the relationship $\rho_j = (\ln CI^+(x_{mi}) - \ln CI^-(x_{mi}))^{-2}$, where CI represents the 95% confidence interval, which can be approximately expressed as:

$$CI^{\pm}(x_{mi}) = \hat{\epsilon}_k(x_{mi})\{1 \pm \frac{1.96}{\sqrt{(n-k+1)\hat{\epsilon}_k(x_{mi})}}\} \tag{9}$$

Therefore, it is seen in Eqs. (8) and (9) that the weight factor ρ_j decreases as the level x_{mi} increases, which implies that the extrapolation scheme puts more emphasis on the more reliable data points. Moreover, it should be noted that there is a level x_{mi} beyond which the weight factor ρ_j is no longer defined since the CI^- estimated by Eq. (9) would be negative as the levels exceed x_{mi}.

2.4. Proposed method 1: Gumbel parameters based on Weibull tail fitting of maxima

The requirement related to a large number of simulations for the GPP and the GEV methods has led to the development of alternative methods for the estimation of the extreme value distribution from a small number of simulations, so that the individual maxima from the limited time history can be used for the extreme value estimation. One possible model is that, if the individual maxima follow the 3-parameter Weibull distribution, the extreme value which is followed by the Gumbel distribution can be determined based on the distribution parameters of the Weibull model.

The 3-parameter Weibull distribution function commonly used as a model for individual response maxima is given as

$$F_{X_m}(x) = 1 - \exp\left\{-\left(\frac{x-\mu}{\sigma}\right)^{\lambda}\right\} \tag{10}$$

where μ, σ and λ are the location, scale and shape parameters, respectively. As recommended, those three parameters are determined by the moment method (Farnes and Moan, 1993). It is also recommended that only the global maxima, i.e. the largest maxima between positive slope up-crossings of the threshold should be selected. This method reduces the potential correlation between selected maxima and eliminates the majority of smaller amplitude extremes that are less significant for estimation of extreme values distribution. Fig. 1 illustrates the global maxima which are marked by red color, i.e. X_m.

After obtaining the Weibull parameters, the extreme value which is assumed to follow the Gumbel distribution has the approximated scale and location parameters (Bury, 1975):

$$\begin{aligned} &\mu_{x_e} + \frac{0.57722}{\alpha_{x_e}} = u + \sigma\left\{(\ln n)^{\frac{1}{\lambda}} + \frac{0.57722}{\mu}(\ln n)^{\frac{1-\lambda}{\lambda}}\right\} \\ &\alpha_{x_e} = \frac{\lambda}{\sigma}(\ln n)^{\frac{1-\lambda}{\lambda}} \end{aligned} \tag{11}$$

where n is the number of global maxima for a given time duration; μ_{x_e} and α_{x_e} are the location and scale parameters of the Gumbel distribution function, respectively; μ, σ and λ are Weibull distribution parameters, as given in Eq. (10).

Since the aim of the fitting is to obtain a Weibull model to be used for estimation of extremes, it is important that the fitting procedure gives a good fit to the upper tail data. Therefore, a threshold value is used so that the effect of the lower tail data (below the threshold) can be neglected. The idea of only considering observations above a threshold value for estimation of parameters in a proposed distribution model has been applied for extreme value prediction of the response of a single flexible riser (Sodahl, 1991). In the present study, the Weibull tail fitting method will be used.

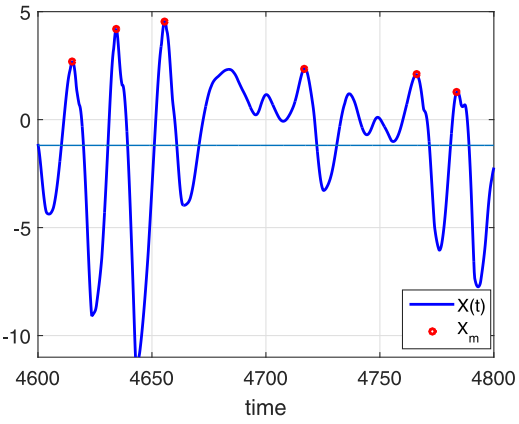


Fig. 1. Definition of global maxima. X_m: individual global maxima. X(t): realization of a stochastic process. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The choice of threshold is based on experience. In this paper, the method for determining the threshold is based on the mean and standard deviation of the original time series. For instance, for a given 3-h simulation, the smallest threshold is set to be the mean value, e.g. $x_0 = \mu_x$, and largest threshold is to be the 1.4 times the standard deviation, e.g. $x_0 = \mu_x + 1.4STD$, based on the recommendations by Moriarty et al. (2004). Other threshold values are inserted for refinement; totally, five threshold values are applied, i.e. $x_0 = \mu_x + \eta_x$, where $\eta_x = [0, 0.5, 1, 1.2, 1.4] * STD$.

2.5. Proposed method 2: Translation process method based on CDF mapping

The basic idea of the translation process method is to translate the non-Gaussian process to a standard Gaussian process, so that the extreme value distribution of the non-Gaussian process can be estimated from that of the standard Gaussian process. The translation function is given by,

$$\begin{aligned} &F_X(x) = \Phi(u) \\ &x = g(u) = F_X^{-1}[\Phi(u)] \end{aligned} \tag{12}$$

where u is the standard Gaussian process; $\Phi(u)$ is the corresponding CDF; $g(\cdot)$ is the translation function from the X-space to the U-space, and F_X^{-1} is the inverse function of F_X. The extreme value distribution is then calculated via the Gaussian extreme value theory,

$$\begin{aligned} F_{X_e}(x) &\approx \exp\left\{-\nu^+(x)T\right\} \\ &\approx \exp\left\{-\nu_0^+ T \exp\left\{-\frac{g^{-1}(x)^2}{2}\right\}\right\} \end{aligned} \tag{13}$$

where $\nu^+(x)$ is the up-crossing rate at a threshold level x, which is identical to the mean up-crossing rate in the U-space at level $u = g^{-1}(x) = \Phi^{-1}\{F_X(x)\}$ since the translation function is monotonic; ν_0 is the zero up-crossing rate; T is the time duration and $g^{-1}(x) = \Phi^{-1}F_X(x)$ is the inverse translation function.

Therefore, an appropriate distribution model $F_X(x)$ which is able to capture the upper tail well is essential. In this section, a combined distribution model is used so that the behavior of the upper tails can be captured. The idea of using a combined distribution model for CDF mapping has been applied by, e.g., Peng et al. (2014). In this section, a different combined distribution model for the upper tail fitting is presented.

An empirical distribution based on the measured data, i.e. a kernel sampling density, is applied in the lower tail probability region, and a 3-parameter Weibull distribution is applied in the upper tail. The details

Table 1
Riser and buoyancy element properties.

	Unit	Bare riser	Buoyancy modules
Outside diameter	[m]	0.25	0.63
Inside diameter	[m]	0.05	0.05
Mass density	[kg/m]	100	100
EI	[kNm^2]	104	104
Content density	[kg/m^3]	1000	1000
Length	[m]	110	50

of the kernel sampling density is not given here since this paper is focused on analysis of the extreme value distribution which is governed by the upper tail. The expression of the CDF of the upper tail data larger than the tail marker x_0 followed by the Weibull distribution is given as:

$$F_{3wb}(X \le x | X \ge x_0) = \frac{F_X(x) - F_X(x_0)}{1 - F_X(x_0)} \tag{14}$$

where F_{3wb} is the CDF of 3-parameter Weibull distribution, as given in Eq. (10), and x_0 is the tail marker(or threshold), and corresponds to the location parameter μ in Eq. (10). The CDF of the maxima in the upper tail is then calculated by,

$$F_X(x) = F_{X_{kel}}(x_0) + (1 - F_{X_{kel}}(x_0))F_{3wb}(x), \quad x \ge x_0 \tag{15}$$

where $F_{X_{kel}}(x)$ is the CDF of the kernel sampling distribution which is applied for the data smaller than the tail marker x_0. Moreover, $F_X(x_0) = F_{X_{kel}}(x_0)$, ensuring that the CDF is continuous at $x = x_0$. After determining the distribution model $F_X(x)$, the translation function is established by CDF mapping, i.e. Eq. (12).

3. Riser system modeling

Before further assessment, it is necessary to address some challenges related to conducting the riser collision analysis. Firstly, the wake effect significantly influences the relative distance between risers. However, application of the wake models in the finite element software is still a challenge. Secondly, even though there are many existing methods for extreme value estimation, it is still a challenge to find a suitable definition of a random variable to describe the stochastic properties of the riser collision problem. In this section, these aspects will be addressed by examining a pair of risers in a steep-wave configuration subjected to combined current plus waves.

Time domain simulation is necessary for calculating the non-linear riser response under combined current and wave loads. Especially, the non-linearity is increased when the wake effect is taken into account. The wake effect is considered by combining the finite element software Riflex (1987) along with the Blevins wake model (Blevins, 2005). The Riflex is specially designed to handle static and dynamic analyses of slender marine structures. The Blevins wake model expresses the drag and lift forces on the downstream riser as a function of the relative distance between risers.

3.1. Description of riser system

The semi-submersible floater is modeled as a rigid body, and the motion of the floater is specified through the linear motion transfer functions in 6 degrees-of-freedom (DOFs). Amplitudes and phase angles at the center of gravity are presented in Fig. 2 for the surge motion.

The risers in the steep-wave configuration are illustrated in Fig. 3. Each of the risers have a total length of 160 m with a diameter of 0.25 m. The buoyancy elements are attached along the riser over a length of 50 m with a diameter of 0.63 m starting 10 m from the lower end. The main riser properties are summarized in Table 1. The initial relative distance between the top-ends of riser is $L_0 = 10$ m. The water depth is 100 m.

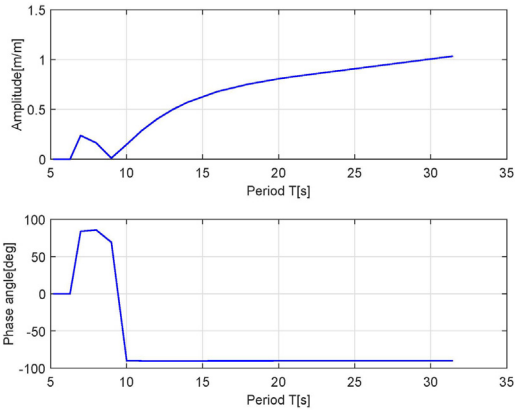


Fig. 2. First order motion transfer function amplitude and phase angle for surge.

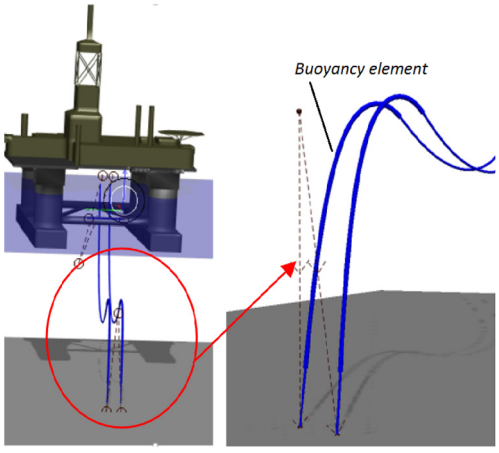


Fig. 3. Riser configuration.

3.2. Environmental conditions

The wave condition is described by a JONSWAP wave spectrum with a peak factor of $\eta = 3.3$, a significant wave height $H_s = 14$ m and a peak period $T_p = 18$ s. In addition, a linearly decreasing current with surface velocity $V_c = 1$ m/s and seabed velocity $V_c = 0.8$ m/s has been included. Both waves and current are perpendicular to the risers.

3.3. Hydrodynamic forces

In Riflex, the hydrodynamic forces are calculated based on two-dimensional strip theory. The wave-induced excitation forces (Froude–Krylov and diffraction forces) are computed by a long wavelength approximation which involves added mass and potential damping of the actual cross section together with the wave kinematics. For a slender structure, the main component of the hydrodynamic forces is the viscous drag force. The viscous load is computed using the drag term in the modified Morison equation, taking into account the relative motion between the riser and the fluid flow. The Morison equation per unit length is expressed as:

$$\begin{aligned} F(t) = {} & \frac{1}{2} C_D D(v(t) - \dot{x}(t))|v(t) - \dot{x}(t)| \\ & + \rho C_M \frac{\pi}{4} D^2 \dot{v}(t) - \rho(C_M - 1)\frac{\pi}{4} D^2 \ddot{x}(t) \end{aligned} \tag{16}$$

where ρ is the water density; C_D and C_M are the drag and inertia coefficients, respectively; v and $\dot{v}$ are the water fluid velocity and acceleration, respectively; $\dot{x}$ and $\ddot{x}$ are the riser velocity and acceleration, respectively. For current plus wave flow, v is given as a superposition of current and wave orbital velocities.

3.4. Wake model

The wake effect generated by the upstream riser needs to be accounted for. Usually, the wake effect is translated into a reduced flow velocity over the downstream riser. In this study, the wake model developed by Blevins (2005) is used, by which the reduction of the local flow velocity over the downstream cylinder is transformed to the reduction of the drag coefficient. The formula is given as,

$$C_D(L,T) = C_{D0}\left\{1 - k_1\left(\sqrt{\frac{C_{D_0}D_u}{L}}\right)\exp\left(\frac{-k_2T^2}{C_{D_0}D_uL}\right)\right\}^2 \tag{17}$$

where (L,T) is the longitude and transverse locations of the downstream cylinder with respect to the upstream cylinder; C_D is the downstream cylinder drag coefficient based on the local flow velocity; C_{D0} is the reference drag coefficient based on the undisturbed flow velocity; D_u is the diameter of the upstream cylinder; the parameters $k_1 = 1$ and $k_2 = 4.5$ are constants, and determined by fitting curve to the experimental data at $L/D = 3$, 5, 9 and 20.3 using the least-squares method (Price and Paidoussis, 1984). However, more data is required in order to validate this model.

The Blevins model also comprises an inward lift force on the downstream cylinder, towards the wake center-line. The lift force is proportional to the transverse gradient of the drag force, according to the Rawlins' postulate. In the present study, however, the downstream riser is placed at the wake center-line so that the lift force caused by the asymmetry flow can be neglected. Actually, the Blevins model is valid for a relative distance between the centers of cylinders larger than 2–3 times the diameter of the upstream cylinder. At distances less than this value, the interference becomes more complex with a negative suction force involved. More details about the lift force induced by the wake effect are given in Blevins (2005).

For the inertia force, a typical value of 2 is used for the inertia coefficient, and it is considered independent of the relative distance. The upstream riser is considered as a single and isolated cylinder, and the hydrodynamic force is also considered to be independent of the relative distance.

3.5. Time domain simulation

The time domain simulations are conducted by using the commercial finite element analysis software Riflex, where each riser is modeled by means of line elements. Riflex is specially designed to handle static and dynamic analyses of flexible risers and other slender structures. The static analysis methods comprise catenary analysis, available for a limited range of systems. The dynamic analysis methods comprise linear and nonlinear time domain analysis.

The wake effect in steady current is firstly considered in the static analysis. Since the drag force acting on the downstream riser depends on the response, i.e. the relative distance between risers, an iteration procedure is necessary for determining the final static equilibrium deflection. However, Riflex does not have the functionality to give a varying $C_D(x)$ as input, and therefore such a representation is achieved by combining Riflex with an automatic workflow developed by the author. Firstly, Riflex runs a static analysis based on the initial distance and calculates the relative distance between two segments. Then the shortest relative distance x and the associated $C_D(x)$ is computed. The workflow generates new input files and automatically calls Riflex for a new analysis. This procedure is repeated until a converge is achieved. The risers' equilibrium static position is then determined. More details about the implementation of the wake model in the static analysis can be found in Fu et al. (2017).

The 3-h dynamic responses of the risers have been simulated 50 times with different random seeds for generating time series of waves. When computing the dynamic response of interacting risers, the calculated C_D for the downstream riser obtained from the static analysis is applied in the dynamic analysis.

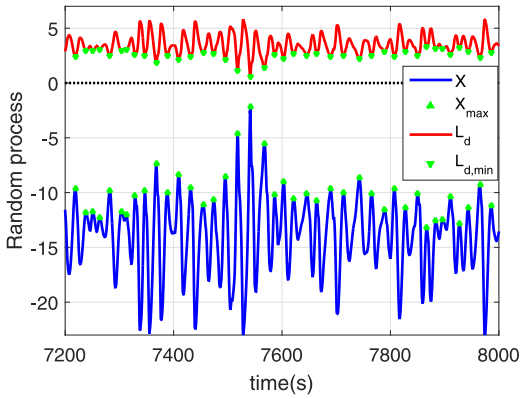


Fig. 4. Time series of the process X and minimum relative distance L. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Statistics of the process X.

	S1		S2		S3	
	mean	CoV	mean	CoV	mean	CoV
Mean	−15.53	0.20%	**−13.80**	0.26%	−15.55	0.21%
STD	2.88	1.00%	2.30	0.97%	**3.46**	1.07%
Skewness	−1.10	3.61%	−0.87	3.63%	−0.73	3.79%
Kurtosis	4.22	3.67%	3.46	2.82%	3.16	2.60%

3.6. Definition of random process

The relative distance is essential for estimating the riser collision probability. The basic principle for calculating the relative distance is that each riser is modeled as a line represented by a series of line segments, according to their properties and geometries. Each segment comprises a series of beam elements. For each pair of segments, the minimum distance can be checked element-by-element at each time step. This check is carried out for all combinations of elements until the point with the smallest distance is detected within the segments.

This calculated minimum relative distance is taken as the random process, which can be used further for extreme value analysis. However, it is convenient to transform the extreme minimum value problem to the non-dimensional extreme maximum value problem by introducing the following process:

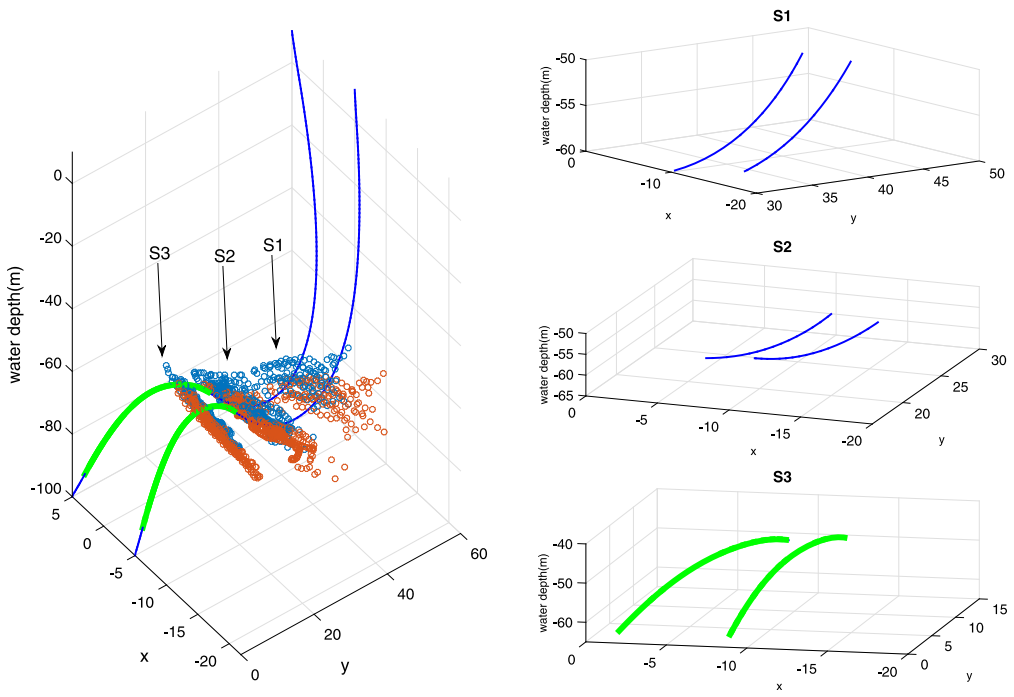
$$X(t) = -L(t)/D \tag{18}$$

where $L(t)$ is the minimum center-to-center distance between two critical segments at time t. In this case, $X(t) < 0$, and the risers physically collide when $X(t) = -1$. Fig. 4 compares the time history of $X(t)$ with the corresponding maxima and $L(t)$ with the corresponding minima.

3.7. Identification of critical segments

As mentioned, the shortest distance between each pair of segments is checked element-by-element at each time step. This check is carried out for all combinations of elements until the point of smallest distance is detected within the segments, which is very time consuming, especially if the simulation duration is long and the number of elements is large. This procedure also implies that the exact location of shortest distance might change at each time step. However, it should be noted that some element combinations may not be relevant for calculation of the shortest distance. Therefore, some critical segments where the collisions are likely to occur can initially be identified.

A 10-min simulation is preformed in order to find the critical segments along the risers. The minimum distance is checked element-by-element in Riflex (SINTEF, 2017). The locations where the minimum distance occurs are recorded, as illustrated in Fig. 5. The blue



(a) Critical location. (b) Critical segments.

Fig. 5. Identification of critical segments. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
Statistics of the global maxima X_m.

S1	mean	CoV	STD	CoV	skewness	CoV	kurtosis	CoV
mean	−11.87	0.46%	1.57	3.85%	0.48	37.25%	3.43	15.45%
mean + 0.5STD	−11.74	0.45%	1.44	4.03%	0.84	20.68%	3.61	19.40%
mean + 1.0STD	−10.79	0.86%	1.15	7.17%	1.22	19.91%	4.50	27.11%
mean + 1.2STD	−10.29	1.01%	1.03	9.63%	1.38	20.29%	4.95	29.26%
mean + 1.4STD	−9.76	1.24%	0.93	13.49%	1.50	22.57%	5.31	32.82%
S2	mean	CoV	STD	CoV	skewness	CoV	kurtosis	CoV
mean	−10.46	0.47%	1.32	4.20%	0.37	67.99%	3.84	21.23%
mean + 0.5STD	−10.40	0.48%	1.19	4.57%	0.87	27.89%	4.13	27.25%
mean + 1.0STD	−9.67	0.77%	0.96	8.57%	1.40	23.60%	5.53	35.20%
mean + 1.2STD	−9.21	1.05%	0.90	11.94%	1.51	24.95%	5.83	38.91%
mean + 1.4STD	−8.69	1.52%	0.85	16.02%	1.50	28.66%	5.64	43.09%
S3	mean	CoV	STD	CoV	skewness	CoV	kurtosis	CoV
mean	−12.45	0.36%	1.22	3.15%	0.18	77.60%	3.40	9.58%
mean + 0.5STD	−12.42	0.31%	1.06	2.94%	0.82	17.74%	3.65	14.63%
mean + 1.0STD	−11.65	0.54%	0.84	5.89%	1.24	15.55%	4.56	20.86%
mean + 1.2STD	−11.18	0.74%	0.76	8.17%	1.26	18.74%	4.65	26.21%
mean + 1.4STD	−10.69	1.01%	0.69	10.94%	1.30	24.95%	4.74	34.40%

and orange circles are the critical dynamic locations of the upstream and downstream risers, respectively. The static equilibrium deflection shapes are plotted in the same figure, and the buoyancy modules are presented by the green lines. It is found from the figure that there are three segments which have high risk of collision, indicated as $S1$, $S2$ and $S3$. More details for identification of critical segments can be found in Fu et al. (2017).

4. Comparison of different methods

The assessment of the different methods for estimation of the extreme values will be done based on the data obtained from the 50 3-h simulations conducted in Section 3. Table 2 summarizes the first four moments of the random process X (defined in Eq. (18)) for different critical segments, i.e. the mean, standard deviation (STD), skewness and kurtosis. The coefficients of variations (CoVs) of these statistical moments are also shown in the same table. The CoV is defined as the ratio between the STD and the mean value, denoting the variation of estimation from each short-term simulation. Generally, the higher statistical moments present larger variations. It should be noted that the skewness of X is negative, indicating that the lower tail of the distribution X is longer. From the table it also appears that the smallest mean value is found for S2, indicating a higher probability of collision at S2. It is also found that even though the segment S3 has the largest mean value, it gives the largest STD, indicating a high risk of collision.

The statistics of the global maxima from the 50 3-h simulations, i.e. X_m are summarized in Table 3, including the first four moments and the associated CoVs for different thresholds. The mean value and STD of the extreme value, i.e. X_e estimated from 50 3-h simulations are given in Table 4.

Table 4
Statistics of the extreme values X_e.

	S1		S2		S3	
	Mean	STD	Mean	STD	Mean	STD
X_e	−8.20	0.648	−5.36	1.129	−6.15	0.990

Table 5
Collision probabilities estimated based on the GEV method.

Location	μ	σ	γ	P_{f3h}
S1	−8.50	0.48	0.03	3.06×10^{-6}
S2	−5.88	0.83	0.05	5.59×10^{-3}
S3	−6.60	0.75	0.02	9.94×10^{-4}

Table 6
Collision probabilities estimated based on the Gumbel paper method.

Location	μ	α	P_{f3h}
S1	−8.50	1.78	1.59×10^{-6}
S2	−5.88	1.03	6.43×10^{-3}
S3	−6.60	1.75	1.37×10^{-4}

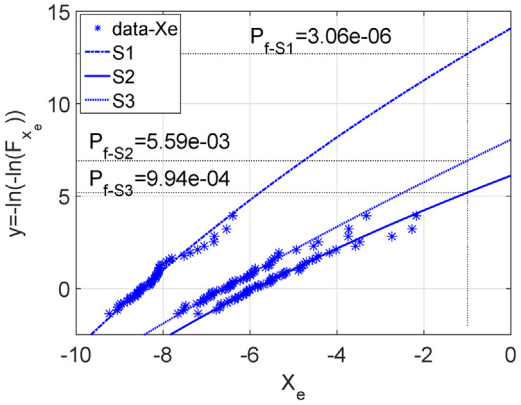


Fig. 6. Gumbel paper plot of the simulated 50 3-h extremes and fitted GEV distributions for three critical locations.

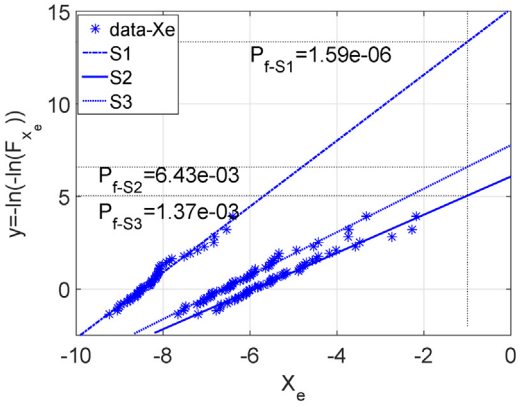


Fig. 7. Gumbel paper plot of the simulated 50 3-h extremes and fitted Gumbel distributions for three critical segments.

4.1. GPP and GEV method

The extreme values, extracted from the 50 3-h simulations, are assumed to be GEV distributed. The distribution parameters are determined by using the maximum likelihood estimates (Anderson and Mathématicien, 1958). In order to compare with the Gumbel model, the results for segments $S1$, $S2$ and $S3$ obtained from the GEV model are plotted in the Gumbel probability paper in Fig. 6. The associated distribution parameters and the 3-h collision probabilities, i.e. $P_{f3h} = 1 - F_{X_e 3h}$, are summarized in Table 5. From the results it is found that the fitted curves are nearly straight lines, and the shape parameter γ is quite close to zero, which is desired for practical application of the 2-parameter Gumbel distribution instead of the 3-parameter GEV distribution. Therefore, in the following sections, the GPP method will be used as a reference method for benchmarking the results of the other ones.

Fig. 7 presents the fitted Gumbel distributions in the Gumbel probability paper by applying the least-square-fitting method. Table 6 lists the associated distribution parameters and the calculated 3-h collision probabilities. The results show that collision most likely occurs for S2, and the probability to have a collision at S1 is much lower.

4.2. Threshold and tail marker

Proposed method 1(PM1)

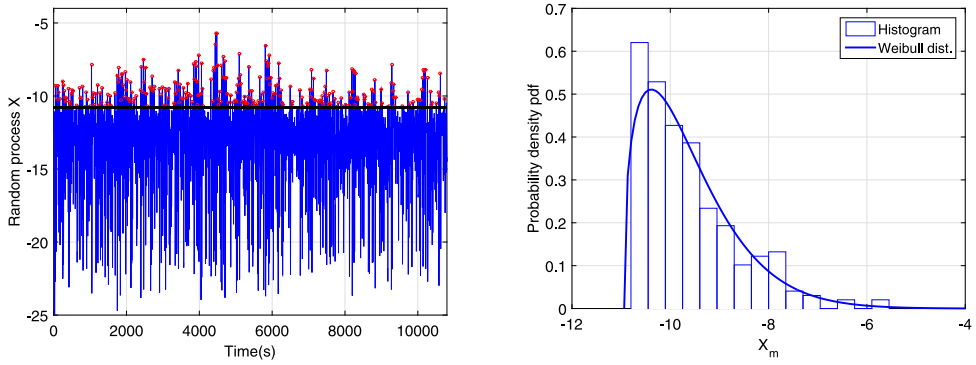
As mentioned previously, five different thresholds are investigated in this section, i.e. $x_0 = \mu_x + [0, 0.5, 1, 1.2, 1.4] * STD$. Fig. 8 presents an example of the selected global maxima with threshold $x_0 = \mu_x + STD$, which are identified from one 3-h time history sample. Fig. 9 shows the estimated maxima and extreme distributions of one 3-h time history for different threshold values in terms of the probability of exceedance, i.e. $P_{f3h} = 1 - F_{X_e 3h}$. The estimated Weibull distributions of the global maxima are illustrated in Fig. 9(a). The estimated Gumbel distributions of the extreme values are presented in Fig. 9(b). The empirical extremes extracted from 50 3-h simulations are plotted in the same figure. From Fig. 9 it appears that the choice of threshold has a significant effect on the statistics and the shape of the distribution. The lower threshold values, i.e. $\eta = 0$ and $\eta = 0.5STD$, preserve the largest number of maxima from the time series. However, the distribution is heavily weighted to lower values, and the fitted distribution does not agree well for the upper tail data. As the threshold value increases, i.e. $\eta = 1.2STD$ and $\eta = 1.4STD$, the weight of the upper tail data becomes important, but the amount of data is reduced. From the results it is found that $\eta = STD$ captures most numbers of the extreme value data for the present study. Therefore, the threshold $\eta = STD$ will be adopted when the Weibull tail fitting method (PM1) is applied.

Proposed method 2(PM2)

For the translation process method based on CFD mapping of a combined parent distribution (PM2), the tail marker is also important. Fig. 10(a) shows the translation functions of the PM2 with different tail markers, i.e. $\eta = [0, 0.5, 1.0, 1.2, 1.4] * STD$. Generally, it is observed from the figure that, for this highly skewed non-Gaussian processes, the PM2 method can give a reliable estimation in the upper tail region. The CDF mapping gives acceptable estimates as the upper tail behavior is well captured. From Fig. 10(b) it appears that there are more extremes to be captured as the tail marker increasing. However, the increase of η from $\eta = STD$ to a higher value does not change the results significantly, thus $\eta = STD$ is considered to be adequate for extreme estimation.

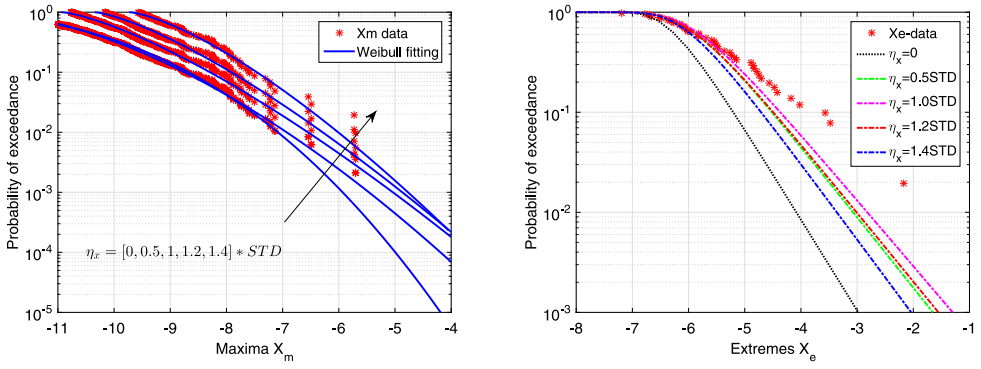
ACER

Regarding the ACER methods, the empirical ACER functions $\hat{e}_k(x)$ for different orders of k are given in Fig. 11(a), which are based on the data from the 50 3-h histories. It appears that, for the lower range of the individual maxima, there is a noticeable variation of the empirical ACER functions for different orders of k, which implies a significant effect of dependence between the data points. However, the result is almost not affected by a further increase of k when $k \geq 2$ in the upper range, which means that there is no dependence effects that need to be accounted for. Thus, $k = 2$ is considered to be adequate for the estimation of extremes. The advantage of the case with the second



(a) Global Maxima. (b) PDF

Fig. 8. Global maxima with threshold $\eta = \sigma$ of a three hour simulation *NO. 20* at S2. (a): Time history and global maxima over threshold. (b): Probability density function. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



(a) Weibull distribution for maxima . (b) Gumbel distribution for extremes

Fig. 9. Maxima and extreme values distribution of a three hour simulation *NO. 20* for different thresholds at S2. (a): Weibull distribution of maxima. (b): Gumbel distribution of extremes. PM1: Proposed method 1-Gumbel method based on Weibull tail fitting of maxima. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

order empirical ACER function is that the result is the most accurately estimated one because most data are available for its estimation.

The estimation of ACER functions requires a better fitting of $\hat{e}_k(x)$ at higher thresholds, and the estimation is not sensitive to the initial tail marker. Therefore, the tail marker $\eta = \sigma$ is chosen, which is consistent with the selection for the previous tail fitting method. Fig. 11(b) presents the empirical function $\hat{e}_2(x)$ and its 95% CI obtained from the time series, as well as the estimated ACER function in the upper tail region and the corresponding estimated 95% CI provided by the extrapolation scheme.

4.3. Comparison of different methods

Due to the statistical uncertainties which are inherent in the stochastic response process, repeated simulations are required in order to obtain a reliable estimation. When the methods PM1 and PM2 are applied, the distribution of the extreme value is determined based on the averages of the expected value and the standard deviation for each simulation, given by

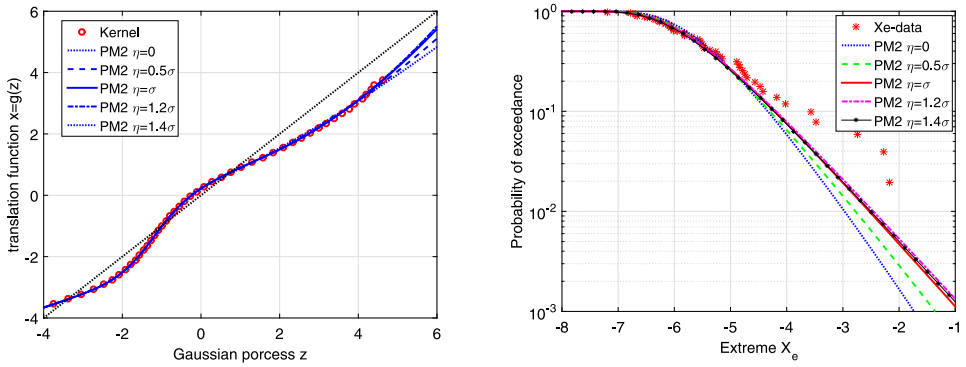
$$E[X_e] = \frac{1}{N}\sum_{i=1}^{N} E_i[X_e]$$
$$STD[X_e] = \{\frac{1}{N-1}\sum_{i=1}^{N} | \ E_i[X_e] - E[X_e]|\}^{\frac{1}{2}} \quad (19)$$

where N is the number of the time histories provided by the simulations and $E_i[X_e]$ is the expected value for simulation number i.

As mentioned previously, the purpose of this paper is to estimate the extreme value distribution by using limited data. Fig. 12 illustrates the collision probability for the different methods at S2 versus the number of simulations N. The threshold value for the ACER, the PM1 and the PM2 methods is chosen as $\eta = STD$ according to the previous discussion. Generally, it appears that the GPP and GEV methods give slightly higher collision probabilities compared with the other methods due to the fact that the results are estimated by fitting to the largest maxima. Regarding the convergence, it appears that the results obtained by applying the PM2, the ACER and the GPP are almost not affected by a further increase of N when $N \geq 20$. When the data is limited, e.g. $N \leq 15$, the accuracy of the GPP method is significantly influenced by the number of simulations, while the PM2 and the ACER methods are able to give an acceptable result when the data is limited.

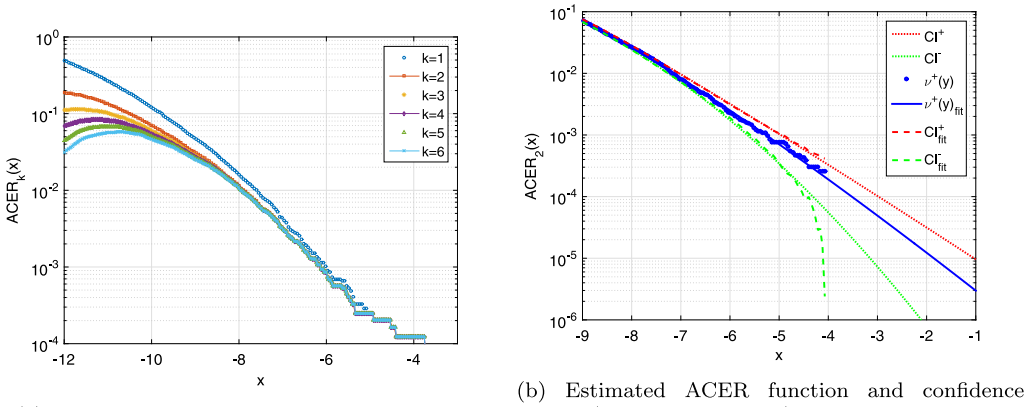
Finally, the estimated extreme value distributions in terms of probability of exceedance for the three critical segments are shown in Fig. 13. The extracted 50 extreme samples are plotted in the same figure. From the figure it also appears that all the methods are able to capture most of the largest maxima for all three segments. For segment S1, it appears that the risk of collision is significantly lower, and the extreme samples are not good enough to predict the collision events. The collision probabilities for different methods are given in Table 7.

As expected, both segments S2 and S3 give higher risk of collision due to the smallest mean value or the largest STD of relative distance, respectively. Fig. 13(c) shows that the collision probabilities obtained from PM1 and PM2 agree well with the results obtained from GPP and GEV for the segment S3, which indicates the potential applicability of the proposed methods. However, Fig. 13(b) illustrates that the methods PM1 and PM2 result in lower collision probabilities compared with the methods GPP and GEV. This is partly due to the dependence between



(a) Translation function. (b) Estimated extreme distribution.

Fig. 10. Extreme values distribution based on translation function at S2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



(a) Experimental ACER function for k=1, 2...6. (b) Estimated ACER function and confidence bands. (k=2 and $\eta = STD$)

Fig. 11. ACER function at S2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

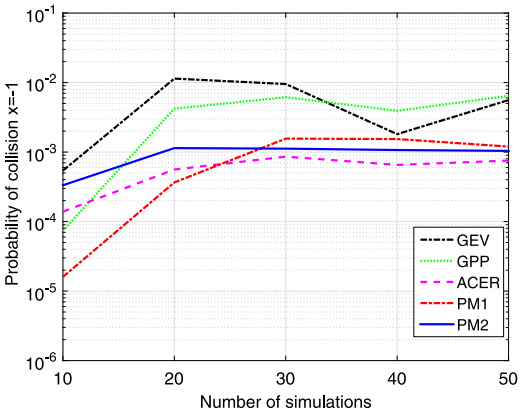


Fig. 12. Convergence study on the number of simulations for different methods. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the selected maxima. Compared with segment S3, the relative distance for segment S2 shows a stronger non-Gaussian behavior. Even when the global maxima are used, the dependence between the maxima, especially the small maxima, decreases the weighting of the larger maxima, leading to a lower collision probability.

Unfortunately, the ACER method fails to give a satisfactory estimation because it is not able to capture the largest extreme values. A possible reason is that the weighting of the smaller maxima is too high. The results also emphasize the importance of the tail marker.

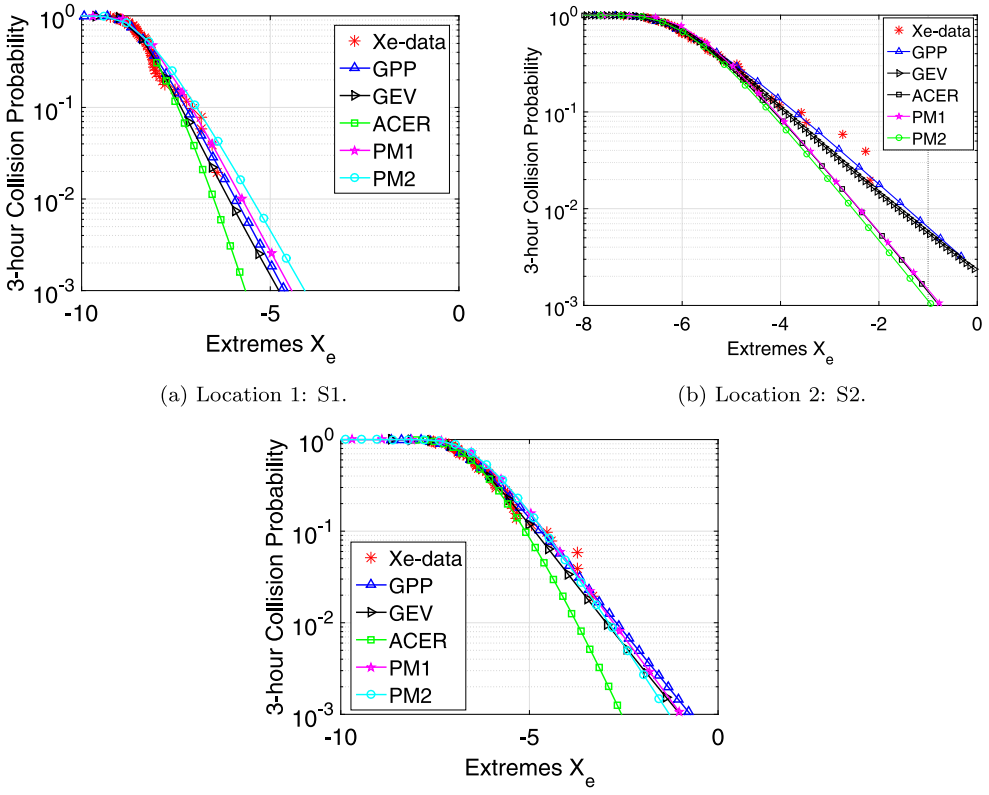
Table 7
Collision probabilities estimated from different extreme value analysis methods.

Methods	GPP	GEV	ACER	PM1	PM2
$P_{f-3h,S1}$	–	–	–	–	–
$P_{f-3h,S2}$	6.43×10^{-3}	5.60×10^{-3}	1.4×10^{-3}	1.50×10^{-3}	1.20×10^{-3}
$P_{f-3h,S3}$	1.51×10^{-3}	1.08×10^{-3}	3.34×10^{-5}	1.10×10^{-3}	7.32×10^{-4}

5. Conclusions

This paper evaluated the performance of different methods for the short-term extreme value analysis for the riser collision problem. A pair of tandem arrangement risers in a steep-wave configuration, subjected to combined current and wave loads are modeled. The Blevins wake model is used to calculate the reduced drag force caused by the wake effect. The minimum relative distance between the risers is computed at each time step. The random process is obtained by changing the sign of the minimum distance in order to deal with the maxima extreme value problem.

The performance of the Gumbel probability paper method, the generalized extreme value method, the average conditional exceedance rate method and the two proposed methods, i.e. the Gumbel method based on Weibull tail fitting of maxima and the translation process method based on a combined parent distribution, are evaluated. The results show that the Gumbel distribution is a good model for the estimation of the riser collision probability when a large number of simulations are available. However, the proposed method PM2 (which is based on the CDF mapping of the combined distribution) gave a



(a) Location 1: S1.

(b) Location 2: S2.

(c) Location 3: S3.

Fig. 13. Extreme values distribution for different locations. GPP: Gumbel probability paper method; GEV: general extreme value method; ACER: average conditional exceedance rate method; PM1: Proposed method 1-Gumbel method based on Weibull tail fitting of maxima; PM2: Proposed method 2-Translation process based on CDF mapping of a combined parent distribution. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

satisfactory estimation when the amount of data is limited. The proposed method 1 (PM1) also resulted in a good prediction with limited data in the present study. The potential weaknesses and strengths of the proposed methods are given. However, a wider range of collision scenarios need to be studied in order to further validate the methods. It should be noted that the selection of an appropriate threshold value has a significant effect on the Weibull tail fitting. The performance of the ACER method is somewhat unstable for the present study.

CRediT authorship contribution statement

Ping Fu: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft. **Bernt J. Leira:** Supervision, Writing – review & editing. **Dag Myrhaug:** Supervision, Writing – review & editing. **Wei Chai:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Anderson, T.W., Mathématicien, E., 1958. An Introduction to Multivariate Statistical Analysis, Vol. 2. Wiley New York.

Blevins, R., 2005. Forces on and stability of a cylinder in a wake. J. Offshore Mech. Arct. Eng. 127 (1), 39–45.

Bury, K.V., 1975. Statistical Models in Applied Science. Wiley.

Chai, W., Leira, B.J., Næss, A., 2018. Probabilistic methods for estimation of the extreme value statistics of ship ice loads. Cold Reg. Sci. Technol. 146, 87–97.

Cunnane, C., 1989. Statistical distributions for flood frequency analysis. Oper. Hydrol. Rep. (WMO).

Ding, J., Chen, X., 2014. Assessment of methods for extreme value analysis of non-Gaussian wind effects with short-term time history samples. Eng. Struct. 80, 75–88.

Duggal, A.S., Niedzwecki, J.M., 1993. Regular and random wave interaction with a long flexible cylinder. In: 12th Intl Conf on Offshore Mechanics and Arctic Engng.

Duggal, A.S., Niedzwecki, J.M., 1994. Probabilistic collision model for a pair of flexible cylinders. Appl. Ocean Res. 16 (3), 165–175.

Farnes, K., Moan, T., 1993. Extreme dynamic, non-linear response of fixed platforms using a complete long-term approach. Appl. Ocean Res. 15 (6), 317–326.

Fu, P., Leira, B.J., Myrhaug, D., 2015. Experiments and analysis related to flow field around two circular cylinders. In: The Twenty-Fifth International Ocean and Polar Engineering Conference. International Society of Offshore and Polar Engineers.

Fu, P., Leira, B.J., Myrhaug, D., 2017. Reliability analysis of wake-induced collision of flexible risers. Appl. Ocean Res. 62, 49–56.

Gaidai, O., Storhaug, G., Naess, A., 2016. Extreme large cargo ship panel stresses by bivariate ACER method. Ocean Eng. 123, 432–439.

Gaidai, O., Storhaug, G., Naess, A., 2018. Statistics of extreme hydroelastic response for large ships. Mar. Struct. 61, 142–154.

Gumbel, E.J., 1958. Statistics of extremes. 1958. Columbia Univ. Press, New York, p. 247.

He, J.W., Low, Y.M., 2013. Predicting the probability of riser collision under stochastic excitation and multiple uncertainties. J. Offshore Mech. Arct. Eng. 135 (3), 031602.

Karpa, O., Naess, A., 2013. Extreme value statistics of wind speed data by the ACER method. J. Wind Eng. Ind. Aerodyn. 112, 1–10.

Leira, B.J., Holmås, T., Herfjord, K., 2002. Probabilistic analysis and design in relation to riser-riser collision. In: The Twelfth International Offshore and Polar Engineering Conference. International Society of Offshore and Polar Engineers.

Moriarty, P.J., Holley, W., Butterfield, C.P., 2004. Extrapolation of Extreme and Fatigue Loads using Probabilistic Methods-No. NREL/TP-500-34421. National Renewable Energy Laboratory Golden, CO(US).

Naess, A., Gaidai, O., 2008. Monte Carlo methods for estimating the extreme response of dynamical systems. J. Eng. Mech. 134 (8), 628–636.

Næss, A., Gaidai, O., 2009. Estimation of extreme values from sampled time series. Struct. Saf. 31 (4), 325–334.

Naess, A., Karpa, O., 2013. Estimation of extreme values by the average conditional exceedance rate method. J. Probab. Stat. 2013.

Naess, A., Moan, T., 2012. Stochastic Dynamics of Marine Structures. Cambridge University Press.

Peng, X., Yang, L., Gavanski, E., Gurley, K., Prevatt, D., 2014. A comparison of methods to estimate peak wind loads on buildings. J. Wind Eng. Ind. Aerodyn. 126, 11–23.
Price, S.J., Paidoussis, M.P., 1984. The aerodynamic forces acting on groups of two and three circular cylinders when subject to a cross-flow. J. Wind Eng. Ind. Aerodyn. 17 (3), 329–347.
Riflex, 1987. Review of flow interference between two circular cylinders in various arrangements. Theory Manual.
Rustad, A.M., Larsen, C.M., Sørensen, A.J., 2008. FEM modelling and automatic control for collision prevention of top tensioned risers. Mar. Struct. 21 (1), 80–112.
SINTEF, O., 2017. Riflex theory manual.
Sodahl, N., 1991. Methods for design and analysis of flexible risers.
Winterstein, S.R., 1987. Moment-Based Hermite Models of Random Vibration. Lyngby.
Winterstein, S.R., 1988. Nonlinear vibration models for extremes and fatigue. J. Eng. Mech. 114 (10), 1772–1790.