Ocean Engineering 357 (2026) 125440

Contents lists available at ScienceDirect

Ocean Engineering

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

Research paper

Design loads for wave impacts — The *Probabilistic Adaptive Screening* (PAS) method for extreme non-linear hydrodynamic loads and responses of maritime structures

Sanne M. van Essen [a,b,*], Harleigh C. Seyffert [a]

[a] Department of Maritime & Transport Technology, Delft University of Technology (TU Delft), Delft, The Netherlands
[b] Ships Department, Maritime Research Institute Netherlands (MARIN), Wageningen, The Netherlands

ARTICLE INFO

Keywords:
Extreme value prediction
Design loads
Wave impacts
Ships
Offshore structures
Coastal structures
Non-linear loads
Probabilistic design
Reliability
Green water
Slamming

ABSTRACT

Wave impact loads on maritime structures can cause casualties, damage, pollution of the sea and operational delays. Consequently, their extreme values should be accounted for in the design of these structures. However, this is challenging, as wave impact events are both rare and highly complex, requiring both high-fidelity simulations and long analysis durations to reliably quantify the associated design loads. Moreover, existing extreme value prediction methods are neither specifically developed nor adequately validated for wave impact phenomena. We therefore introduce the new *Probabilistic Adaptive Screening* (PAS) method for predicting extreme non-linear loads on maritime structures. The method integrates copula-based statistical dependence modelling with multi-fidelity screening and adaptive sampling. This framework enables efficient extreme value prediction by statistically mapping low-fidelity indicator variables to high-fidelity impact loads. The method allows for efficient linear potential flow indicators to be used in the low-fidelity stage, even for strongly non-linear load cases. The statistical framework of the method is validated against four weakly and strongly non-linear test cases, including non-linear waves, ship vertical bending moments, green water impact loads, and slamming loads. It is concluded that PAS with optimal settings accurately estimates both the short-term distributions and extreme values in these test cases, with most probable maximum (MPM) values within 2–15% of the reference brute-force Monte-Carlo Simulation (MCS) results. In addition, PAS achieves this performance very efficiently, requiring in the order of 1–3% of the high-fidelity simulation time needed for conventional MCS. These results demonstrate that PAS can reliably reproduce the statistics of both weakly and strongly non-linear extreme load problems, while significantly reducing the associated computational cost compared to MCS.

1. Introduction and background

1.1. Wave impacts

A *wave impact* on a ship, offshore, or coastal structure is a dynamic loading event that occurs when a water wave hits the structure, causing a rapid transfer of momentum. It may occur when such a maritime structure encounters large and steep waves, experiences large wave-induced motions, or some combination of both (see Fig. 1). The resulting loads can cause significant damage, endanger crew or inhabitants, cause pollution of the sea by lost cargo, decrease performance of the structure, or (in rare occasions) endanger the structure itself. Examples include green water and slamming on ships, wave-in-deck loads on offshore structures, and impacts on wind turbines, breakwaters, dams, jetties, bridges, and other coastal infrastructure. Severe wave impact accidents are documented on sailing ships (Dallinga and Gaillarde, 2001; Kapsenberg, 2018; Pereira, 2023; Reuters, 2023) and offshore structures (Halsne et al., 2022; Ersdal and Kvitrud, 2000; Zhang et al., 2025; Viste-Ollestad et al., 2016; Havtil, 2025). Wave impacts also affect the structural reliability of various coastal structures, ranging from breakwaters (Cuomo et al., 2011a; Zhou et al., 2023), offshore wind turbines (Han et al., 2025), bridges (Cuomo et al., 2009), port infrastructure (Lucio et al., 2024), lighthouses (Antonini et al., 2021), to residential buildings (McCann et al., 2024; Hansom et al., 2015). This makes accurate prediction of the distributions and extreme values of strongly non-linear loads, such as those due to wave impacts, essential. Waves and wave-induced processes are inherently stochastic, necessitating probabilistic design approaches. Even if a high-fidelity (HF) model could perfectly predict the loads associated with a specific

* Corresponding author at: Department of Maritime & Transport Technology, Delft University of Technology (TU Delft), Delft, The Netherlands.
E-mail addresses: s.m.vanessen@tudelft.nl, s.v.essen@marin.nl (S.M. van Essen), h.c.seyffert@tudelft.nl (H.C. Seyffert).

https://doi.org/10.1016/j.oceaneng.2026.125440
Received 28 November 2025; Received in revised form 18 March 2026; Accepted 2 April 2026
Available online 14 April 2026

Fig. 1. Two examples of wave impacts on marine structures: a wind turbine foundation near Fécamp in 2023 (left, photo: K. King) and research vessel *Discoverer* on the Bering Sea in 1979 (right, photo: R. Behn/NOAA).

wave event, this information alone has limited practical value. Without knowledge of the event's probability of occurrence, it is impossible to assess the overall risk or to make reliable design decisions. In other words, the statistics of wave events and wave-induced loads are at the core of the problem.

1.2. Requirements for accurate prediction of extreme wave impact loads

As also discussed in van Essen and Seyffert (2025a), obtaining extreme values for wave impact loads is particularly challenging because such events are both rare and highly non-linear. Accurate and converged statistics of rare events require long-duration analyses, while HF simulations are needed to resolve the detailed physics. HF models in this context can include Computational Fluid Dynamics (CFD) simulations or basin experiments. HF models in this context include CFD simulations and basin experiments. Although CFD can increasingly reproduce wave impact loads for individual events (see Section 1.3), simulating durations in the order of the lifetime of the ship remains computationally infeasible for practical design applications. This motivates the use of dedicated extreme value prediction methods (EVPM), which estimate design loads without assessing very long time series with an HF model. Here, EVPM refers to any method that can obtain extreme values of strongly non-linear responses or loads. By combining probabilistic methods with low- and high-fidelity (LF and HF) simulations, EVPMs focus computational effort on events that matter most for design. While the statistics of wave events remain a key aspect of the problem, the choice of LF and HF models and the physical processes they represent is also crucial, as these decisions directly influence the accuracy of the predicted extreme loads. In this way, an EVPM connects the stochastic nature of waves, model selection, and design considerations. As concluded from van Essen and Seyffert (2023), an EVPM for strongly non-linear loads should then:

1. Integrate multiple fidelity levels, combining suitable HF models for accurate physical modelling with LF models to efficiently capture rare-event statistics;
2. Account for the practical limitation that only a few wave events can be simulated with an HF model during the design stage;
3. Allow for some wave non-linearity in the LF model, since linear models neglect important factors such as wave breaking and higher-order interactions that drive wave impact complexity;
4. Preserve consistent LF and HF time profiles, because peak loads are not always most relevant for structural dynamics (rise times or impulses can be more critical);
5. Have a flexible framework that allows case-specific LF and HF models while preserving the overall methodology, enhancing general applicability;
6. Provide accurate extreme values for strongly non-linear loads, validated using real-world cases;
7. Be efficient and practical for use in design.

1.3. Low- and high-fidelity models

Multi-fidelity EVPMs are only as reliable as the underlying physical modelling and the selected combination of LF and HF models. The HF model must resolve all relevant physical processes with sufficient numerical or experimental fidelity, including appropriate spatial and temporal resolution and geometric representation. For wave impact problems, the available HF modelling approaches are currently CFD simulations or basin experiments. Historically, experiments were the only reliable source of wave impact load data and they remain the primary reference in most existing standards and guidelines (DNV, 2019a, 2018, 2019b; BV, 2015, 2019; ABS, 2020, 2021; ITTC, 2017a). However, several studies have demonstrated that CFD can accurately replicate experimental breaking wave events (e.g., Düz et al., 2020) and wave impact loads when both wave kinematics and ship motions are properly represented (e.g., Bandringa and Helder, 2018; Bandringa et al., 2020; Pákozdi et al., 2022; Gramstad et al., 2023; Bunnik et al., 2015; Scharnke and Helder, 2023; van Essen et al., 2021). While numerous HF modelling choices and numerical settings exist, a detailed discussion is beyond the scope of the present paper. As discussed in the next section, many existing EVPMs also rely on an LF *indicator* variable that is used to identify critical events for subsequent HF evaluation. This LF variable may, but does not need to, represent the same physical quantity as the HF response at a lower resolution. The first requirement to define an LF indicator is that it exhibits similar *order statistics* (ranked values arranged in ascending or descending order) to those of the target HF variable. While this requirement is more easily satisfied when LF and HF models represent the same variable, this study demonstrates that physically related but different indicators can be effective, particularly for strongly non-linear wave impact problems. A second essential requirement is that the LF model should be computationally inexpensive, as it must be evaluated over durations corresponding to several times the lifetime of the structure. This does not mean that the LF model has to be linear; it can be linear or weakly non-linear, and be an analytical model, potential flow model, higher-order spectral wave model, or even a coarse-mesh CFD model. Identifying a suitable LF indicator model that satisfies both requirements is challenging and typically relies on engineering judgement and prior experience, as the suitability of an indicator cannot be assessed beforehand. This reliance on expert model selection is common to most EVPMs. Guidance is therefore usually drawn from previous studies. van Essen and Seyffert (2023) provides a detailed review of suitable indicator variables for different types of wave impact loading. Most of these studies employ either peak values of (relative) wave elevation or related kinematic measures obtained from linear or weakly non-linear potential flow simulations, or wave impact loads computed using coarse-mesh CFD as LF indicators. Furthermore, van Essen et al. (2021) demonstrated that LF indicator results can provide effective inputs for HF CFD simulations.

Table 1
Most important nomenclature.

x'/x''	LF/HF version of variable x	RWE	Relative Wave Elevation	n_w	# wave encounters in MCS
$\hat{x}$	MPM of variable x	VBM	Vertical Bending Moment	N	# seeds in MCS
AS	Adaptive Screening method	C	Wave crest height	P_{exp}	PoE level corr. to T_{exp} and $T_{p,e}$
CFD	Computational Fluid Dynamics	$C(u,v)$	Fitted copula model	R_{19}/R_{bow}	RWE at station 19 or bow
EVPM	Extreme Value Prediction Method	$[\mathbf{d}^*, \mathbf{h}^*]$	Predicted HF distribution	T_{exp}	Target/exposure duration
HF	High-Fidelity	$\left[\mathbf{d}_{H,b}^{\text{cop}}, \mathbf{h}_b^{\text{cop}}\right]$	Distribution HF copula draw b	$T_1/T_{1,e}$	Mean wave (encounter) period
LF	Low-Fidelity	$\left[\mathbf{d}_L^{\text{mcs}}, \mathbf{l}^{\text{mcs}}\right]$	Distribution all LF samples in MCS	$T_p/T_{p,e}$	Peak wave (encounter) period
MCS	Monte-Carlo Simulation	$\left[\mathbf{d}_L^{\text{sel}}, \mathbf{l}^{\text{sel}}\right]$	Distribution present LF samples	V_s	Forward speed of the ship
MD	Acquisition function	F_s	Bow-flare slamming force	V_{hog}/V_{sag}	Hogging/sagging VBM
MVPD	Acquisition function	F_x	Green water force	$V_{r,19}$	RRV at station 19
MPM	Most Probable Maximum	$\mathbf{h}^{\text{sel}}$	Present HF samples	z	# bootstrapping copula draws
PAS	Probabilistic Adaptive Screening	$l_{0,\text{init}}/l_0$	Input/updated LF threshold	ϵ_1/ϵ_2	Case-specific stopping criteria limits
POT	Peak-over-threshold	m	# present HF samples	μ	Wave heading w.r.t. structure
RRV	Relative Rise Velocity	n	# LF indicator peaks in MCS	ω/ω_e	Wave (encounter) frequency

1.4. Existing EVPMs

Extreme value prediction is similar to reliability analysis; the former focuses on extreme loads, while the latter considers failure probability $P_f = P(R - S \leq 0)$, where R is resistance and S loads. Since they are related, we review both, but structural resistance is not considered further. A comprehensive literature review of maritime EVPMs is given in van Essen and Seyffert (2023, 2025a) added recent literature on both EVPMs and reliability methods. Here we summarise the key studies relevant to the present work and briefly discuss developments since the review paper. Response-conditioning methods generate critical wave events conditioned on a given response, using transfer functions, wave spectra, and phase assumptions. They produce single (Adegeest et al., 1998) or multiple wave profiles (Taylor et al., 1997; Torhaug et al., 1998; Jensen, 2009), usually via FORM/SORM. They are efficient and event-based but mostly assume linear Gaussian waves; extensions using higher-order models (Takami et al., 2023; Dermatis et al., 2025b,a) are promising but have not been validated for strongly non-linear impact cases. Screening methods use LF indicator variables, as described above, to identify critical events for HF analysis by ranking waves according to their LF responses (Stansberg, 2008; Bunnik et al., 2018, 2019; Stansberg, 2020; van Essen et al., 2021). Such methods meet most EVPM criteria and could suit strongly non-linear cases, but are associated with a non-conservative bias when the indicator is poorly chosen. This comes from assuming a monotonic relation between the order statistics of the LF and HF variables; when a non-perfect LF indicator seeks the largest magnitude event in a time record, it will invariably sometimes indicate the second, or third (or so on) largest event instead of the true extreme. This always leads to a biased underestimation of the extremes. Such a bias can be mitigated with a safety factor, but this is not ideal. Different types of rare event sampling methods also reduce the number of HF simulations compared to Monte-Carlo simulation (MCS). Examples are importance sampling (Tabandeh et al., 2022; Chiron et al., 2023; Fernandez Castellon et al., 2023), subset sampling (Chan et al., 2022), adaptive sampling, or combinations (Cheng et al., 2023; Yuan et al., 2024; Cheng et al., 2025). Adaptive sampling (or sequential sampling/Bayesian model updating / Bayesian experimental design/active learning) seems most suitable for an EVPM, as it can guide exploration towards the extremes by learning from previous iterations. Adaptive sampling uses surrogates, which can be polynomials, polynomial chaos expansion (Nguyen and Manuel, 2024b,a), support vector regression (Roy et al., 2018; Fang et al., 2024), neural networks (Chojaczyk et al., 2015; Bao et al., 2024), or especially Gaussian Process Regression (GPR, or Kriging) (Fuhg et al., 2021; Forrester et al., 2007; Kennedy and O'Hagan, 2000). Applications of adaptive sampling with GPR include extreme weakly non-linear wave-induced loads on maritime structures (Mohamad and Sapsis, 2018; Gramstad et al., 2020; Guth and Sapsis, 2022; Guth et al., 2023; Tang et al., 2023; Abaei et al., 2024). Adaptive sampling has to stop at convergence; criteria are studied in Huang et al. (2006), Zhang et al. (2024). Multi-fidelity sampling for ship motions was studied by Kim et al. (2025). Response Surface Methods (RSM) use similar surrogates for reliability problems (see e.g., Kim and Lee, 2015; Zhao et al., 2017; Kim et al., 2011; Wilkie and Galasso, 2021; Kingston et al., 2011; Saraygord Afshari et al., 2022), with optimisation strategies discussed in Marrel and Iooss (2024b,a), Li and Wang (2022), Li et al. (2020). Adaptive sampling methods with GPR satisfy most EVPM criteria, but their implementations remain case-specific and untested for strongly non-linear wave-induced loads. Summarising, some existing methods meet multiple EVPM criteria, but their implementations remain case-specific and untested for strong wave impacts. Consequently, industry practice relies on standards (e.g., DNV, 2019a, 2018, 2019b; BV, 2015, 2019; ABS, 2020, 2021; ITTC, 2017a), specifying plate thicknesses or using experiments with multiple 0.5-3 h wave seeds to estimate short-term extreme responses.

To address these limitations of the existing methods, we previously developed the Adaptive Screening (AS) method (van Essen and Seyffert, 2025a). AS is a multi-fidelity approach for predicting extreme non-linear HF wave-induced loads, combining elements from screening, adaptive sampling and GPR. It uses an LF indicator to select critical events for HF evaluation, and builds a surrogate HF distribution from which the design loads are derived, iteratively adding HF samples until convergence. The steps of the method are briefly discussed in Appendix A. AS was validated with good results for a weakly non-linear case (second-order waves) and a moderately non-linear case (vertical bending moments), and further tested in a pilot study of a strongly non-linear green water impact on a containership.[1] Further work (van Essen and Seyffert, 2025b) studied the effect of the *acquisition function*, one of the elements of AS. Most data and all scripts are available in van Essen and Seyffert (2025e,c). The strength of AS for strongly non-linear loads lies in its flexibility, allowing a wide range of LF indicators, including non-linear ones if necessary, and in its efficiency, as it minimises the number of HF simulations. This makes it well-suited for problems with sparse datasets. However, for complex cases, the LF indicator may need to be a relatively expensive model, such as the coarse-mesh CFD used in Case 3 of van Essen and Seyffert (2025a), which increases computational cost. As with other screening methods, a poorly chosen indicator can introduce a systematic non-conservative bias. Consequently, AS meets all requirements in Section 1.2 for simple cases, but in complex problems it cannot always be both accurate and efficient. AS is used as reference method throughout this study.

Existing methods, including AS, generally do not capture the full probabilistic dependence or joint extremes of LF and HF variables, which may explain some limitations. Copula models provide a flexible framework for multivariate distributions, separating marginal behaviour from dependence. They can be used to interpolate the joint

[1] In van Essen and Seyffert (2025a), we used the term *non-linear responses* to emphasise the method's applicability to hydrodynamic motions, accelerations, green water loads etc. However, as discussed in Section 1.4, extreme value prediction aligns with reliability analysis, where *loads* is the standard term. We therefore now use *loads* for all hydrodynamic responses.

distribution learned from a limited number of samples. Further details are given in Appendix B. Copula models have recently been applied to reliability analysis of slopes in spatially variable soils with limited samples (Xu and Liu, 2026; Wang et al., 2025), systems of dependent equipment components (Qiao et al., 2025; Bru-Cordero et al., 2025), reliability-based design optimisation for correlated variables (Zhang et al., 2026; Lu et al., 2025), and assessment of joint risks for breakwater overtopping (Mares-Nasarre et al., 2024). In extreme value analysis, copulas have been used to model joint tail distributions of multiple variables (Krupskii and Joe, 2019) and the dependence between mooring line loads and sea state parameters (Santjer et al., 2025). Copula models have long been used by oceanographers to capture the joint statistics of dependent environmental variables (Vanem et al., 2024; Duan et al., 2024; Fang et al., 2024; de Waal and van Gelder, 2005) and to account for spatial dependence between extreme environmental variables (Beck et al., 2020; Ghosh and Mallick, 2010; Davison et al., 2012). For wave impact loading, copula models have been used to describe the joint distribution of quasi-static and impulsive components of measured deck pressures (Wei et al., 2022), wave impact maxima on a coastal bridge conditioned on rise time (Serinaldi and Cuomo, 2011), extreme stresses at multiple deck locations of a container vessel (Gaidai et al., 2016), interactions among parameters in experimental dam break impacts (Shen and Liu, 2024), maxima and rise times of breakwater loads (Cuomo et al., 2011b), and to derive the most likely wave impact load profile on a floating wind turbine based on wave conditions (Yang et al., 2025). Copula models have previously been integrated with multi-fidelity approaches: (Kramer et al., 2019; Peherstorfer et al., 2016) developed multi-fidelity importance sampling, and Proppe (2019) combined multi-fidelity modelling with a copula to establish a reliability model hierarchy. Copulas have also been paired with adaptive sampling: (Li et al., 2025; Ma et al., 2026) combined Bayesian model updating with MCS and copulas for reliability analysis, while (Bracken et al., 2018) used a Bayesian model with a copula to capture joint annual extremes of time-varying climate variables.

Despite their promise, copula-based approaches have not yet been combined with multi-fidelity or adaptive sampling frameworks to improve the efficiency of extreme value prediction for strongly non-linear loads such as wave impacts. The closest work (Peherstorfer et al., 2016; Kramer et al., 2019) relies on importance sampling rather than adaptive sampling and has not been tested on real-world strongly non-linear cases. Adaptive sampling is generally considered more flexible and robust for complex tail behaviour.

1.5. Objectives & contribution of the present paper

As discussed above and in van Essen and Seyffert (2025a), a new EVPM is needed for wave impact loads, as serious maritime incidents still occur and existing methods are insufficiently validated for this problem. This study is also motivated by the observation that statistical aspects in maritime applications are often treated less rigorously than they could be, despite the inherently stochastic nature of waves and their responses. This work therefore focuses on the statistical framework of the EVPM rather than on LF or HF model details. We propose a new variant of AS that addresses its limitations for complex, strongly non-linear cases. Unlike classical AS, which relies heavily on LF order statistics as surrogates for HF extremes, the proposed method introduces a probabilistic coupling between LF indicators and HF loads. This reduces dependence on expensive LF models and enables the use of simpler, cheaper indicators while maintaining robust and accurate extreme-value predictions. The related objectives of this paper are to:

1. Introduce a modification of AS, called *Probabilistic Adaptive Screening* (PAS), which improves the statistical handling of extreme values of strongly non-linear loads, and complies with the requirements in Section 1.2.
2. Validate that the method can (more) accurately and efficiently predict short-term extreme values (than AS and MCS) for a range of realistic strongly non-linear applications.
3. Demonstrate that the approach can be applied using practical LF indicator models derived from linear potential flow calculations, instead of relying on coarse-mesh CFD.

The novelty of PAS lies in integrating existing method elements into a unified multi-fidelity framework. By combining screening and adaptive sampling with probabilistic LF-HF modelling, PAS enables efficient and statistically consistent prediction of extreme wave impact loads. While PAS can in principle predict both long- and short-term extremes, this work focuses on short-term extremes due to validation data availability. Overall, PAS offers a new, efficient, and accurate EVPM tailored to wave impact loads.

Section 2 presents the new PAS method, and Section 3 introduces four case studies for its validation and comparison with AS and MCS. In addition to wave impact cases, simpler problems are included to assess robustness and illustrate performance across varying levels of non-linearity and modelling complexity. Section 4 describes the performance metrics, Section 5 presents and discusses the results, and Section 6 concludes the study. Table 1 provides a brief nomenclature.

2. The Probabilistic Adaptive Screening (PAS) method

Multi-fidelity extreme value methods link LF indicators and HF loads at different statistical levels (see Fig. 2). Retaining more information (left) improves accuracy but increases complexity and simulation cost, while simplification (right) reduces cost at the risk of losing critical information. AS uses LF order statistics as surrogates for HF extremes, linking them at the level of the marginal distributions in a deterministic way. While efficient, it struggles with strongly non-linear problems, where a single LF value can be associated with a range of HF values. This makes its performance sensitive to outliers in the LF indicator, and costly LF models seem to be required for good wave impact results. The proposed *Probabilistic Adaptive Screening* (PAS) method replaces the deterministic link between LF and HF marginal order statistics with a probabilistic connection between paired LF-HF peaks (as indicated on Fig. 2). PAS retains the screening and adaptive sampling of AS, but replaces the order-statistics surrogate with a probabilistic copula model fitted to the LF-HF pairs. This captures their variability, enabling accurate HF distribution estimates from few HF samples and reducing sensitivity to the LF indicator in strongly non-linear cases. This allows for flexible dependence modelling and efficient prediction of wave impact design loads from simple linear LF indicators.

2.1. Steps

The goal of PAS in the present implementation is to predict the most probable maximum (MPM) value in exposure duration T_{exp} of a target HF non-linear variable. The appropriate T_{exp} depends on the problem and user requirements, and is typically between 20 min and 3 h (depending on the ergodicity of the operating conditions at sea). PAS uses a peak-over-threshold (POT) approach. The method consists of the steps outlined below, numbered according to the schematic overview of PAS in Fig. 3.

Step 1 **LF indicator:** Define an LF indicator variable with an expected strong order-statistics relationship with the target HF response, as in other response-conditioning and screening methods (see Section 1.3). The indicator signal is not necessarily the same physical quantity as the target HF response at a lower fidelity level; it may also be a different but physically related signal. An ideal indicator exhibits a monotonic correspondence between the highest LF indicator values and the largest HF response events. A review of suitable wave impact indicators is provided in van Essen and Seyffert (2023).

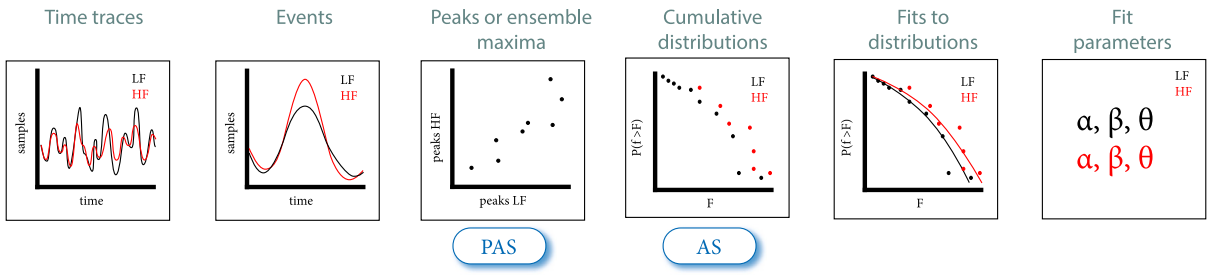


Fig. 2. Some of the possible *statistical levels* where multi-fidelity methods can derive or learn the relation between an LF indicator variables (black) and HF non-linear loads (red), including the location of AS and PAS. Modified from van Essen and Seyffert (2025a).

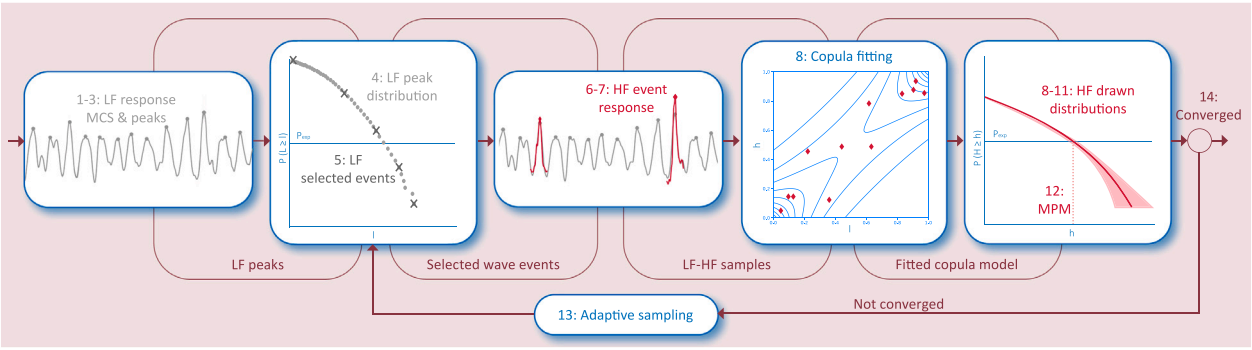


Fig. 3. Schematic illustration of PAS, where the numbers roughly correspond to the method steps in Section 2.1. The left plot only shows a small part of the MCS time traces, and only a few HF samples are included in the middle and right distributions to illustrate the principle.

Step 2 **LF Monte-Carlo simulation:** Perform LF MCS for a large number N of wave seeds, each of duration T_{exp}. See van Essen et al. (2023), Scharnke et al. (2023) for an example of the required N for converged extreme values of wave crests, green water impact forces, and wave-in-deck impact forces. The total MCS duration T_{tot} follows from $T_{\text{tot}} = N \cdot T_{\text{exp}}$. Also determine the total number of zero up-crossing encountered wave crests n_w within T_{tot}. This can be estimated as $n_w \approx NT_{\text{exp}}/T_{1,e}$, or counted directly from the underlying wave record. The mean wave period is taken as $T_1 = T_p/1.98$ (for a JONSWAP spectrum with $\gamma = 3.3$). The corresponding mean encounter period for a sailing ship follows from Eq. (1), with wave frequency $\omega = 2\pi/T$, ship forward speed V_s, and wave heading μ relative to the vessel (with π denoting head seas).

$$\omega_e = \omega - \omega^2 V_s \cos\mu/g \quad \rightarrow \quad T_{1,e} = T_1 \big/ \left(1 - (2\pi V_s \cos\mu)/(T_1 g)\right) \tag{1}$$

Step 3 **LF peaks over threshold:** Identify the n ordered LF indicator peaks $\mathbf{l}^{\text{mcs}}$ that exceed the threshold $l_{0,\text{input}}$ in the MCS duration T_{tot}, using a POT crossing procedure (see Eq. (2)). Such a threshold can be straightforward to define for some problems, for example when only relative wave elevation (RWE) values above a deck level lead to an impact. By ignoring all LF samples below this threshold, excessive sampling in the region where the HF response is likely to be zero can be avoided.

$$\mathbf{l}^{\text{mcs}} = \left(l_1, l_2, \ldots, l_n\right), \qquad l_j > l_{0,\text{input}}, \quad j = 1, \ldots, n \tag{2}$$

Step 4 **LF distribution:** Calculate the LF exceedance probability for all indicator peaks $\mathbf{d}_L^{\text{mcs}} = \{d_{L,i}^{\text{mcs}} | i = 1, 2, \ldots, n\}$, related to the number of wave encounters, by applying Eq. (3). n_w from Step 2 forms the common basis for all exceedance distributions in PAS. By thus normalising both LF indicators and HF responses, we correct their exceedance probabilities for differing numbers of peaks and express them on a common exposure scale, which enables consistent relations between LF and HF distributions and physical durations. The largest LF indicator peak value has an exceedance probability of $1/n_w$ and the smallest n/n_w. This yields LF indicator peak exceedance probability distribution $\left[\mathbf{d}_L^{\text{mcs}}, \mathbf{l}^{\text{mcs}}\right]$, with $\mathbf{d}_L^{\text{mcs}}$ in ascending order.

$$\mathbf{d}_L^{\text{mcs}} = \frac{n}{n_w} \cdot P(\mathbf{l}^{\text{mcs}} \geq l) \tag{3}$$

Step 5 **Initial LF samples:** Select m initial samples from the LF exceedance distribution. The selected set is called $\left[\mathbf{d}_L^{\text{sel}}, \mathbf{l}^{\text{sel}}\right]$, where $\mathbf{d}_L^{\text{sel}} = \{d_{L,k}^{\text{mcs}} | k = 1, 2, \ldots, m\}$ and $\mathbf{l}^{\text{sel}} = \{l_k^{\text{mcs}} | k = 1, 2, \ldots, m\}$. Different sampling strategies can be used to choose the indices k and the associated exceedance probability levels $[p_1, p_2, \ldots, p_m]$, as explained in Section 2.5. Using Eq. (4), we select the LF MCS exceedance distribution elements closest to these levels and assemble them in subset $\left[\mathbf{d}_L^{\text{sel}}, \mathbf{l}^{\text{sel}}\right]$.

$$d_{L,k}^{\text{sel}} = \arg\min_{d \in \mathbf{d}_L^{\text{mcs}}} |p_k - d|, \quad k = 1, 2, \ldots, m \tag{4}$$

Step 6 **Corresponding HF samples:** Obtain the HF loads for the selected indicator peaks in $\left[\mathbf{d}_L^{\text{sel}}, \mathbf{l}^{\text{sel}}\right]$ by running an HF model for the corresponding wave events. The HF model must capture all relevant physical processes, including appropriate grid resolution, time stepping, and numerical settings. For wave impacts, this typically requires fine-mesh CFD or model experiments, with robust initialisation of HF events based on LF simulations (approaches for this are discussed in Section 5.7). The selected samples may turn out to be *true positives* (the LF indicator correctly predicted an impact) or *false positives* (the LF indicator predicted an impact that did not occur). Only true positives (with HF load values greater than zero) are included in the matched LF-HF dataset, to avoid introducing a strong non-conservative bias for rare events. The resulting HF dataset is $\mathbf{h}^{\text{sel}} = \{h_k | k = 1, 2, \ldots, m\}$, where h_k is the HF maximum for event k. False positives are also removed from $\left[\mathbf{d}_L^{\text{sel}}, \mathbf{l}^{\text{sel}}\right]$ to keep the LF-HF datasets consistent, while still being counted in the total number of evaluated HF samples and iterations.

Step 7 **Temporary outlier removal:** Temporarily remove LF-HF pairs whose HF responses strongly deviate from the overall trend. A quadratic model $\hat{h}_k = \text{poly}_2(l_k)$ is fitted to the matched data, residuals $r_k = h_k - \hat{h}_k$ are computed, and samples with $|r_k| > 3 \cdot \text{RMS}(r)$ are identified as outliers. These samples are excluded only for the current iteration and are reconsidered later as new LF-HF pairs are added. The remaining data are then used for copula fitting, which improves the stability of the fit.

Step 8 **Copula fitting:** Fit a joint copula probability model to the LF-HF paired samples $[\mathbf{l}^{\text{sel}}, \mathbf{h}^{\text{sel}}]$. Before fitting the model, the samples are converted to empirical pseudo-observations $[\mathbf{u}, \mathbf{v}]$ on the unit interval [0,1]. This is expressed in Eq. (5), where rank() denotes the statistical rank of the LF or HF elements, and $F_{\mathbf{l}^{\text{sel}}}$ and $F_{\mathbf{h}^{\text{sel}}}$ are the empirical marginal distributions of the LF and HF samples, respectively. This is necessary because copula models are defined on $[0,1]^2$. The copula model selection and fitting procedures are explained in Section 2.2, and the fitting is illustrated in Fig. 4. This yields a fitted copula C for the joint exceedance probability $P(H \geq h, L \geq l \mid H > h_0, L > l_0)$, where L and H are random variables corresponding to the LF and HF samples. The copula is only defined over the sample range, so $l_0 = \min(\mathbf{l}^{\text{sel}})$ (which may equal or exceed the initial $l_{0,\text{input}}$) and $h_0 = \min(\mathbf{h}^{\text{sel}})$.

$$\begin{aligned} u_i &= \frac{1}{m+1}\,\text{rank}(l_i^{\text{sel}}) \approx F_{\mathbf{l}^{\text{sel}}}(l_i^{\text{sel}}) \\ v_i &= \frac{1}{m+1}\,\text{rank}(h_i^{\text{sel}}) \approx F_{\mathbf{h}^{\text{sel}}}(h_i^{\text{sel}}), \qquad i = 1, \ldots, m \end{aligned} \tag{5}$$

Step 9 **Draw conditional HF samples:** Draw z=10 Monte Carlo realisations of HF values from the fitted copula model, conditioned on all LF values in $\mathbf{l}^{\text{mcs}}$ from Step 4 above l_0 (because the copula is not defined below this value). We call this new conditioning set $\mathbf{l}_0^{\text{mcs}}$ (with elements $l_{0,j}^{\text{mcs}} | j = 1, \ldots, n_0$). Multiple realisations are drawn to quantify sampling variability. First, theoretical marginal distributions $F_{l,t}$ and $F_{h,t}$ are fitted to the LF and HF samples, as explained in Section 2.3. Next, the LF pseudo-observations $\mathbf{u}^{\text{mcs}}$ are computed using the fitted LF marginal CDF (see Eq. (6)). For each realisation $b = 1, 2, \ldots, z$ and each conditioning point j, a uniform random variable $r_{j,b}$ is drawn, and the corresponding conditional copula quantile $v_{j,b}^{\text{cop}}$ is computed from the fitted copula model C (see Eq. (7), where ε is a small constant to avoid numerical instabilities). $v_{j,b}^{\text{cop}}$ represents the rank of the HF value conditional on the LF value. The HF samples $h_{j,b}^{\text{cop}}$ (assembled in $\mathbf{h}_b^{\text{cop}}$) in real space are then obtained using the inverse of the fitted HF marginal, as formulated in the last part of Eq. (7).

$$\mathbf{u}^{\text{mcs}} = F_{l,t}(\mathbf{l}_0^{\text{mcs}}) \tag{6}$$

$$\begin{aligned} r_{j,b} &\sim \text{U}(\varepsilon, 1-\varepsilon) \\ v_{j,b}^{\text{cop}} &= C_{V|U}^{-1}\left(r_{j,b} \mid u_j^{\text{mcs}}\right) \\ h_{j,b}^{\text{cop}} &= F_{h,t}^{-1}\left(v_{j,b}^{\text{cop}}\right), \qquad j = 1, \ldots, n_0,\ b = 1, \ldots, z, \quad \varepsilon = 10^{-6} \end{aligned} \tag{7}$$

Step 10 **HF distribution realisations:** Compute an exceedance probability distribution for each set of drawn HF samples $\mathbf{h}_b^{\text{cop}}$. The drawn samples from the copula model are biased, as we only fitted the copula to samples above a threshold. They are debiased by relating the exceedance probabilities to the number of wave encounters instead of the number of LF conditioning values (similar as for the LF distribution in Step 4). This is formulated in Eq. (8), which defines exceedance probabilities $\mathbf{d}_{H,b}^{\text{cop}}$ corresponding to $\mathbf{h}_b^{\text{cop}}$. The largest HF predicted sample drawn from the copula then has an exceedance probability of $1/n_w$ and the smallest n_0/n_w. For each realisation we now have an HF exceedance probability distribution dataset $\left[\mathbf{d}_{H,b}^{\text{cop}}, \mathbf{h}_b^{\text{cop}}\right]$.

$$\mathbf{d}_{H,b}^{\text{cop}} = \frac{n_0}{n_w} \cdot P(\mathbf{h}_b^{\text{cop}} \geq h) \tag{8}$$

Step 11 **Mean HF distribution:** Compute the mean distribution $[\mathbf{d}^*, \mathbf{h}^*]$ and estimate its U95% sampling uncertainty $\mathbf{h}^{*U95}$ over these realisations (see Eq. (9) and (10), where σ_h is the standard deviation over the realisations). Since the exceedance probability range $\mathbf{d}_{H,b}^{\text{cop}}$ is fully determined by the conditioning values, it is the same for every realisation b.

$$\mathbf{d}^* = \mathbf{d}_H^{\text{cop}}, \quad \text{with} \quad \mathbf{d}_H^{\text{cop}} = \mathbf{d}_{H,b}^{\text{cop}}\ \forall b \tag{9}$$

$$\mathbf{h}^* = \frac{1}{z}\sum_{b=1}^{z} \mathbf{h}_b^{\text{cop}} \qquad \text{and} \qquad \mathbf{h}^{*U95} \approxeq \mathbf{h}^* \pm 1.96\,\boldsymbol{\sigma}_h \tag{10}$$

Step 12 **Estimate MPM:** Estimate the T_{exp}-duration MPM $\hat{H}$ from the mean predicted distribution $[\mathbf{d}^*, \mathbf{h}^*]$. The exposure duration and prevailing wave conditions define the target exceedance probability P_{exp} (Eq. (11)), where n_{exp} is the average number of wave encounters during T_{exp}. The MPM follows from the POT exceedance distribution by evaluating it at P_{exp} (Eq. (12)). For rare events, this provides a close approximation of the probability that the maximum within the exposure period exceeds the specified threshold.

$$P_{\text{exp}} = 1/n_{\text{exp}} \approxeq T_{1,e}/T_{\text{exp}} \tag{11}$$

$$\mathbf{d}^*(\hat{H}) = P_{\text{exp}}, \quad \text{therefore:} \quad \hat{H} = \mathbf{d}^{*-1}(P_{\text{exp}}) \tag{12}$$

Step 13 **Adaptive sampling:** Start the adaptive sampling procedure, iterating over Step 6 to Step 13. Each iteration, an acquisition function is applied to define a new sample. The new samples are selected from $\left[\mathbf{d}_L^{\text{mcs}}, \mathbf{l}^{\text{mcs}}\right]$ defined in Step 4, without replacement. The acquisition function is discussed in Section 2.4. If the new HF sample is not a false positive, it is added to the LF selected set $\left[\mathbf{d}_L^{\text{sel}}, \mathbf{l}^{\text{sel}}\right]$ of Step 5 and the number of selected samples m is updated, after which Step 6 to Step 13 are repeated. If the new sample is a false positive (with HF value zero), it is only counted towards to total numbers of evaluated iterations and HF samples. A new sample is then selected using the acquisition function, and the procedure is repeated.

Step 14 **Convergence:** When convergence is reached according to a stopping criterion (Section 2.6), the result is the converged prediction for the HF distribution $\mathbf{h}^*$ over prediction range $\mathbf{d}^*$, the associated MPM value $\hat{H}$, and the MPM uncertainty.

The steps of PAS are presented in general terms, making the method flexible and adaptable to any strongly non-linear problem with suitable LF and HF analysis tools. The following sections describe the specific choices made for the applications in this study: copula fitting and selection (Section 2.2), marginal distributions (Section 2.3), acquisition functions (Section 2.4), initial sampling (Section 2.5), and stopping criteria (Section 2.6). The selection of LF and HF models as well as the applied LF thresholds are case-specific and are discussed in the corresponding case descriptions. Assumptions are discussed in Section 2.7, and implementation details are given in Section 2.8.

PAS is not a physics-free method, as both the LF and HF solvers are based on hydrodynamic principles. PAS establishes a statistical mapping between multi-fidelity models that already embed the relevant physics. Physical consistency is inherited from the governing equations, while the statistical component efficiently transfers information across fidelity levels. Similar multi-fidelity principles are used in many alternative EVPMs, as discussed in Section 1. PAS should therefore be applied with appropriate LF and HF models, selected on a case-by-case basis.

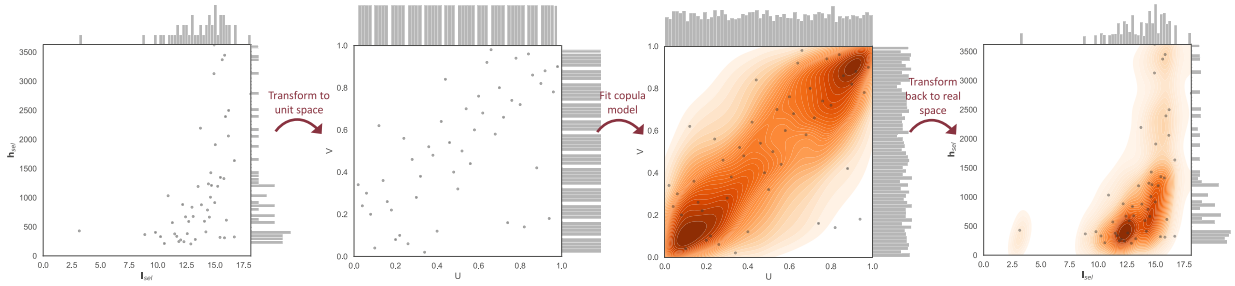


Fig. 4. Example copula-fitting procedure, with from left to right: scatter diagram of LF-HF data in real space, the corresponding empirical pseudo-observations in uniform U,V-space, the fitted copula model in U,V-space, and the transformed data from the copula model in real space (transformed back using fitted marginal distributions).

The brief description of AS in Appendix A indicates which steps are inherited from AS and which steps are new or modified. The similarity requirements between the LF indicator order statistics and the HF load order statistics are less stringent in PAS than in AS. This simplifies indicator selection and, by permitting lower-fidelity indicators, accelerates Step 2 compared to AS.

2.2. Copula fitting and selection

As mentioned in the introduction, a copula model can be used to model the dependence between two variables, here the LF and HF variables. The basics of copula theory are provided in Appendix B. In Step 8, the best-fitting copula for the paired LF-HF data is not known a priori and may vary between iterations. We therefore fit several candidate copulas at each iteration and select the best model. We consider Gaussian, Frank, Clayton, Gumbel, and Student-T copulas (Appendix B.2), which together cover symmetric and asymmetric dependence, with no, one-sided, or symmetric tail dependence. These relatively simple models are robust for small sample sizes. This makes them suitable for repeated fitting within the iterative framework. When an LF threshold is applied, the data are treated as conditional on LF tail events, and all candidate copulas are fitted accordingly. Extreme-value copulas (Güdendorf and Segers, 2010) were also tested, but showed unstable fits for small samples and no performance gain for larger samples. This is consistent with their reliance on asymptotic tail behaviour (Lavaud, 2018). First, each candidate copula was fitted to the paired empirical pseudo-observations $[\mathbf{u}, \mathbf{v}]$ from Eq. (5). We used the default fitting methods from the used copula packages; see Section 2.8 for implementation and fitting details.

Formal goodness-of-fit tests such as Cramér–von Mises or Kolmogorov–Smirnov do not work well for small datasets, as we have in PAS. The goal is to select the most plausible copula at each iteration, not to identify the true model. We therefore use the Akaike Information Criterion (AIC, Akaike, 1974) in Eq. (13) to choose among the five candidates. The log-likelihood ℓ in Eq. (14) is computed from the copula density c at pseudo-observations $[\mathbf{u}, \mathbf{v}]$, using fitted parameters $\hat{\boldsymbol{\theta}}$ (via maximum likelihood or rank-based inversion of Kendall's tau). AIC balances fit and complexity; most copulas have one parameter k, except Student-T with two. To limit volatility in model selection across iterations, a new copula is adopted only if its AIC is at least 4.0 lower than the previous model; otherwise, the previous model is retained (following Burnham and Anderson, 2002).

$$\mathrm{AIC} = 2k - 2\ell(\hat{\boldsymbol{\theta}}) \tag{13}$$

$$\ell(\hat{\boldsymbol{\theta}}) = \sum_{i=1}^{m} \log c(u_i, v_i; \hat{\boldsymbol{\theta}}) \tag{14}$$

2.3. Marginal distribution fitting

In Step 8, the LF-HF paired samples are transformed to empirical pseudo-observations, without requiring parametric marginals. Full marginal distributions are needed only in Step 9 to interpolate conditional copula samples back to physical space. For small samples, interpolation using parametric fits is preferred, as linear interpolation can create kinks and kernel smoothing may distort tails. The copula is valid only over the fitted LF-HF ranges, so these marginals are used solely for interpolation.

For this purpose, we fit three-parameter Weibull cumulative distributions (Eq. (15)) to the available LF and HF samples separately. Weibull-type tails correspond to a negative shape parameter ($\xi < 0$) in the generalised extreme value (GEV) or generalised Pareto (GP) framework. While the GP distribution represents the asymptotic limit for threshold exceedances, its maximum likelihood estimates are highly variable and sensitive to the threshold in small samples (e.g., Davidson and Smith, 1990; Jonathan et al., 2021). The three-parameter Weibull provides a more stable and accurate interpolation of the empirical tail over the observed range and is commonly used in the field (see e.g., ITTC, 2017b; DNV, 2018; Haver, 2017). We apply least-squares fitting, as recommended for tail modelling (e.g., DNV, 2019a).

$$P(X \le x) = 1 - \exp\left(-\left(\frac{x-\theta}{\alpha}\right)^{\beta}\right) \tag{15}$$

Because the Weibull fit is formally valid only for independent POT values, an optional upper-tail fitting mode is provided when no threshold is set. For very small samples (< 10), all data are used; for large samples (> 30), only the upper 80% is fitted, with a linear transition from 100 to 80% for intermediate sizes (10–30). When a threshold is specified, the full dataset is always used.

2.4. Adaptive sampling and acquisition function

Step 13 of PAS uses an acquisition function to define the next best sample to evaluate. We select one new HF sample per iteration. In van Essen and Seyffert (2025b), AS was evaluated with seven acquisition functions. Those distributing samples uniformly along the logarithmic exceedance probability axis or concentrating near P_{exp} were most effective. However, the HF distribution in PAS is constructed differently, and copula fitting requires samples across a wider range. Accordingly, PAS uses the maximum-distance (MD) acquisition function, which selects new samples based on their largest separation from existing HF points on the logarithmic exceedance probability axis.

For each iteration i, we have the set of LF selected samples $\mathbf{d}_L^{\mathrm{sel}}$ and the set of HF selected samples $[\mathbf{d}_H^{\mathrm{sel}}, \mathbf{h}^{\mathrm{sel}}]$ in Step 5, with $\mathbf{d}_H^{\mathrm{sel}} = \{d_1^s, d_2^s, \ldots, d_m^s\}$ ordered such that $d_1^s < d_2^s < \cdots < d_m^s$. Here we drop the iteration index i for convenience. We now define p_{new} as the optimal exceedance probability for adding a new LF sample to the set. The new point is located halfway between $\ln(d_k^s)$ and $\ln(d_{k+1}^s)$, where the gap $[\ln(d_{k+1}^s) - \ln(d_k^s)]$ is the largest among all pairs of consecutive

selected samples. The function operates on the logarithm of exceedance probability to focus on the tail of the distribution.

$$p^{\mathrm{md}}_{\mathrm{new}} = \underset{k\in\{1,\ldots,m-1\}}{\arg\max} \left(\frac{\ln(d^s_{k+1}) - \ln(d^s_k)}{2} + \ln(d^s_k) \right) \tag{16}$$

This would be the best new sample value to add to $\mathbf{d}^{\mathrm{sel}}_{L,i}$ on a continuous scale (including iteration subscript i again). However, in PAS, we only have the discrete LF samples in the available LF MCS sample pool of Step 4. We therefore select the sample closest to $p^{\mathrm{md}}_{\mathrm{new},i}$ from the available pool $\mathbf{d}^{\mathrm{mcs}}_{L} \setminus \mathbf{d}^{\mathrm{sel}}_{L,i}$, without replacement (i.e., excluding already selected samples; see Eq. (17)). The new sample $d_{L,i}$ is added to the existing LF pool $\left[\mathbf{d}^{\mathrm{sel}}_{L,i}, \mathbf{l}^{\mathrm{sel}}_{i}\right]$ in Step 5 for the next iteration.

$$d_{L,i} = \underset{d\in\mathbf{d}^{\mathrm{mcs}}_{L}\setminus\mathbf{d}^{\mathrm{sel}}_{L,i}}{\arg\min} \left|\ln(p^{\mathrm{md}}_{\mathrm{new},i}) - \ln(d)\right| \tag{17}$$

The risk associated with this MD function when the physical threshold in the input is not properly set is that it may oversample zero HF values, which do not contribute to characterising the HF extremes. To mitigate this, we implemented a stabilising mechanism: if 5 consecutive zero HF values (or false positives) occur and the exceedance probability of a new sample exceeds that of any existing sample in $\left[\mathbf{d}^{\mathrm{sel}}_{L}, \mathbf{l}^{\mathrm{sel}}\right]$, the LF threshold l_0 is increased by 2% of the maximum LF value in $\mathbf{l}^{\mathrm{mcs}}_{0}$ for the next iteration.

Choosing MD as the acquisition function results in a semi-adaptive scheme: the next sample is determined by the previously selected conditioning points rather than the outcomes of the previous iteration, so adaptivity only enters through the stopping criterion and the adaptive threshold explained above. Fully adaptive sampling functions can still be employed in PAS, but in practice they tend to perform worse than MD.

2.5. Initial sampling

Bayesian and adaptive sampling methods can be sensitive to initial samples, as early data shape the first predictions and subsequent updates. Poor initial coverage may slow convergence or bias results. Although PAS combined with MD is not fully Bayesian, initial sampling can still influence outcomes. To reduce this sensitivity, we use an initial sampling strategy that is consistent with the MD acquisition function: four samples selected at uniformly spaced indices over $\ln(\mathbf{d}^{\mathrm{mcs}}_{L})$ from Step 4.

2.6. Stopping criterion

We stop the iterative adaptive sampling procedure when new iterations no longer produce significant changes in the predicted exceedance probability distributions. This is done using a stopping criterion consisting of two parts. The first part sets a limit ϵ_1 for the mean absolute difference between each set of subsequently predicted distributions, averaged over a number of the last iterations. The second part sets a limit ϵ_2 for the coefficient of variation (COV) of the MPM value, averaged over a number of the last iterations. The complete formulations for this stopping criterion are in Appendix C. This appendix also explains how this criterion is different from the criterion used for AS in van Essen and Seyffert (2025a).

2.7. Assumptions

The most important assumption in PAS is the similarity between the order statistics of the LF indicator and the HF response in Step 10. Most alternative methods, such as response-conditioning methods and other screening techniques discussed in van Essen and Seyffert (2023), also rely on this assumption. Its validity depends on the indicator choice; a poorly chosen indicator can significantly reduce the result accuracy. However, introducing the copula fitting technique in PAS reduces the weight of this assumption, as it makes the LF-HF coupling probabilistic instead of deterministic. This reduces the impact of outliers in the order statistics comparison. Another critical assumption in Step 6 is that the HF tool accurately calculates the true HF event response. Previous studies have shown that CFD can effectively predict wave impact loads if wave kinematics and ship motions are modelled well (as discussed in Section 1.3). Additionally, van Essen et al. (2021) demonstrated that screening results could serve as effective inputs for such calculations. Finally, PAS assumes that the LF-HF mapping of each problem fits one of the candidate copula models discussed in Section 2.2. The dynamic best-fitting selection procedure in Step 8 already reduces the method's reliance on a single copula model. The present study shows that these assumptions work for a wide range of considered test problems.

2.8. Implementation

The PAS framework, including its iterative procedure, acquisition functions, and stopping criteria, was implemented in `Python` (v3.8.20). It supports the integration of copula models from various sources. In this study, we use Gumbel, Clayton and Frank copula models from the Python copulas package (v0.12.0), and Gaussian and Student-T copula models from the `R` (v4.5.2) library `copula` (v1.1.6) (Hofert et al., 2016, 2018) via an embedded `Python` interface (using package `rpy2`, v3.6.4). `R` provides access to a wider range of models than `Python`, including extreme-value copulas. For the copulas from the `Python copulas` package, model parameters were estimated using the default rank-based procedure based on inversion of Kendall's tau (which has a fixed relation with the copula parameters, see Appendix B.2). For the copulas from the `R copula` package, model parameters were estimated using maximum likelihood via the `fitCopula` function (method = 'ml'). All scripts are available in the 4TU repository: (van Essen and Seyffert, 2025d). Although the code is still research-grade and not yet optimised for speed, the computational time required to run the PAS scripts for a single iteration is negligible compared to the LF and, in particular, the HF physical models. Running PAS for one iteration typically takes only a few seconds to one minute on a single laptop core.

3. Test cases

We applied PAS (and AS for reference) to four cases, each with a variation. Table 2 summarises the cases, with the corresponding wave and ship operating conditions listed in Table 3, and the ground-truth MPM values from the validation dataset provided in Table 4. The cases increase in complexity. Case 1 (waves) is a simple analytical wave-only, weakly non-linear problem, where the LF and HF models resolve the same variable at different fidelity levels. Case 2 (vertical bending moments, VBM) is a fully numerical, weakly non-linear ship response problem, again with LF and HF models resolving the same variable. Case 3 (green water) introduces strongly non-linear HF wave impact loads and combines numerical LF indicators with experimental HF measurements, demonstrating that PAS can link LF and HF models even when they resolve different physical variables. Additional complexity is due to threshold behaviour and the presence of false positives. Finally, case 4 (slamming) represents an even more non-linear case, again combining numerical LF indicators with experimental HF measurements, adding the complexity of large outliers in the LF-HF relation. The magnitude of slamming loads strongly depends on relative wave-ship timing and shape, explaining the occurrence of such outliers. The green water and slamming cases include two sea states, while the wave and VBM cases each use other wave conditions. PAS accounts for these wave conditions in the underlying LF and HF models. By considering eight sub-cases involving in total four wave conditions, we evaluate the robustness and accuracy of PAS across a range of wave environments, including steep and strongly non-linear waves. Together, these cases cover a large spectrum of non-linearity, complexity, and LF-HF model combinations. The focus is on strongly non-linear wave

Table 2
Summary of considered cases, where AC = accommodation, BF = bow flare, GW = green water, PF = potential flow, RWE = relative wave elevation, RRV = relative rise velocity, VBM = vertical bending moment and T_{MCS} = MCS duration. The variable P_{exp} for an exposure duration (T_{exp}) of 1 h for all cases is defined in Eq.(11), and n_w in Step 2 of PAS. Cases 1a,b and 2a are identical to those evaluated with AS in van Essen and Seyffert (2025a).

Case	Structure	HF variable	HF source	LF indicator	LF source	Conditions Table 3	T_{MCS} [h]	n_w [–]	P_{exp} [–]
1a	–	2nd-order waves	Analytical	Lin. waves	Analytical	A	50:00	16,364	2.55×10^{-3}
1b	–	2nd-order waves	Analytical	Lin. noisy waves	Analytical	A	50:00	16,364	2.55×10^{-3}
2a	Ferry	Hogging VBM	Non-linear PF	Hogging VBM	Linear PF	B	30:00	14,359	1.74×10^{-3}
2b	Ferry	Sagging VBM	Non-linear PF	Sagging VBM	Linear PF	B	30:00	14,359	1.74×10^{-3}
3a	Ferry	GW force AC	Experiments	RWE bow	Linear PF	Heavy	34:49	15,198	1.91×10^{-3}
3b	Ferry	GW force AC	Experiments	RWE bow	Linear PF	Extreme	23:43	10,900	1.82×10^{-3}
4a	Ferry	Slam force BF	Experiments	RRV st19	Linear PF	Heavy	34:49	15,198	1.91×10^{-3}
4b	Ferry	Slam force BF	Experiments	RRV st19	Linear PF	Extreme	23:43	10,900	1.82×10^{-3}

Table 3
Summary of irregular wave conditions (all with a JONSWAP spectrum) and operational conditions used in the cases, where H_s is significant wave height, T_p is peak wave period, $s = 2\pi H_s/(gT_p^2)$ is non-dimensional wave steepness, γ is the JONSWAP peak enhancement factor, h is water depth, μ is the wave direction w.r.t. the ship (180 deg is head waves), V_x is the forward speed of the ship or structure, V_y is its transverse drift speed, and $T_{p,e}$ is the corresponding peak wave encounter period (Eq.(1)). The names of the last two conditions were defined in the corresponding experimental campaign.

Name	H_s [m]	T_p [s]	s [–]	γ [-]	h [m]	μ [deg]	V_x [m/s]	V_y [m/s]	$T_{p,e}$ [s]	Comment
A	10.0	11.0	0.053	3.3	30.0	–	–	–	–	No ship
B	13.2	10.0	0.085	3.0	1000.0	180	5.1	0.0	7.5	
Heavy	8.1	9.4	0.059	3.3	179.9	150	2.4	0.4	8.2	
Extreme	8.3	10.0	0.053	3.3	179.9	150	5.0	0.6	7.8	

impact problems, where alternative extreme value prediction methods are limited or inapplicable. Accordingly, wave impact cases 3 and 4 are most relevant for PAS. The simpler cases, which may also be solvable with more efficient methods or even direct HF MCS, are included to demonstrate the robustness and performance of PAS across increasingly complex scenarios.

3.1. Cases 1 (second-order waves) and 2 (vertical bending moments)

Cases 1 and 2 are identical to those used in van Essen and Seyffert (2025a) to validate AS, so their description is only briefly repeated in Appendix D.
Summarising case 1:

- HF target extreme value: one-hour MPM of second-order wave crest height $\widehat{C''}$.
- LF indicator: linear wave crest height C'_{good} (case 1a), linear noisy wave crest height C'_{worse} (case 1b).

Summarising case 2:

- HF target extreme value: one-hour MPM of weakly non-linear hogging $\widehat{V''_{hog}}$ (case 2a) and sagging $\widehat{V''_{sag}}$ (case 2b) VBM.
- LF indicator: linear hogging V'_{hog} (case 2a) and sagging V'_{sag} (case 2b) VBM.

3.2. Case 3 (strongly non-linear): green water impact forces

Case 3 considers a strongly non-linear problem: predicting extreme green water impact forces on the accommodation of a 190 m ferry (the same vessel as in Case 2) in extreme irregular waves, under two severe, steep bow-quartering wave conditions (Heavy and Extreme in Table 3). The HF green water impact forces in longitudinal ship direction on the accommodation, called F''_x, were obtained from long-duration model experiments. Details of the experiments are described in van Essen et al. (2023, 2024), with selected aspects (e.g., a discussion of potential scale effects) summarised in Appendix E.2. Details of the test conditions are provided in Table 2, showing test durations of 34:49 h for the Heavy condition and 23:43 h for the Extreme condition. As discussed in van Essen et al. (2023), this is sufficiently long to have a converged HF validation reference for the one-hour MPM. Some pictures of the experiments are shown in Fig. 5. The impact forces were measured at a sample frequency of 800 Hz full-scale with 40 force panels (van de Bunt et al., 2021) on the front of the accommodation (total area 9.0 × 14.4 m full-scale), with F''_x representing the instantaneous sum of all panel measurements. Previous studies have shown that the relative wave elevation (RWE) around the ship's bow can be an effective indicator for green water events (see e.g., van Essen et al., 2021; Hoffman and Maclean, 1970; Hong et al., 1993; Ogawa et al., 2002; van 't Veer and Vlasveld, 2014). Accordingly, the selected LF indicator for this case is R'_{bow}, generated using a linear potential flow simulations. The location of the RWE probe in the simulations mirrors that in the experiments (see Fig. 5). van Essen and Seyffert (2025a) demonstrated that accurate green water predictions were achieved for a pilot case using AS in combination with an LF indicator generated from coarse-mesh CFD. If linear potential flow can be used instead to generate an indicator for green water impact loads (in combination with PAS), this can have a large advantage in computational time. The LF-HF peak scatter plots are shown in Appendix E.1. The plot shows that, as expected, the LF and HF peaks for this case have less similar order statistics than those of case 2. Because the experimental green water force peaks (ground truth HF validation data) have a large tail variability, we fitted a three-parameter Weibull distribution (Eq. (15)) to the 30% highest force peaks, to derive the reference 'true' distribution. Appendix E.3 provides more details on the potential flow simulations used to generate the LF data. We used an LF threshold of 12.0 m in Step 3 of PAS. This value was chosen slightly above the local freeboard at the bow (10.5 m). RWE values below this level are expected to generate many false positives. Due to the ferry's large bow flare angle, events with RWEs only marginally exceeding the freeboard are often deflected to the sides, preventing water from flowing onto the deck and thus avoiding green water loads on the accommodation. Summarising:

- HF target extreme value: one-hour MPM of green water impact force $\widehat{F''_x}$ in the Heavy (case 3a) and Extreme (case 3b) condition.
- LF indicator: linear potential flow RWE at the bow R'_{bow}.

3.3. Case 4 (strongly non-linear): bow-flare slamming forces

Case 4 studies another strongly non-linear problem, based on the same experiments as case 3. Here we predict extreme values of the bow-flare slamming forces on the ferry, sailing in the same two irregular

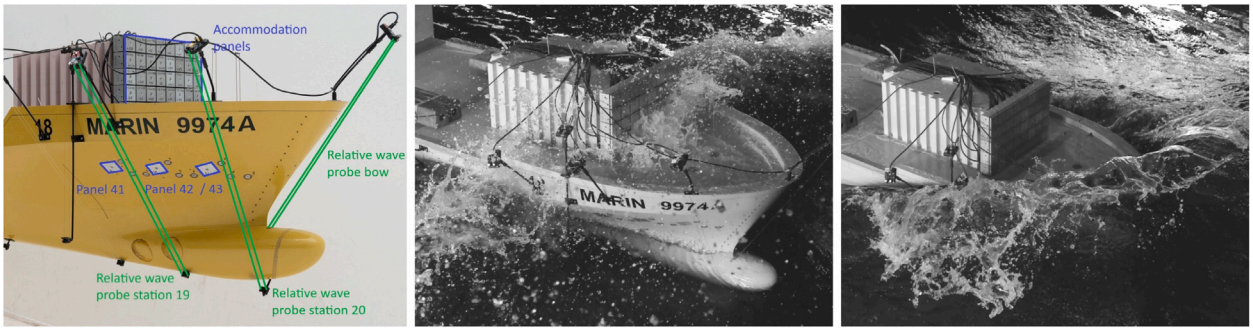


Fig. 5. MARIN ferry 2 and the relevant instrumentation around the bow before the *CRS SCREAM* experiments (left), example green water impact (middle) and example bow-flare slamming impact (right).

Heavy and Extreme bow-quartering wave conditions. The HF slamming impact forces, denoted F_s'', were obtained by summing the measured peak forces for each wave event across three force panels located in the bow flare (see Fig. 5). The size of each of these panels 41, 42 and 43 was 1.8 × 1.8 cm full-scale, and they were sampled at a frequency of 800 Hz full-scale. Previous studies have shown that the relative velocity between the wave crest and the ship's bow flare can be an effective indicator for slamming events (see e.g., Bunnik et al., 2018, 2019; Kapsenberg, 2018; Ochi, 1964). Motivated by this, we adopt the relative rise velocity (RRV) at station 19, denoted $V'_{r,19}$, as the LF indicator. For each wave event, the RRV is defined as the peak RWE divided by its rise time, with the rise time measured from the RWE's zero up-crossing to the subsequent peak. The RWE used to compute the RRV is obtained from linear potential flow simulations on the weatherside at station 19 (located at 19/20 of the ship length from the stern; see Fig. 5). The LF-HF peak scatter plots are shown in Appendix E.1, which shows that the LF and HF peaks for this case have a distinctly different relation than those of the other cases. The true HF MPM values from the experiments in Table 4 were again derived from a 3-parameter Weibull fit to the 30% highest force peaks. Appendix E.3 provides more details on the potential flow simulations used to generate the LF data. We did not use an LF threshold value in Step 3 of PAS for the RRV, because no clear threshold can be identified for the occurrence of slamming events at the panels. This would differ if, for example, peaks in RWE would be used as the indicator, as this would naturally introduce a threshold at some distance above the panel underside measured from the calm waterline (2.875 m). Summarising:

- HF target extreme value: one-hour MPM of bow-flare slamming force $\widehat{F_s''}$ in the Heavy (case 4a) and Extreme (case 4b) condition.
- LF indicator: linear potential flow relative rise velocity at station 19 $V'_{r,19}$.

4. Method used to validate PAS

In a practical design procedure for a maritime structure, HF CFD simulations or experiments would need to be performed for new wave events at every iteration in Step 6 to obtain the required HF loads. In the present validation study, however, we circumvent this requirement: rather than running CFD simulations, we extract the corresponding HF load responses directly from the validation datasets for each selected wave event. This allows us to validate the statistical framework of PAS without mixing it with CFD load prediction accuracy (following the same approach used in van Essen and Seyffert (2025a) for AS). Earlier studies have shown before that accurate HF wave event impact loads can be obtained with CFD (see Section 1.3).

For cases 1 and 2, obtaining identical deterministic wave inputs for the LF and HF variables was straightforward. For cases 3 and 4, this was achieved by combining experimentally measured waves recorded next to the ferry (propagated to the vessel's centre of gravity), with the linear frequency-domain potential flow database of RWE response functions. This ensured that the wave phasing of the LF time traces closely matched that of the experimental waves. These simulations were previously performed in van Essen et al. (2024), where the full procedure is described in detail.

The performance of PAS compared to the validation material is assessed using three metrics: M_1 for efficiency, and M_2 and M_3 for accuracy. M_1 (Eq. (18)) is the ratio of HF samples required to reach convergence, n_c, to the total number of HF samples without PAS, n_w (the total number of wave events in the validation dataset). M_2 measures the mean deviation of the MPM prediction over the last 10 true positive iterations at convergence, $\widehat{H}_{i_c,10}$, from the ground truth $\widehat{H}_t$ (Eq. (18)), where i_c is the converged iteration number. Finally, M_3 quantifies the maximum deviation of the predicted distribution from the ground truth over an exceedance probability range $[0.5P_{\text{exp}}, 2P_{\text{exp}}]$ (Eq. (19)), again averaged over the last 10 true positive iterations at convergence, with $\mathbf{h}^*_{\text{ran},i_c,10}(k)$ and $\mathbf{h}_{\mathbf{t,ran}}(k)$ denoting the mean predicted and true distributions over this range with elements k, respectively. Convergence of PAS can be somewhat erratic; the averaging over a number of iterations reduces this noise. Smaller absolute values of all three metrics indicate better performance. Note that M_2 and M_3 may also be negative, indicating whether predictions are conservative. In theory, the convergence criteria should take care of the accuracy of the prediction. However, M_2 and M_3 are used to check whether the results converge towards the ground truth and not towards a biased value. These criteria are similar, but not identical, to those used for AS in (van Essen and Seyffert, 2025b).

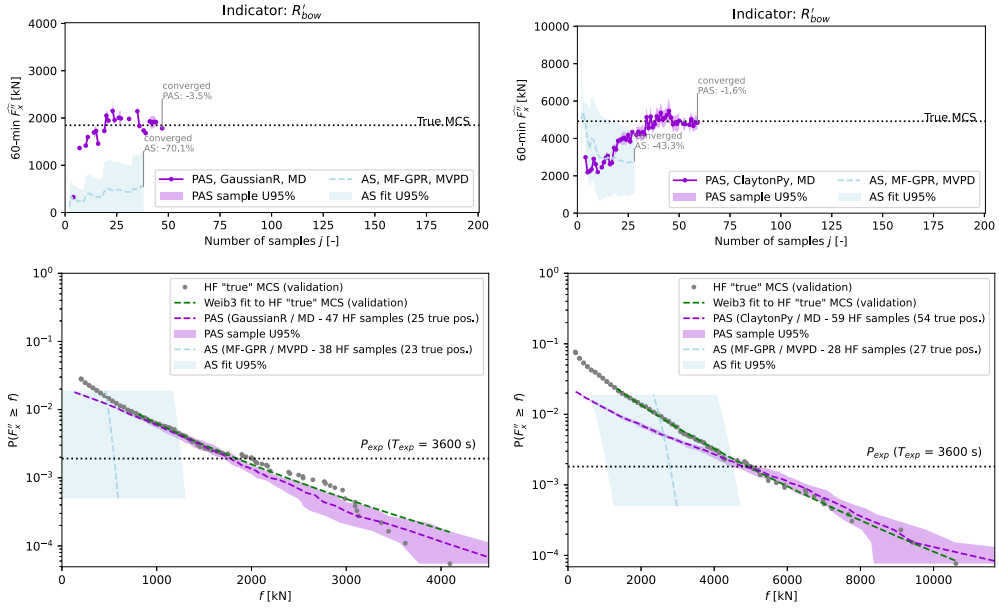
$$M_1 = \frac{n_c}{n_w} \qquad M_2 = \frac{\widehat{H}_{i_c,10} - \widehat{H}_t}{\widehat{H}_t} \tag{18}$$

$$\boldsymbol{\Delta}(k) = \mathbf{h}^*_{\text{ran},i_c,10}(k) - \mathbf{h}_{\mathbf{t,ran}}(k) \quad \rightarrow \quad k^{\text{absmax}} = \arg\max_k |\boldsymbol{\Delta}(k)| \quad \rightarrow \quad M_3 = \frac{\boldsymbol{\Delta}(k^{\text{absmax}})}{\widehat{H}_t} \tag{19}$$

The performance of PAS is compared to that of AS and the brute-force MCS approach. An overview of the steps in AS that differ from PAS and some relevant formulations and settings are included in Appendix A. It is important to note that throughout this study, PAS and AS are evaluated using the same analytical and linear potential-flow LF indicators. This represents a deviation from the approach used for pilot-case 3 in van Essen and Seyffert (2025a), where the AS method was used in combination with an LF indicator derived from coarse-mesh CFD simulations. That AS-coarse mesh CFD combination yielded accurate distributions of green water impact loads on a containership in the pilot study, but using this in practice would be associated with a high computational cost.

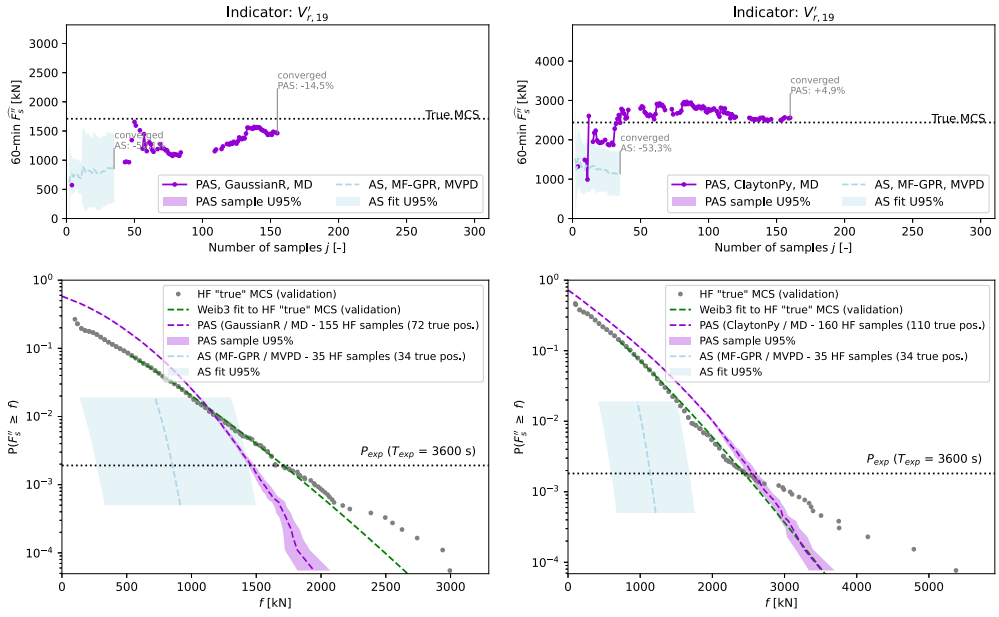
5. Results and discussion

Before comparing PAS and AS results to each other and to the MCS ground truth, we define stopping criteria for $S(j)$ in Eq. (30): ϵ_1 for



(a) Case 3a, Heavy condition (b) Case 3b, Extreme condition

Fig. 6. Case 3 - green water: convergence of one-hour MPM as a function of number of samples (top) and final converged distributions (bottom) from AS and PAS, both with linear potential flow indicator R'_{bow}. The copula in the name of the PAS results is the utilised model in the last iteration.



(a) Case 4a, Heavy condition (b) Case 4b, Extreme condition

Fig. 7. Case 4 - slamming: convergence of one-hour MPM as a function of number of samples (top) and final converged distributions (bottom) from AS and PAS, both with linear potential flow indicator $V'_{r,19}$. The copula in the name of the PAS results is the used model in the last iteration.

the maximum absolute distribution difference between iterations, and ϵ_2 for the coefficient of variation (COV) of the one-hour MPM. We used $\epsilon_2 = 0.1$ for all cases (standard deviation 10% of mean). The ϵ_1 value is case-specific, as it shares the units of the predicted quantity. For consistency across cases, we selected round numbers approximately 1%–2% of the maximum HF value in each validation dataset: $\epsilon_1 = 0.2$ m wave crest height for Case 1, 2×10^7 Nm VBM for Case 2, 200 kN green-water force for Case 3, and 20 kN slamming force for Case 4. These limits are more lenient than those used for case 1 and 2 with AS in van Essen and Seyffert (2025a). To ensure a fair comparison, we applied these to both AS and PAS. Section 5.6 includes a brief discussion of the sensitivity of the results for these stopping criteria.

The resulting converged distributions from AS and PAS for the wave impact cases are presented in Fig. 6 (cases 3a,b) and Fig. 7 (cases 4a,b), together with the iterative convergence of the MPM for each case. The corresponding figures for the simpler cases 1 and 2 can be found in Appendix D. The performance of the methods across all cases is summarised in Fig. 8, using the metrics defined in Section 4. As a reminder, M_1 measures efficiency relative to the HF ground truth, M_2 the deviation of the predicted HF MPM, and M_3 the maximum deviation

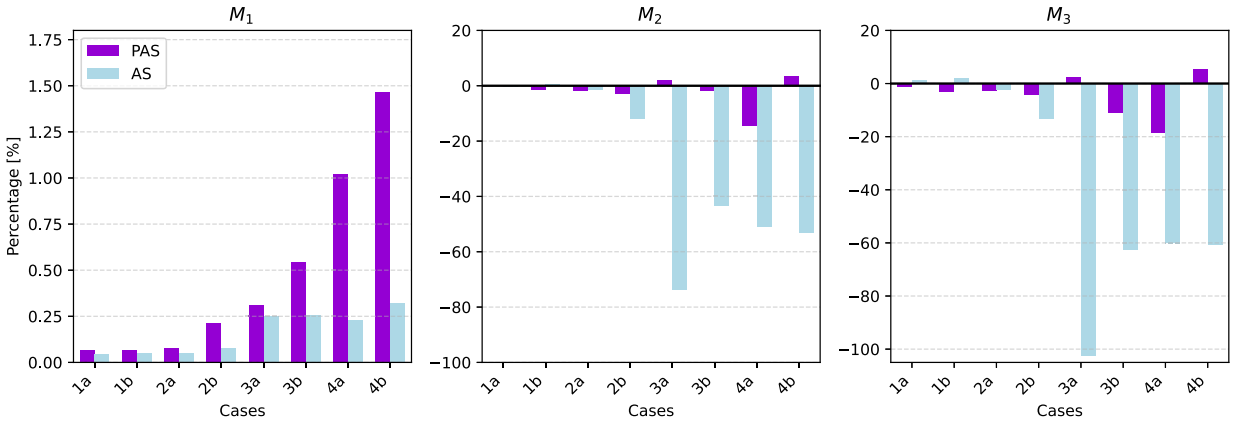


Fig. 8. Performance metrics M_1, M_2 and M_3 defined in Section 4 for all considered cases.

Table 4
Summary of the results for all cases.

	Case 1a	Case 1b	Case 2a	Case 2b	Case 3a	Case 3b	Case 4a	Case 4b
Predicted one-hour MPM at convergence								
unit	m	m	Nm	Nm	kN	kN	kN	kN
HF MCS[a]	10.53 (U95% 0.10)		1.10×10^9	9.00×10^8	1.85×10^3	4.92×10^3	1.71×10^3	2.44×10^3
AS	10.55	10.76	1.06×10^9	8.06×10^8	0.55×10^3	2.79×10^3	0.85×10^3	1.14×10^3
PAS	10.59	10.53	1.11×10^9	9.44×10^8	1.78×10^3	4.84×10^3	1.46×10^3	2.56×10^3
Required number of HF samples for convergence								
HF MCS[b]	16,364	16,364	14,359	14,359	15,198	10,900	15,198	10,900
AS	7	8	7	11	38	28	35	35
PAS	15	15	15	31	47	59	155	160
LF duration to simulate (analytical/linear potential flow)								
HF MCS	–	–	–	–	–	–	–	–
AS	50 h	50 h	30 h	30 h	34:49 h	23:43 h	34:49 h	23:43 h
PAS	50 h	50 h	30 h	30 h	34:49 h	23:43 h	34:49 h	23:43 h
Expected HF duration to simulate (CFD/experiments)								
HF MCS	50 h	50 h	30 h	30 h	34:49 h	23:43 h	34:49 h	23:43 h
AS (estimate[c])	~2 min	~3 min	~2 min	~4 min	~13 min	~9 min	~12 min	~12 min
PAS (estimate[c])	~5 min	~5 min	~5 min	~10 min	~16 min	~20 min	~52 min	~53 min

[a] Ground truth.
[b] All waves.
[c] Assuming HF events with a duration of 20 s each as discussed in Section 5.

between predicted and ground-truth distributions. The ground truth for each case is always derived from the HF validation MCS datasets. The predicted MPM at convergence is also included in Table 4. The MPM data underlying Fig. 8 differs slightly from the table, as the table shows the results for the converged iteration, whereas the MPM used for M_2 is averaged over the last 10 iterations at convergence (see Section 4).

5.1. Accuracy

The figures in Appendix D and Table 4 show that AS and PAS both provide accurate predictions for cases 1 and 2. PAS outperforms AS for both cases, but still does not predict the true tail of the sagging distribution (case 2a in Fig. 15(b)) well. M_2 in Fig. 8 shows that PAS predicts MPM values within 2% error for these two cases. Similar results were found in van Essen and Seyffert (2025a) for AS, where the USMV acquisition function produced slightly more accurate results than the current AS with MVPD for Case 2. However, the results for cases 3 and 4 clearly highlight the advantage of PAS over AS when combined with a linear potential flow indicator. Although PAS also relies on a screening indicator, the inclusion of probabilistic copula fitting substantially reduces the bias associated with indicator quality. This is particularly important when the LF and HF order statistics are not very similar, a situation in which AS performs poorly while PAS remains robust. The M_2 values show that PAS is able to predict the target extreme MPM values with an accuracy (averaged over the last 10 iterations before convergence) within 2% for both variations of case 3, and within 15% for both variations of case 4, whereas AS shows consistently non-conservative deviations of up to 75%. With the current settings, PAS predicts the one-hour MPM over all four cases within 15% accuracy at convergence. Similar results are observed for M_3; the maximum deviations of the PAS distributions compared to the ground truth are within 19% of the MPM for all test cases. It should be kept in mind that the present results rely on combining both PAS and AS with linear potential flow indicators, offering a computationally cheap way to examine these extreme responses.

5.2. Uncertainty

When comparing the extreme values from the validation datasets with those obtained using (P)AS, their respective uncertainties must be considered. The distributions in Figs. 6, 7, 14 and 15 also show the U95% sample uncertainty from the fitted copula model for PAS, and the U95% MF-GPR uncertainty for AS. These uncertainty measures are not directly comparable, but they show that the sample uncertainty from PAS is small around P_{exp} for all eight case variations (whereas the MF-GPR uncertainty band from AS is much larger).

For cases 1 and 3, the sample variability across multiple realisations of the same wave condition is available for the validation datasets.

This is not directly applicable to the present validation, as the LF simulations reproduce only one wave realisation (see Section 4); instead, the relevant uncertainty is the repeatability of the experimental HF measurements. Although not directly available, previous studies (e.g., van Essen, 2021 on wave impact loads measured in a seakeeping basin) indicate that this uncertainty is comparable to or smaller than the sample variability. We therefore compare the PAS uncertainty with the HF sample variability across multiple realisations of the same wave condition to assess whether our validation is meaningful.

For case 1, the HF validation sample uncertainty is 0.1 m (~1% of the true MPM value, see Table 4 and van Essen and Seyffert (2025a)). The U95% sample uncertainty of the PAS results at the converged iteration is 0.0 m (case 1a) and 0.07 m (case 1b), which corresponds to ~0-0.7% of the true MPM values. For green water case 3, the root-mean-square errors (RMSE) of extreme values in the HF validation data are detailed in van Essen et al. (2023, 2024), with the latter showing convergence to well below 0.1% of the one-hour MPM in both test conditions. Assuming U95% $\approx$ 1.96 RMSE, this corresponds to an uncertainty below 0.2% of the MPM. By comparison, the U95% uncertainty of the one-hour MPM predicted by PAS is 176 kN (Case 3a) and 722 kN (Case 3b), roughly 10% and 15% of the true MPM values, respectively. The uncertainty of extreme values in the HF validation data was not assessed for Cases 2 and 4, but it is expected to lie between that of Cases 1 and 3 for Case 2, and to be comparable to Case 3 for Case 4.

Because the uncertainty in the extreme values of the validation data is small and of similar or lower order than that of the PAS results, the validation comparison is meaningful.

5.3. Efficiency

When accuracy would be the only metric for the performance of an EVPM, MCS would always be preferred. The objective of PAS is to make accurate predictions *in a more efficient way*. M_1 in Fig. 8 shows that both AS and PAS reduce the required HF samples to under 1.5% of those needed for MCS across all cases. PAS is the more reliable method for accurate and efficient predictions, especially for complex cases. It requires fewer than 0.2% of HF events compared to MCS for Cases 1a,b and 2a,b, and less than 1.5% for Cases 3a,b and 4a,b. The exact number depends on the stopping criteria, which are case-specific; optimal limits for new cases depend on design requirements, acceptable risks, target quantiles, and possibly class society regulations. For simpler problems (Cases 1 and 2), AS is nearly as accurate and more efficient, making it a viable alternative.

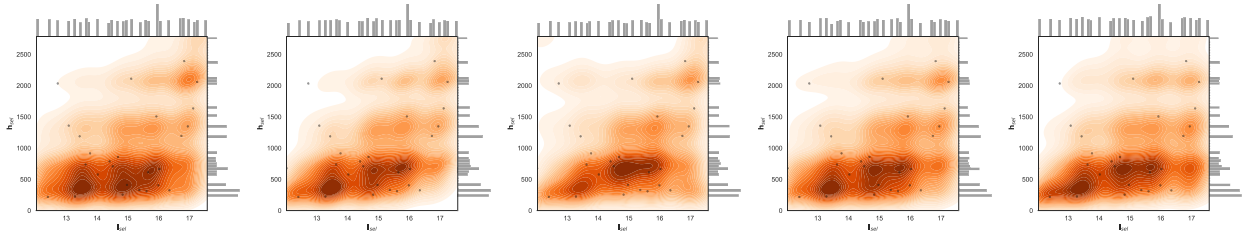
In all cases, this greatly reduces the number of HF events compared to MCS or experiments. Assessing PAS feasibility for a new vessel requires translating these numbers into simulation durations. As Section 4 notes, CFD was not performed for the selected HF events in the present validation. We can therefore only estimate the required durations. However, extensive literature shows that HF CFD can accurately reproduce wave impact loads when it closely matches experimental wave and ship motion data. The duration per event varies in these studies, ranging from $4T_p$ (Gramstad et al., 2023), 50 s (Bunnik et al., 2015), 35 s (Bandringa and Helder, 2018), ~20 s (Bandringa et al., 2020), to even as short as 3 s (Pákozdi et al., 2022). There are also studies that initialise the CFD simulations from an LF tool instead of experiments (as done in a screening method); this can be done with 52 s (Guth et al., 2023) or 12 s (van Essen et al., 2021) event durations. When fully non-linear wave kinematics and ship motions are provided by the LF tool, only short simulations per event are needed. In contrast, longer durations are required if waves and motions must be initialised from linear wave elevations or further from the structure. In our study, it is reasonable to assume that the waves in CFD can be initialised close to the structure, and that the ship motions can be imposed from the linear potential flow simulations. This should make it possible to reduce the duration per HF event to ~20 s (following the procedure in the green water impact study Bandringa et al., 2020). Using this assumption, we can translate the required number of HF events to HF simulation durations (see Table 4). This results in total durations to simulate with HF CFD or experiments in the order of 2 to 53 minutes, which seems feasible in the design stage of a ship. In comparison, the HF validation datasets used between 23 and 50 h of HF experiments. Based on this, it is estimated that the PAS converged results can be obtained by analysing between 0.04% and 2.2% of the HF MCS durations with an HF tool across all cases. Note that these percentages are roughly twice the M_1 values, because with typical wave periods of around 10 s, a 20 s event corresponds to two wave encounters.

The computational time required for these simulations (in CPU hours, CPUh) also depends on the HF tool used to simulate these 2-53 minutes, as well as the chosen grid size, domain, and time-step settings (in case of CFD) for each case. When using PAS, the full MCS duration must be simulated with an LF tool, but with frequency-domain potential flow or analytical methods (as in all cases here) the cost is negligible compared to HF simulations. For example in cases 2 to 4, the frequency-domain `SEACAL` calculations (see Section 2.8, 1–2 speeds and 1–2 headings) required approximately 0.25 CPUh per case on a standard laptop, with analytical linear wave simulations requiring even less. Conservatively, we estimate a maximum of 1 CPUh for LF simulations, including the transformation from frequency to linear time domain. Running PAS itself requires at most 1 min per iteration (0.017 CPUh), resulting in a total of approximately 2–3 CPUh for all iterations. Overall, the expected computational cost is roughly 1 CPUh for LF simulations, 2-3 CPUh for PAS, plus the time for HF CFD or experimental simulations, which typically cover up to one hour of simulated time divided over 7–160 events. The HF simulations dominate the total cost. A future study implementing the full PAS workflow with HF simulations will enable a more precise estimate.

5.4. Possible error sources

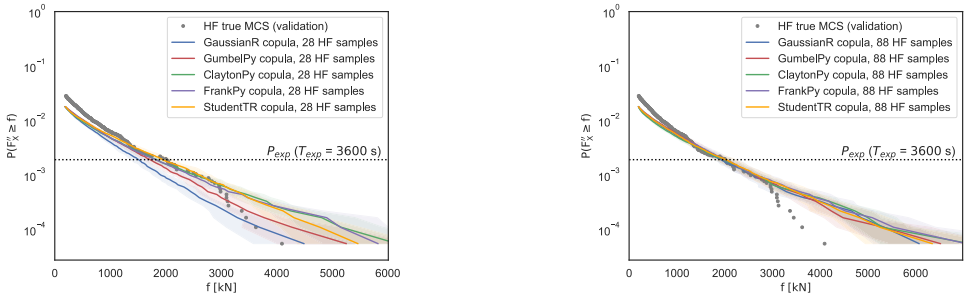
PAS is not guaranteed to converge to the exact HF exceedance distribution, even with increasing LF-HF samples. This is inherent to exceedance probability estimation, which is based on binary events (exceeds or not), and is therefore very sensitive to individual observations, especially in the tails. For instance, a single low-indicator but high-HF-load value can substantially shift the HF exceedance distribution, as also observed in AS and other screening studies (Bunnik et al., 2018, 2019; van Essen and Seyffert, 2025a). Consequently, adding more samples does not always improve tail estimates monotonically. New observations change the empirical pseudo-observations, the fitted copula and marginal parameters, and redistribute exceedance probabilities. Finite-sample effects mean that these updates can introduce noise, slow convergence, or shifting tail behaviour rather than refining it. In this context, convergence means stabilisation of the fitted model under the chosen sampling and modelling assumptions, rather than asymptotic convergence to the true distribution. This risk is lower with a copula model than in sampling approaches without it, as the copula introduces a probabilistic joint dependence structure rather than a deterministic one. The outlier removal feature also dampens this effect.

A risk when combining copula modelling with adaptive sampling is bias introduced by using targeted samples. In general, adaptive sampling can produce non-*i.i.d.* data (independent and identically distributed) and deform dependence estimates by over-representing certain regions. Copula models treat the transformed marginals as if they were draws from a uniform distribution (see e.g., Nelsen, 2006; Genest and Favre, 2007). Small or concentrated samples can violate this assumption and increase bias and instability. In the present case, this risk is mitigated by using an acquisition function that distributes samples approximately uniformly over the logarithm of the LF exceedance probability space, rather than focusing on tails or uncertainty. This leads to a balanced coverage over the probability levels.



(a) Gumbel (AIC -5.1) (b) Gaussian (AIC -3.9) (c) Student-T (AIC -2.8) (d) Frank (AIC -1.7) (e) Clayton (AIC -0.7)

Fig. 9. Different copula models fitted to the LF-HF sample data at converged PAS iteration 47 (51 HF samples of which 28 true positives), for green water case 3a. The plots are ordered by lowest AIC (best-fitting) on the left to highest AIC on the right.



(a) It. 47 (51 HF samples of which 28 true positives) (b) It. 150 (154 HF samples of which 88 true positives)

Fig. 10. Exceedance probability distributions (mean and U95%) resulting from 10 conditional draws from the copula models fitted to the LF-HF samples at iteration 47 and 150, for green water case 3a.

Large differences in order statistics between the LF indicator and the target HF variable (i.e., a poor indicator choice) can also distort the fitted dependence structure, especially in the tails, and lead to biased HF exceedance estimates. For example, in the green water case, a large bow flare may deflect water to the sides. A high RWE (LF indicator) would then not correspond well to the green water load on the deck (HF load). Other potential errors come from imperfect copula choices and uncertainty in picking the best model from a limited set, errors in marginal distribution fitting and interpolation, Monte-Carlo sample variability in conditional simulation, and the fact that the stopping criterion reflects stability between iterations rather than consistency with the true exceedance distribution (necessary because the true distribution is unknown in a real design case).

5.5. Sensitivity to copula model

Since PAS results are largely driven by the fitted copula, and this step differs from AS, we evaluated the sensitivity of the results to the selected copula model in more detail.

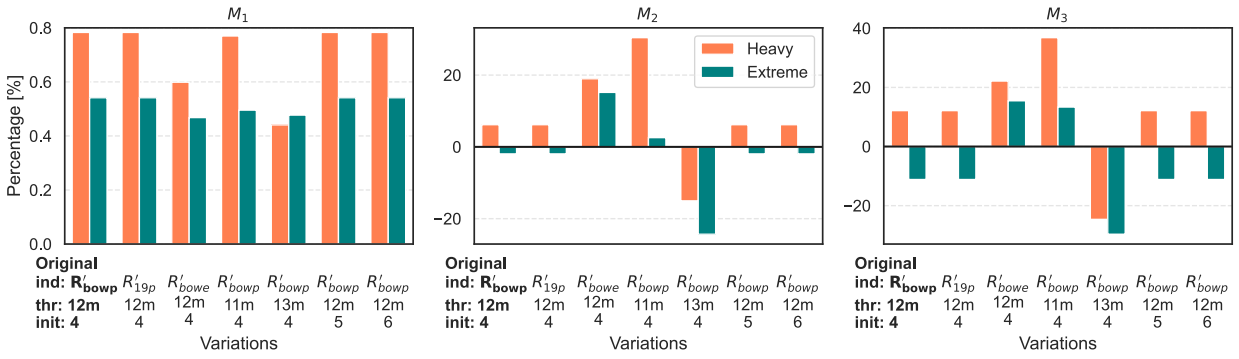
Firstly, convergence of the copula selection process is discussed in Appendix F. As explained in Section 2.2, the AIC-based copula selection in PAS adapts to the currently available HF samples rather than seeking a single 'true' copula model. The preferred copula may therefore change as new samples are added. The appendix shows that, over all cases, the Gaussian copula is most often selected. This is likely due to its robustness with sparse data combined with the stabilising effect of the AIC threshold. When an alternative model is chosen, it is typically the Clayton copula, consistent with the observed dominance of lower-tail dependence (especially for cases 3 and 4, see Fig. 16).

Secondly, we assess the sensitivity of predicted distributions to the copula model, by analysing green water Case 3a at iteration 47. Fig. 9 shows the 28 true positive samples at this iteration, together with the five candidate copula models fitted to these samples. This shows that the fitted copulas are quite similar overall, but there are also differences. The plots are ordered by AIC value from low (left, best-fitting) to high (right); the Gumbel and Gaussian copulas fit the dataset best at this iteration. Fig. 10(a) shows the drawn conditional distributions from each of these models at the same iteration. This shows that the distribution differences are limited at P_{exp}, but the tails can be significantly different. Selecting the most-likely copula model per iteration instead of *a priori* selecting one model can therefore pay off in the prediction of the extremes. The same is plotted for PAS iteration 150 in Fig. 10(b). This shows that the difference in the tail reduces with an increasing number of samples. In other words, for a large number of samples, all copula models seem to converge to a similar distribution.

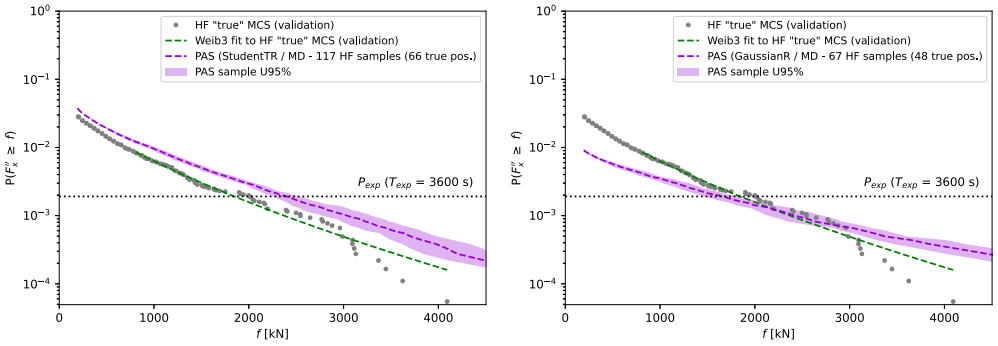
5.6. Sensitivity to other parameters

In addition to the copula model, PAS includes many internal parameters and user-defined inputs that can be adjusted. For most settings, we selected values that perform robustly across the considered problems, as demonstrated by the four cases in this study. Here, we present a sensitivity analysis of the most influential choices: the LF indicator, LF threshold and initial sampling, and we further discuss the impact of the acquisition function and stopping criteria.

First, we consider the sensitivity of the green water results (Case 3) for indicator choice, LF threshold, and initial sampling. We ran PAS with three LF indicators: bow RWE from potential flow ($R'_{\text{bow}p}$, the original case with added subscript p to emphasise its source), station-19 RWE from potential flow (R'_{19p}), and bow RWE from experiments ($R_{\text{bow}e}$). We also tested two alternative LF thresholds (11 m and 13 m, versus the original 12 m) and two additional initial sample sizes (5 and 6, versus 4), all selected as in Section 2.5. Results are summarised in Fig. 11, using a modified COV stopping criterion of 5% instead of 10%. While 10% was identified as the optimal choice across all four cases, a stricter criterion was required for case 3 alone to ensure convergence across all its variations; consequently, the results for the original case differ slightly from those shown in Fig. 6. A similar plot is shown for two variations of the slamming case in Fig. 12 (with the same COV criterion as used for Fig. 7). Here, we ran PAS with two LF



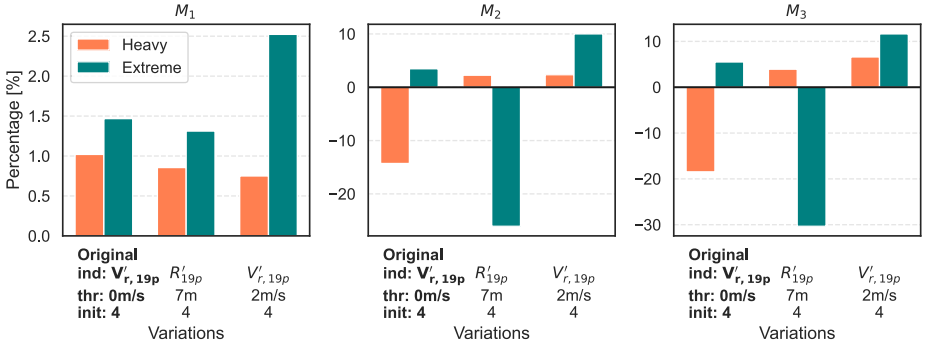
(a) Performance metrics M_1, M_2 and M_3.



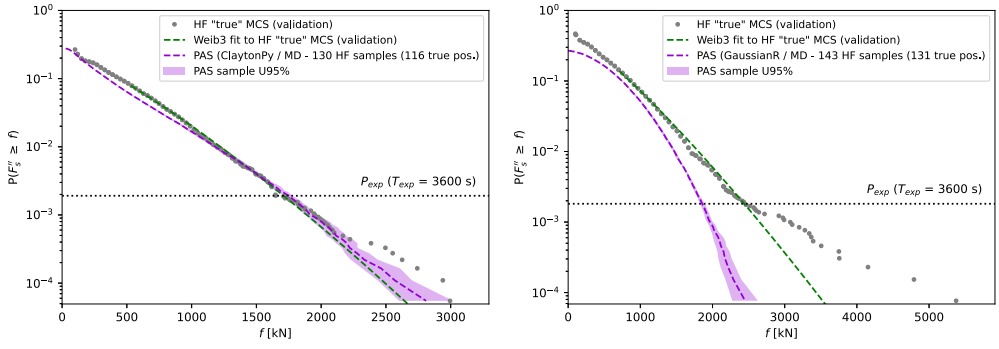
(b) PAS Heavy distribution, threshold 11 m

(c) PAS Heavy distribution, threshold 13 m

Fig. 11. Sensitivity of the green water results for indicator choice, LF threshold choice and number of initial samples (with modified COV criterion 5%, so results of the original case differ slightly from Fig. 6), expressed as performance metrics, with some example distribution plots. Ind = indicator, thr = LF threshold, init = number of initial samples, and R_{Xp} = RWE at location X from potential flow, R_{Xe} = RWE at location X from experiments.



(a) Performance metrics M_1, M_2 and M_3.



(b) PAS Heavy distribution, indicator R'_{19p}

(c) PAS Extreme distribution, indicator R'_{19p}

Fig. 12. Sensitivity of the slamming results for indicator choice and LF threshold choice, expressed as performance metrics, with some example distribution plots. Ind = indicator, thr = LF threshold, init = number of initial samples, and $V'_{r,Xp}$ = RRV at location X from potential flow, R'_{Xp} = RWE at location X from potential flow.

indicators: the RRV at station 19 from potential flow ($V'_{r,19_p}$, the original case) and the RWE at station 19 from potential flow (R'_{19_p}). We also varied the LF threshold from no threshold (original case) to 2 m/s. For both cases, all other settings were kept unchanged. The results shown are converged according to the combined criterion, implying some dependence on the stopping limits. Nevertheless, the comparison provides a clear indication of the sensitivities.

Based on the order-statistics assumption that is the foundation of screening-based methods, the choice of indicator is expected to strongly influence results. This was also observed in van Essen and Seyffert (2025a), where AS with a coarse-mesh CFD RWE indicator produced accurate green water distributions, whereas the potential-flow RWE indicator used here did not. The copula model in PAS was introduced to reduce sensitivity to indicator choice by modelling LF-HF relation probabilistically rather than deterministically, thereby limiting the influence of outliers. The three green water indicators in Fig. 11 support this idea: the difference in results between the two potential-flow indicators is small, indicating that the specific choice does not strongly affect the results. This is likely because these variables are internally statistically consistent. Interestingly, using the experimental bow RWE as indicator yields poorer results than the linear potential-flow RWE, probably due to measurement noise in the experimental signal. For the two slamming indicators in Fig. 12, we see something slightly different; both perform well in one case but less so in the other. This is also visible when comparing the predicted slamming distributions with indicator R'_{19_p} in Figs. 12(b) and 12(c) to those of the original indicator $V'_{r,19(p)}$ in Fig. 7. This suggests that neither indicator is fully adequate and that indicator choice is more critical for slamming than for green water. As discussed before, we did not know 'the best' indicators for our cases, but selected a good candidate indicator based on experience and earlier work. There could be even better indicators, with fewer false positives for these cases, especially for slamming (such as e.g., the Wagner model (Hermundstad and Moan, 2007) or momentum theory (Kapsenberg, 2018)).

Fig. 11 also shows the sensitivity of the green water results to the selected LF threshold. The original threshold of 12 m provides very good results under both test conditions, whereas lowering the threshold to 11 m or increasing it to 13 m both leads to deteriorated performance. Setting the threshold is challenging. A value that is too high creates unbalanced sampling that focuses on extreme events, reducing the quality of the copula fit to the full dataset. A value that is too low on the other hand complicates the definition of the HF distribution zero-crossing, and reduces the efficiency by increases the number of false negatives (though this effect is not yet pronounced for 11 m, see M_1 in Fig. 11). Choosing a threshold slightly above the physical threshold (in this case the freeboard of 10.5 m) appears to be a good compromise, with 12 m performing well. For the slamming case, we also introduced a rise velocity threshold of 2 m/s (see Fig. 12) to assess whether this sensitivity to threshold selection is also valid for other cases. This threshold similarly leads to an overestimation of the MPM and its distribution, as sampling becomes more focused on the tail of the distribution.

Fig. 11 also shows that the influence of the number of initial samples is negligible, as expected since they are defined analogously to the MD function.

In van Essen and Seyffert (2025b), we evaluated seven acquisition functions for AS. For PAS, coverage of the full exceedance probability domain is especially important (as explained in Section 2.4). While we did not repeat the full sensitivity study, we tested most previous functions: only the MD function consistently performed well with the copula, as most other functions focus on subregions. A variant of MD aiming for uniform coverage of the full domain on linear instead of logarithmic scale also worked, but was less efficient and occasionally produced badly-behaving tail predictions.

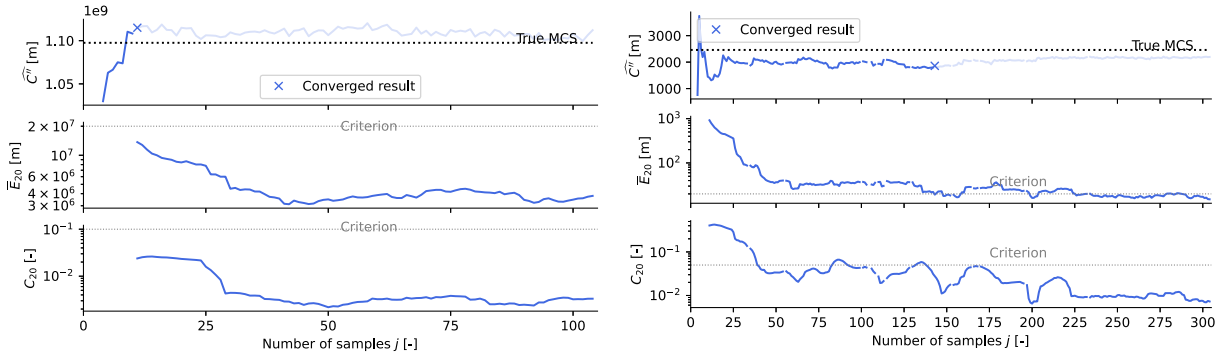
PAS results are sensitive to the stopping criterion settings: averaging windows K_1 and K_2 and convergence thresholds ϵ_1 and ϵ_2 (see Appendix C). Ideal settings depend on the case and are generally unknown: noisier LF-HF statistics may converge best with relatively lenient criteria and large averaging windows, while smoother cases quickly converge with strict limits and short windows. Long averaging windows generally improve stability of the convergence, but also slow it down (which is expensive if every extra sample requires a CFD simulation). For robustness across our four test cases, we chose $K_1 = K_2 = 20$, $\epsilon_1 \approx$1%–2% of the expected maximum HF value, and $\epsilon_2 = 0.1$ (10% COV of the MPM). These settings typically indicate convergence once the MPM and distribution stabilise over 20 iterations, but Fig. 13(b) shows that 'stable' predictions do not always reflect the true converged value; results may still shift with more iterations. The figure also shows that the current averaging window is short for sensitive cases like slamming (Fig. 13(b)) but more than sufficient for simpler cases (Fig. 13(a)). The chosen settings represent a practical balance.

5.7. Future work

In the present validation approach (see Section 4), the HF response is not obtained by explicitly running CFD simulations but is instead extracted directly from the validation dataset. As a result, the potential impact of CFD simulation errors on PAS performance is not accounted for. Two distinct aspects are involved: (i) the numerical accuracy of the CFD simulations themselves, and (ii) the statistical equivalence between LF wave events and their replication in HF CFD simulations. Both aspects are related to substantial research questions, beyond the scope of the present study. With respect to CFD accuracy (aspect i), some insight is available from earlier work, as discussed in Section 1.3. In particular, Bandringa et al. (2020) showed that CFD could successfully reproduce experimental green water loads for case 3a, supporting its suitability for capturing the relevant physics. As noted in Section 2.1, initialising HF event simulations in Step 6 of PAS using the LF screening results is not straightforward (aspect ii). No universally accepted method exists to correlate linear (or weakly non-linear) wave events with equivalent fully non-linear wave events, or to define the corresponding acceptable tolerances in wave parameters. Potential solutions include event-matching procedures (Johannessen and Lande, 2018; Gramstad et al., 2023) or using coarse-mesh CFD as an LF tool (van Essen et al., 2021). The first method is less suitable for sailing ships, and the second requires costly CFD. An alternative is to perform MCS with a fully non-linear wave-only tool (e.g., a higher-order spectral method, HOS) and then apply linear potential-flow screening to the resulting wave traces, allowing HF simulations to be initialised directly from the corresponding HOS events. A promising direction for future research is to perform the full validation of PAS including HF simulations, using a suitable CFD tool paired with this HOS-initialisation approach.

PAS could potentially be further improved and made more robust for broader applications, for example by using different marginal and copula models for tail versus body (to improve tail predictions in the simpler cases), selecting multiple new samples per iteration for parallel HF simulations, or incorporating copula fit quality over the LF range into the acquisition function. PAS could also use event frequencies instead of exceedance probabilities, or adopt a block-maximum/return-period approach instead of POT. This would improve interpretability, align with industry practice, and allow easier application to non-wave problems by referencing fixed-duration blocks rather than variable wave events. However, it would reduce efficiency, since each block spans many waves (e.g., 47 samples would require 47 block simulations instead of 47 wave events). Validating a block-maximum approach in a similar context as used here would require further research to define representative paired peaks, either as the block's maximum HF response or the HF response of the wave with the maximum LF value.

The MPM predicted in Step 12 of PAS corresponds to the $q = e^{-1} \approx 0.368$ quantile of the short-term distribution for linear Gaussian processes (see e.g., Ochi, 1990). Accordingly, the probability that the



(a) Hogging VBM (from Figure 15a).

(b) Slamming, Extreme condition, ind. R'_{19p} (from Figure 12).

Fig. 13. Stopping criteria and convergence limits for simple and complex example cases, both combined with an ϵ_1 criterion in the range of 1%–2% of the maximum HF value in the validation dataset (middle plot), and $\epsilon_2 = 0.1$ (bottom plot). Top plot shows MPM as a function of the number of HF samples.

maximum response over the target exposure duration exceeds the MPM is 62.3%. In offshore design, higher quantiles are often preferred. These can in theory be estimated by replacing P_{exp} in Eq. (12) with $P_q = 1 - q^{m/n}$ (DNV, 2019b), but this requires substantially longer LF MCS to ensure convergence. Further work is therefore needed to assess the method's applicability at higher quantiles.

As noted in Section 1.5, this study validates PAS only for short-term extremes. In practical applications, sea states and long-term statistics must also be handled, which could be addressed with PAS as well, with minor modifications (e.g., to Step 2). Alternatively, a version of the environmental contour method, as described in Winterstein et al. (1993), Huseby et al. (2013), Mackay et al. (2025), could be used. This approach is widely accepted in the offshore industry (e.g., DNV, 2019a), though it can sometimes fail to accurately characterise the joint distribution of environmental variables (Speers et al., 2024; Haselsteiner et al., 2021; de Hauteclocque et al., 2022). Another option is to use another adaptive sampling method for the long-term screening, as proposed by Gramstad et al. (2020), Wang et al. (2024).

Interpreting the significance of PAS wave impact extreme value errors (2%–15% in the MPMs of the green water and slamming case) in an engineering context is challenging. As discussed in Section 1.4, there is no universally accepted extreme value prediction framework for wave impact loads, and in practice the industry usually follows prescriptive standards or relies on direct analysis using experiments. Such approaches are typically combined with substantial safety margins and prediction of higher quantiles. Exact values of these margins are difficult to obtain, but they are generally understood to exceed the 2%–15% level associated with the MPM error reported here. The relationship between such margins and the statistical uncertainty of experimental extreme value estimates is, however, not clearly defined. Embedding the use of PAS and its associated uncertainties within a formal design context, including safety margins and workflow, is considered an important topic for future research.

As a framework for combining LF and HF model data, PAS has no inherent application limitations beyond those of the underlying LF and HF models. It could also be generalisable to non-linear extreme value problems in other fields (possibly with some small modifications to account for non-wave-based problems). Nonetheless, its performance should be validated again when applied to very different types of problems (e.g., in other fields).

6. Conclusions

Based on the present work, it can be concluded that the newly proposed extreme value prediction method PAS provides accurate and efficient estimates of exceedance distributions and extreme values for a range of non-linear wave-induced response problems, compared to the available full brute-force MCS datasets. The method uses a probabilistic approach with copulas to model the joint distribution of low-fidelity (LF) and high-fidelity (HF) variables. This enables the use of efficient linear potential-flow indicators in the LF stage, even for strongly non-linear loads. PAS then predicts extreme HF responses by statistically linking these LF indicators to a limited set of HF samples, effectively learning the dependence between the two fidelity levels. Results are iteratively updated using adaptive sampling, until convergence is reached. PAS with optimal settings achieves a most probable maximum (MPM) accuracy of 2%–15% for all considered cases in the present study, including non-linear waves, vertical bending moments, green water impact loads, and slamming loads. In addition, PAS achieves this performance very efficiently, requiring in the order of 1%–3% of the high-fidelity simulation time needed for conventional MCS over all cases. While PAS is similarly accurate and slightly less efficient than its predecessor AS for the simpler cases (non-linear waves and vertical bending moments), it proves significantly more reliable for the more complex cases (green water and slamming loads). These results demonstrate that PAS can reliably reproduce the statistics of both weakly and strongly non-linear extreme load problems, while significantly reducing the associated computational cost.

PAS is largely insensitive to the choice of copula, though tail predictions can vary for small sample sizes; selecting the best-fitting copula from five candidates at each iteration helps mitigate this sensitivity. PAS green water predictions are less sensitive to indicator choice than those from AS, performing well with different potential-flow indicators. Slamming predictions are more sensitive, suggesting potential for improved indicators. For cases with a physical LF threshold below which many HF loads are zero, the selected input LF threshold strongly affects results: too high reduces the quality of the copula fit across the full exceedance range, too low decreases efficiency. A value slightly above the threshold is a good compromise, while zero is best when no physical threshold exists. PAS is also sensitive to the acquisition function (only the MD function reliably covers the full exceedance probability domain), and to the stopping criteria, which must balance accuracy for complex cases with efficiency for simpler ones. The number of initial samples has negligible effect under the current selection method.

The present study validates the statistical PAS framework. PAS does not impose explicit theoretical limits on structure type or wave conditions. Its applicability is limited by the availability of a suitable LF indicator with order statistics comparable to the HF response. The cases studied here demonstrate feasibility for the investigated scenarios but do not define hard boundaries of the method. Further work should focus on validating the full procedure including CFD load simulations, assessing long-term extremes and extremes at higher quantiles, embedding the use of PAS within a formal design context, and possibly testing it against strongly non-linear problems beyond wave-related scenarios.

CRediT authorship contribution statement

Sanne M. van Essen: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Harleigh C. Seyffert:** Writing – review & editing, Supervision, Funding acquisition.

Declaration of generative AI / AI-assisted technology in the manuscript preparation process

During the preparation of this work the authors used ChatGPT in order to rephrase some individual sentences. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

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.

Acknowledgements

This publication is part of the project "Multi-fidelity Probabilistic Design Framework for Complex Marine Structures" (project TWM.BL.019.007) of the research programme "Topsector Water & Maritime: the Blue route" which is (partly) financed by the Dutch Research Council (NWO). We thank the Cooperative Research Ships (CRS) for permitting us to use the `PySeaWave` package, the `SEACAL` and `PRETTI_R` programs, and the experimental data of cases 3 and 4. We also thank Cees de Valk of KNMI for suggesting copula modelling as a way to reduce bias in sampling methods. Finally, we thank the anonymous Ocean Engineering reviewers as well as Tormod Landet and Guillaume de Hauteclocque of the CRS SPEC working group for their useful comments on the first version of this paper.

Appendix A. Details of the Adaptive Screening (AS) method

Here, the steps of AS are briefly summarised. This is a shortened reproduction of the full description and formulations in van Essen and Seyffert (2025a). Steps 1 to 6 of AS are identical to Step 1 to Step 6 described in Section 2.1. The next three steps are different. They are described here in general terms (for details, see van Essen and Seyffert, 2025a). **Step 7** estimates the sample HF distribution, by assuming that the order statistics of $\mathbf{l}^{\text{sel}}$ and $\mathbf{h}^{\text{sel}}$ are identical. This is a critical screening assumption that only works if a suitable indicator signal is chosen. It indicates that the HF distribution $[\mathbf{d}_H^{\text{sel}}, \mathbf{h}^{\text{sel}}]$ can be estimated using $\mathbf{d}_H^{\text{sel}} \approx \mathbf{d}_L^{\text{sel}}$. **Step 8** defines an exceedance probability range $\mathbf{d}^* \in [0,1]$ around P_{exp}, over which to estimate the HF distribution. **Step 9** constructs the surrogate HF distribution $\mathbf{h}^*$ over $\ln(\mathbf{d}^*)$ using single- or multi-fidelity Gaussian Process Regression (GPR). With single-fidelity GPR, we use the HF sample dataset $[\ln(\mathbf{d}_H^{\text{sel}}), \mathbf{h}^{\text{sel}}]$ as input. With multi-fidelity MF-GPR, we use the same HF sample dataset $[\ln(\mathbf{d}_H^{\text{sel}}), \mathbf{h}^{\text{sel}}]$ as input, *and* LF dataset $\left[\ln(\mathbf{d}_L^{\text{mcs}}), \mathbf{l}^{\text{mcs}}\right]$ from Step 3. After that, steps 10 to 12 of AS are identical to Step 12 to Step 14 of PAS again. As in Section 2.4, AS adds one new HF sample per iteration using an acquisition function. In van Essen and Seyffert (2025a), the USMV function balanced uncertainty and upper-tail sampling. Later, van Essen and Seyffert (2025b) showed that MVPD, which balances sampling around the target probability and the upper tail, is more efficient. When AS is used in this study it is combined with two initial samples, MF-GPR and MVPD (Eq. (20), where $\overline{\mathbf{h}}_n^*$ is the normalised mean HF prediction from the previous iteration and $\mathbf{d}^*$ are the corresponding exceedance probabilities). The closest MCS sample is then selected and added to the LF pool for the next iteration, as in PAS.

$$p_{\text{new}}^{\text{mvpd}} = \arg\max\left[\overline{\mathbf{h}}_n^* \cdot \left(\frac{\mathbf{f_2}}{\max(\mathbf{f_2})}\right)\right] \quad \text{where:} \quad \begin{cases} \mathbf{f_1} = 1 - (\ln(\mathbf{d}^*) - \ln(P_{\text{exp}}))^2 \\ \mathbf{f_2} = \mathbf{f_1} - \min(\mathbf{f_1}) \end{cases} \tag{20}$$

Appendix B. Copula basics

B.1. General theory

Copulas are multivariate distribution functions with uniform marginals on the interval $[0,1]$. If F is a joint cumulative distribution function (CDF) with marginals $F_1, F_2, \ldots, F_d$, then there exists a copula C defined in Eq. (21). If the marginals are continuous, then the copula C is unique and can be written as Eq. (22) (Sklar's theorem, Sklar (1959)). Copulas model dependence between random variables by separating marginal distributions from the dependence structure, allowing flexible joint modelling, particularly in the tails. Compared to empirical distributions, they require less data and support reliable interpolation and extrapolation, making them well suited for surrogate modelling, uncertainty quantification, and multivariate extreme value analysis.

$$F(x_1, x_2, \ldots, x_d) = C(F_1(x_1), F_2(x_2), \ldots, F_d(x_d)) \tag{21}$$

$$C(u_1, u_2, \ldots, u_d) = F(F_1^{-1}(u_1), F_2^{-1}(u_2), \ldots, F_d^{-1}(u_d)) \tag{22}$$

B.2. Copula families and formulations

Copula families are commonly grouped into *elliptical* and *Archimedean* copulas. Elliptical copulas, derived from distributions such as the Gaussian or Student-T, model symmetric dependence, while Archimedean copulas use a generator function and allow asymmetric and tail-dependent behaviour. This work uses Gaussian and Student-T copulas from the former class, and Clayton, Gumbel, and Frank copulas from the latter, restricted to the bivariate case. Overall dependence between variables is measured by Kendall's tau τ, which can be estimated empirically and related to copula parameters via the formulations below. Similarly, lower and upper tail dependence are quantified by the coefficients λ_L and λ_U. For example, $\lambda_U = 0.81$ indicates an 81% probability that one variable is extreme given that the other is.

The **Gaussian** copula (Eq. (23)) is defined using the standard normal CDF Φ and the bivariate normal CDF Φ_ρ, with correlation parameter $\rho \in [-1,1]$. It models symmetric dependence with no tail dependence; $\lambda_L = \lambda_U = 0$, and $\tau = \frac{2}{\pi}\arcsin(\rho)$. The **Student-T** copula (Eq. (24)) uses the univariate and bivariate Student-T CDFs, t_ν and $t_{\rho,\nu}$, parameterised by correlation $\rho \in [-1,1]$ and degrees of freedom $\nu > 0$. It exhibits symmetric upper and lower tail dependence for finite ν, with $\tau = \frac{2}{\pi}\arcsin(\rho)$. The **Clayton** copula (Eq. (25)) is parameterised by $\theta > 0$ and captures lower-tail dependence only, with $\lambda_U = 0$, $\lambda_L = 2^{-1/\theta}$ and $\tau = \theta/(\theta+2)$. The **Gumbel** copula (Eq. (26)) has parameter $\theta \geq 1$ and captures upper-tail dependence only, with $\lambda_L = 0$, $\lambda_U = 2-2^{1/\theta}$ and $\tau = 1-1/\theta$. The **Frank** copula (Eq. (27)) is defined by a dependence parameter $\theta \in \mathbb{R}\backslash 0$, where the sign of θ determines positive or negative dependence. It has no tail dependence ($\lambda_L = \lambda_U = 0$) and

$\tau = 1 - \frac{4}{\theta}\left(1 - D_1(\theta)\right)$, where $D_1(\theta) = \frac{1}{\theta}\int_0^\theta \frac{t}{e^t - 1}, dt$ is the Debye function.

$$C_\rho^{\text{gaussian}}(u, v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)) \tag{23}$$

$$C_{\rho,\nu}^{\text{Student-T}}(u, v) = t_{\rho,\nu}(t_\nu^{-1}(u), t_\nu^{-1}(v)) \tag{24}$$

$$C_\theta^{\text{clayton}}(u, v) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-1/\theta}, \quad \theta > 0. \tag{25}$$

$$C_\theta^{\text{gumbel}}(u, v) = \exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right), \quad \theta \geq 1. \tag{26}$$

$$C_\theta^{\text{frank}}(u, v) = -\frac{1}{\theta}\ln\left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right), \quad \theta \in \mathbb{R} \setminus \{0\}. \tag{27}$$

Appendix C. PAS stopping criterion

The stopping criterion in PAS has two components: one monitors the convergence of the predicted distribution shape, and the other monitors the convergence of the resulting MPM value. The criterion ignores iterations where a false positive sample was added. The first component is based on the mean absolute difference $E(j)$ between successive predicted distributions over an exceedance probability range, $\mathbf{h}^*_{\text{ran}}(j-1)$ and $\mathbf{h}^*_{\text{ran}}(j)$ for iteration j. This range is taken between $P_{\text{exp}}/2$ and $P_{\text{exp}} \cdot 2$, to focus on the distribution shape around P_{exp}. Part one of the stopping criterion $\overline{E}_{K_1}(j)$ is then defined in Eq. (28), by taking the average $E(j)$ over the last K_1 iterations. This is done to ensure a smooth convergence. When $j < K_1$, all available iterations are used, but we set a convergence threshold at $K_1/3$ (convergence can only be declared after this number of iterations).

$$\overline{E}_{K_1}(j) = \frac{1}{\psi}\sum_{i=\chi}^{j} E(i)$$

$$\text{where:} \quad \begin{cases} E(j) = \text{mean}\left|\mathbf{h}^*_{\text{ran}}(j) - \mathbf{h}^*_{\text{ran}}(j-1)\right| & \\ \chi = 1 \text{ and } \psi = j & \text{for } j = 1, 2, \ldots, K_1 - 1 \\ \chi = j - K_1 + 1 \text{ and } \psi = K_1 & \text{for } j \geq K_1 \end{cases} \tag{28}$$

The second part is based on a convergence criterion on the coefficient of variation (COV) of the MPM value over the last K_2 iterations: $C_{K_2}(j)$. This is expressed in Eq. (29), where $\widehat{H}(j)$ is the MPM value predicted in iteration j. Again, we took the COV over the available iterations when $j < K_2$.

$$C_{K_2}(j) = \frac{\sigma_{K_2}(j)}{\mu_{K_2}(j)}$$

$$\text{where:} \quad \begin{cases} \mu_{K_2}(j) = \frac{1}{\psi}\sum_{i=\chi}^{j}\widehat{H}(i) & \\ \sigma_{K_2}(j) = \sqrt{\frac{1}{\psi}\sum_{i=\chi}^{j}\left(\widehat{H}(i) - \boldsymbol{\mu}_{K_2}(i)\right)^2} & \\ \chi = 1 \text{ and } \psi = j & \text{for } j = 1, 2, \ldots, K_2 - 1 \\ \chi = j - K_2 + 1 \text{ and } \psi = K_2 & \text{for } j \geq K_2 \end{cases} \tag{29}$$

The total stopping criterion $S(j)$ is provided in Eq. (30), where $K_1 = K_2 = 20$ limit the influence of outliers and the minimum number of iterations for which convergence can be detected. Stopping criteria ϵ_1, ϵ_2 are chosen case-dependently, using acceptable tolerances for the problem at hand.

$$S(j) = \begin{cases} \text{stop} & \text{if } \left(\overline{E}_{50}(j) < \epsilon_1\right) \cap \left(C_{50}(j) < \epsilon_2\right) \\ \text{continue} & \text{otherwise} \end{cases} \tag{30}$$

Compared to the stopping criterion used for AS in van Essen and Seyffert (2025a), we removed the rejection criterion for distribution shapes that violate distribution assumptions (as drawing from the copula model in Step 9 of PAS always produces a proper distribution), we added a range over which Eqs. (28) and (29) are calculated (this was natural in AS because most steps considered such a range, but not in PAS), we used the mean absolute difference instead of the maximum absolute difference in Eq. (28), we used $K_1 = K_2 = 50$ instead of 20 (to ensure smoother convergence), and we added the convergence threshold.

Appendix D. Details cases 1 and 2

D.1. Case 1 (weakly non-linear): second-order wave crest heights

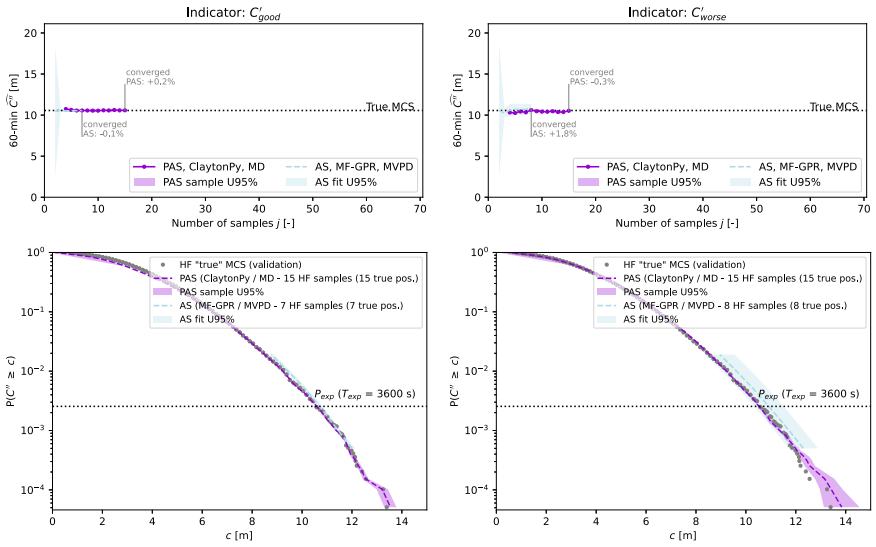
Case 1 from van Essen and Seyffert (2025a,b), with related datasets in van Essen and Seyffert (2025e,c), addresses a weakly non-linear problem. We predict extreme values of 'HF' second-order wave crest heights C'', in long-crested steep waves (condition A in Table 3). The HF validation dataset consisted of analytical second-order wave time traces according to Sharma and Dean (1979). Two LF indicators were defined: linear Gaussian wave crests C'_{good} (case 1a) and crests in the same linear Gaussian waves including an additional noise wave system C'_{worse} (case 1b). The LF MCS of Step 2 was done for 50 h duration. Since the HF variable is analytically tractable, it was easy to generate long HF ground truth time traces for this example case. Appendix E.1 shows a scatter plot of LF indicator versus HF validation response peaks. This plot shows that the LF and HF order statistics align closely, but not perfectly. To generate a new HF sample in each iteration, we used the matched LF and HF peaks. When an event was selected based on its LF indicator value (in Step 5 or Step 13), the corresponding HF value was drawn from the matched LF-HF peaks. This is only possible because we have an analytically traceable HF variable; in reality, each iteration would require a new HF event calculation. Details of the simulation methods are provided in Appendix E.3. We aim to predict the HF one-hour MPM value $\widehat{C''}$. We did not use an LF threshold value of PAS; no false positives were expected for this case.

D.2. Case 2 (weakly non-linear): vertical bending moments

Case 2 from van Essen and Seyffert (2025a,b), with related datasets in van Essen and Seyffert (2025e,c), studies a moderately non-linear problem. We predict extreme values of hogging vertical bending moments (VBM) on a 190 m ferry in extreme irregular head waves (condition B in Table 3). The HF VBM in hogging V''_{hog} and sagging mode V''_{sag} were generated using a non-linear (Froude–Krylov) potential flow simulation, while the LF VBM in hogging V'_{hog} and sagging mode V'_{sag} were generated using a linear potential flow simulation. The hogging peaks were defined by the zero up-crossing troughs in these VBM simulations; the sagging peaks by the zero up-crossing peaks. Unlike in case 1, this HF response is not analytically tractable, but 30-hour HF MCS non-linear simulations to produce validation material were feasible. Details of the simulation methods are provided in Appendix E.3. The LF-HF peak scatter plots are shown in Appendix E.1. This shows that, as expected, the LF and HF VBM peaks have less similar order statistics than those of the first- and second-order wave crests in case 1. For further details on the case, see van Essen and Seyffert (2025a). We aim to predict the HF one-hour MPM values $\widehat{V''_{hog}}$ and $\widehat{V''_{sag}}$. We did not use an LF threshold value of PAS; no false positives were expected for this case.

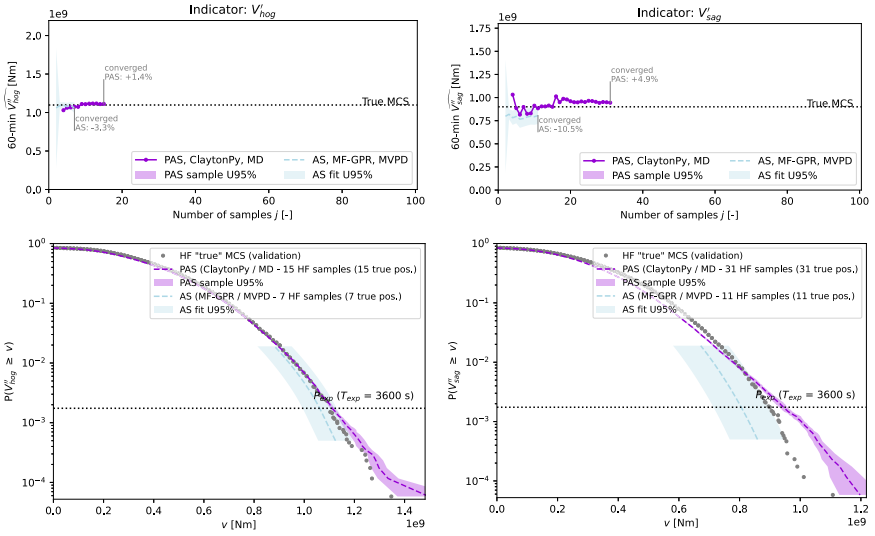
D.3. Cases 1 and 2: convergence and predicted distributions

Similar plots as shown for the complex cases in Section 5, are shown here for the simpler verification cases 1 and 2. Fig. 14 shows these results for cases 1a,b, Fig. 15 and for cases 2a,b. Discussion of these results is included in Section 5.



(a) Case 1a, indicator GoodInd (b) Case 1b, indicator WorseInd

Fig. 14. Case 1 - waves: convergence of one-hour MPM as a function of number of samples (top) and final converged distributions (bottom) from AS and PAS. The copula in the name of the PAS results is the used model in the last iteration.



(a) Case 2a, hogging mode (b) Case 2b, sagging mode

Fig. 15. Case 2 - VBM: convergence of one-hour MPM as a function of number of samples (top) and final converged distributions (bottom) from AS and PAS. The copula in the names of the PAS results is the used model in the last iteration.

Appendix E. Extra case information: implementations, specifications and experimental details

E.1. Scatter plots

Scatter plots of LF indicator peaks versus matched HF response peaks from the validation data for all cases are shown in Fig. 16, illustrating the screening assumption that LF and HF order statistics are comparable. Each plot includes Kendall's τ and tail dependence coefficients λ_U and λ_L (see Appendix B.2), quantifying the rank correlation between LF and HF peaks and indicating the suitability of a copula for modelling their joint behaviour. Perfectly matched order statistics would yield a monotonically increasing line with $\tau = 1$, but the plots show varying correspondence across cases. Cases 3 and 4 additionally reflect the 100 kN per-panel threshold applied in the original experimental data to remove measurement noise and low-frequency inertial effects.

E.2. Experimental details and scale effects

For cases 3, 4 and 5, the HF validation datasets were generated using experiments at a scale of 1:35.986 in the Seakeeping and Manoeuvring Basin of MARIN in Wageningen, The Netherlands. More details about these experiments can be found in van Essen et al. (2023, 2024). We used standard Froude scaling to obtain HF full-scale RWEs, green water and slamming forces. Seakeeping responses are predominantly governed by inertia, so scale effects on most global (relative) motions and first-order loads are expected to be small. Scale effects may occur in the impact loads when air inclusions occur during wave impact events (see e.g. Scharnke and Helder, 2023; Ezeta and Düz, 2025). However,

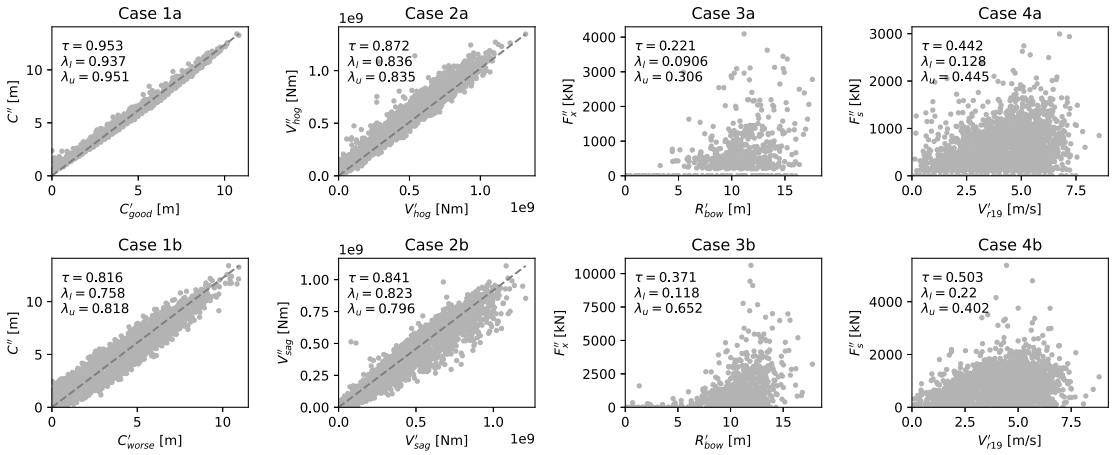


Fig. 16. Scatter plots of LF indicator peaks and matched HF validation peaks for all cases, including the associated Kendall's τ and upper/lower tail dependence coefficients λ_U / λ_L (defined in Appendix B.1).

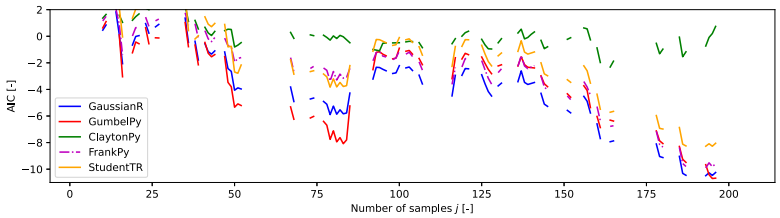


Fig. 17. AIC of the five copula candidate models as a function of the number of HF samples, for green water case 3a (which is converged at 47 HF samples).

no significant air inclusions were observed in the green water impacts of the present dataset. In most cases a horizontal 'jet' was formed over the foredeck, which subsequently hit the accommodation; this does not easily produce air pockets (see e.g., Buchner, 2002). This is supported by the full-scale CFD calculations with single-phase code ComFLOW in Bandringa et al. (2020), which were able to replicate a few of the Froude-scaled largest green water impact force events very accurately for this case. Small air inclusions may locally violate the assumptions underlying Froude scaling for the slamming impacts, but no reliable alternative scaling approach is available. Visual observations during the experiments indicated that the occurrence of such inclusions seems limited.

E.3. Simulation details

For case 1, the HF validation dataset of second-order wave elevation time traces was analytically generated with the `Python PySeaWave` package, using the random phase method and a frequency bandwidth of 0–5 rad/s for second-order interactions. The linear wave LF indicator time traces for case 1a were also derived from this package; the additional noise time traces for case 1b were added to this. For case 2, the HF validation dataset of non-linear VBM time traces was generated using the `PRETTI_R` program (v19.0.1), which is a non-linear Froude–Krylov solver using linear RAO inputs (in this case from linear frequency-domain potential flow diffraction method `SEACAL`). The linear VBM LF indicator time traces were generated with `SEACAL` (v7.2.0) in the zero speed Green's function mode. The ferry hull was described with 3327 panels (a figure of the mesh is included in van Essen and Seyffert, 2025a). For cases 3, 4 and 5, the linear LF vertical bow motions and RWE indicator time traces were generated with `SEACAL` in the Rankine-source mode. The relative rise velocity per wave event used for case 5 was derived from the RWE time traces with a custom `Python` script. The Generalised Pareto fit to the ground truth HF validation data for cases 3, 4 and 5 was made using the `Python scipy.stats.genpareto` package.

Appendix F. Convergence of copula selection process

Firstly, we evaluate how the iterative copula selection procedure converges for all considered cases. The AIC-based copula selection procedure, described in Section 2.2, aims to identify the candidate copula that best fits the currently available HF samples, rather than to determine which copula performs best on the full validation dataset. Convergence to a single copula model for each case is therefore not required; new insights from the Bayesian adaptive sampling procedure, based on new HF samples, may change which model is most likely. This is visualised in Fig. 17, which shows the AIC for case 3a as it evolves over successive iterations. A lower AIC is better. Fig. 18 provides an overview of the selected copula for each case and each iteration (up to a number of iterations past convergence). These figures show that although the Gumbel copula has a slightly lower AIC value when ~5-80 HF samples are available for case 3a, PAS retains the Gaussian copula identified in the first iteration as the most likely model. This is due to the utilised AIC difference threshold of 4.0 (see Section 2.2); only when the AIC difference between models exceeds this threshold is an alternative model considered significantly better. Fig. 17 also shows that the Gaussian and Gumbel models are both good models for the full validation dataset. Fig. 6(a) showed that this case was converged and quite accurate at 47 HF samples already. It is therefore not necessary to arrive at a 'true copula model' before predicting the distributions; the most likely model with fewer HF samples already produces a sufficient result. Fig. 18 shows a consistent preference for the Gaussian copula across cases, likely because it fits very sparse datasets well at low iteration numbers. In later iterations, it is replaced only when a significantly better model is identified (ΔAIC > 4.0). While this may introduce a slight bias in the convergence process (especially as the Gaussian model does not fit all combinations of τ, λ_U and λ_L in Fig. 16 equally well), the results remain acceptable and the procedure is stabilised by the flexibility of the Gaussian copula. When an alternative model is selected, it is most often the Clayton copula, consistent with the generally stronger lower-tail than upper-tail dependence observed,

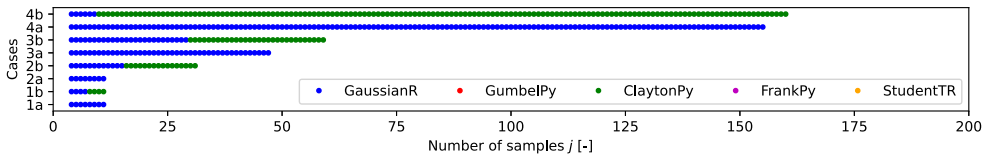
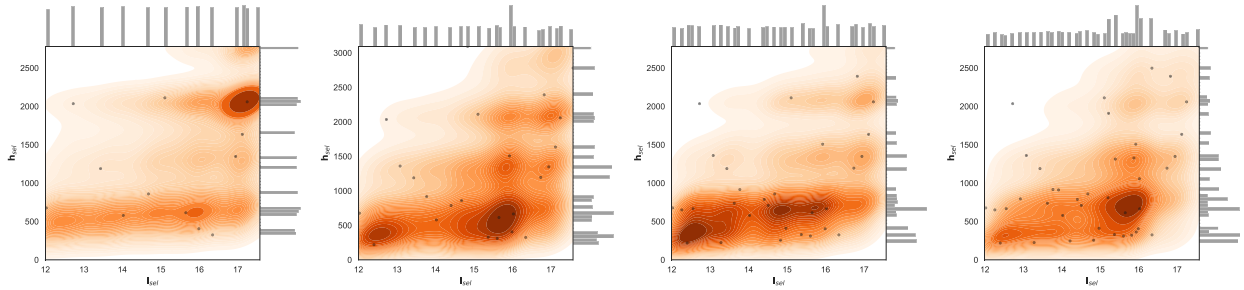
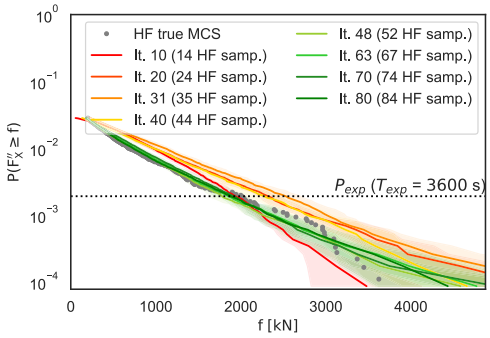


Fig. 18. Selected copula model per case and iteration, plotted until convergence in each case.



(a) It. 20 (24 HF samples). (b) It. 40 (44 HF samples). (c) It. 63 (67 HF samples). (d) It. 80 (84 HF samples).



(e) Distribution convergence.

Fig. 19. Convergence of samples and fitted copula models in real space, including histograms of the associated empirical marginals on the sides, at approx. every 20th iteration (top), and the convergence of the predicted distributions at approx. every 10th iteration (bottom), both for green water case 3a from PAS. This case was considered converged with 47 HF samples (Fig. 6(a)). The nearest iteration is plotted whenever a false positive sample was found in one of the targeted iterations (it. = iteration).

particularly for cases 3 and 4 (see Fig. 16). Finally, Fig. 19 shows the evolution of HF samples, fitted copula model and predicted distributions with an increasing number of iterations for case 3a. This shows that the copula model and distributions move somewhat over the first ~40 iterations, but then stabilise to the converged values (as correctly detected by the convergence criterion at 47 samples).

Data availability

All scripts underlying this publication, the full datasets of case 1 and 2, and the relevant parts of the experimental dataset for cases 3 and 4 are available in the 4TU repository: van Essen and Seyffert (2025d).

References

Abaei, M.M., Leira, B.J., Saevik, S., BahooToroody, A., 2024. Integrating physics-based simulations with gaussian processes for enhanced safety assessment of offshore installations. Reliab. Eng. Syst. Saf. 249, 110235. http://dx.doi.org/10.1016/j.ress.2024.110235.

ABS, 2020. Guidance notes on air gap and wave impact analysis for semisubmersibles. American Bureau of Shipping, Spring, USA, https://pub-rm20.apps.eagle.org/viewer/book-attachment/hnkt3llKpWEbOWbbm3nqnQ/K8aK9uhB5QaI4RoEjIOnoQ-hnkt3llKpWEbOWbbm3nqnQ.

ABS, 2021. Guide for slamming loads and strength assessment for vessels. American Bureau of Shipping, Spring, USA, https://pub-rm20.apps.eagle.org/viewer/book-attachment/GwEbw4kR3F46d2hwrKQcLg/fzVGzQ4EAJe2QQ1F~g4QOw-GwEbw4kR3F46d2hwrKQcLg.

Adegeest, L., Braathen, A., Løseth, R., 1998. Use of nonlinear sea loads simulations in design of ships. In: 12th PRADS Conf.. The Hague, The Netherlands, pp. 53–58.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 19 (6), 716–723.

Antonini, A., Brownjohn, J.M.W., Dassanayake, D., Raby, A., Bassit, J., Pappas, A., D'Ayala, D., 2021. A Bayesian inverse dynamic approach for impulsive wave loading reconstruction. Coast. Eng. 168, 103920. http://dx.doi.org/10.1016/j.coastaleng.2021.103920.

Bandringa, H., Helder, J.A., 2018. On the validity and sensitivity of CFD simulations for a deterministic breaking wave impact on a semi subsmersible. In: 37th OMAE Conf.. ASME, Madrid, Spain, http://dx.doi.org/10.1115/OMAE2018-78089.

Bandringa, H., Helder, J.A., van Essen, S.M., 2020. On the validity of CFD for simulating extreme green water loads on ocean-going vessels. In: 39th OMAE Conf.. ASME, Virtual, Online, http://dx.doi.org/10.1115/OMAE2020-18290.

Bao, Y., Sun, H., Guan, X., Tian., Y., 2024. An active learning method using deep adversarial autoencoder-based sufficient dimension reduction neural network for high-dimensional reliability analysis. Reliab. Eng. Syst. Saf. 247, 110140. http://dx.doi.org/10.1016/j.ress.2024.110140.

Beck, N., Genest, C., Jalbert, J., Mailhot, M., 2020. Predicting extreme surges from sparse data using a copula-based hierarchical Bayesian spatial model. Environmetrics 31, e2616. http://dx.doi.org/10.1002/env.2616.

Bracken, C., Holman, K., Rajagopalan, B., Moradkhani, H., 2018. A Bayesian hierarchical approach to multivariate nonstationary hydrologic frequency analysis. Water Resour. Res. 54, 243–255. http://dx.doi.org/10.1002/2017WR020403.

Bru-Cordero, O.E., Castro, C., Leiva, V., Jaramillo-Elorza, M.C., 2025. Confidence intervals for the reliability of dependent systems: integrating frailty models and copula-based methods. Comput. Model. Eng. Sci. 143 (2), 1401–1431. http://dx.doi.org/10.32604/cmes.2025.064487.

Buchner, B., 2002. Green water on ship-type offshore structures (Ph.D. thesis). Delft University of Technology, Delft, The Netherlands.

Bunnik, T.H.J., Helder, J.A., de Ridder, E.-J., 2015. Deterministic simulation of breaking wave impact and flexible response of a fixed offshore wind turbine. In: 34th OMAE Conf.. ASME, St. John's, Newfoundland, Canada, http://dx.doi.org/10.1115/OMAE2015-41989.

Bunnik, T.H.J., Scharnke, J., de Ridder, E.-J., 2019. Efficient indicators for screening of random waves for wave impacts on a jacket platform and a fixed offshore wind turbine. In: 38th OMAE Conf.. ASME, Glasgow, UK, http://dx.doi.org/10.1115/OMAE2019-95481.

Bunnik, T.H.J., Stansberg, C.T., Pákozdi, C., Fouques, S., Somers, L., 2018. Useful indicators for screening of sea states for wave impacts on fixed and floating platforms. In: 37th OMAE Conf.. ASME, Madrid, Spain, http://dx.doi.org/10.1115/OMAE2018-78544.

Burnham, K.P., Anderson, D.R., 2002. Model selection and multimodel inference: a practical information-theoretic approach, second ed. Springer-Verlag, New York, Chapter 2.7.

BV, 2015. Rule note NR583: Whipping and springing assessment. Bureau Veritas, Paris, France, https://erules.veristar.com/dy/data/bv/pdf/583-NR_2015-07.pdf.

BV, 2019. Rule Note NI638: Guidance for long-term hydro-structure calculations. Bureau Veritas, Paris, France, https://erules.veristar.com/dy/data/bv/pdf/638-NI_2019-02.pdf.

Chan, J., Papaioannou, I., Straub, D., 2022. An adaptive subset simulation algorithm for system reliability analysis with discontinuous limit states. Reliab. Eng. Syst. Saf. 225, 108607. http://dx.doi.org/10.1016/j.ress.2022.108607.

Cheng, K., Papaioannou, I., Lu, Z., Zhang, X., Wang, Y., 2023. Rare event estimation with sequential directional importance sampling. Struct. Saf. 100, 102291. http://dx.doi.org/10.1016/j.strusafe.2022.102291.

Cheng, K., Papaioannou, I., Straub, D., 2025. Enhanced sequential directional importance sampling for structural reliability analysis. Struct. Saf. 114, 102574. http://dx.doi.org/10.1016/j.strusafe.2025.102574.

Chiron, M., Genest, C., Morio, J., Dubreuil, S., 2023. Failure probability estimation through high-dimensional elliptical distribution modeling with multiple importance sampling. Reliab. Eng. Syst. Saf. 235, 109238. http://dx.doi.org/10.1016/j.ress.2023.109238.

Chojaczyk, A., Teixeira, A., Neves, L., Cardoso, J., Soares, C.G., 2015. Review and application of Artificial Neural Networks models in reliability analysis of steel structures. Struct. Saf. 52, 78–89. http://dx.doi.org/10.1016/j.strusafe.2014.09.002.

Cuomo, G., Lupoi, G., Shimosako, K., Takahashi, S., 2011a. Dynamic response and sliding distance of composite breakwaters under breaking and non-breaking wave attack. Coast. Eng. 58, 953–969. http://dx.doi.org/10.1016/j.coastaleng.2011.03.008.

Cuomo, G., Piscopia, R., Allsop, W., 2011b. Evaluation of wave impact loads on caisson breakwaters based on joint probability of impact maxima and rise times. Coast. Eng. 58 (1), 9–27. http://dx.doi.org/10.1016/j.coastaleng.2010.08.003.

Cuomo, G., ichiro Shimosako, K., Takahashi, S., 2009. Wave-in-deck loads on coastal bridges and the role of air. Coast. Eng. 56, 793–809. http://dx.doi.org/10.1016/j.coastaleng.2009.01.005.

Dallinga, R.P., Gaillarde, G., 2001. Hatch cover loads experienced by M.V. Derbyshire during typhoon 'Orchid'. In: Glasgow Marine Fair and Int. Workshop on Safety of Bulk Carriers. Glasgow, Scotland, UK, Glasgow, Scotland, UK.

Davidson, A.C., Smith, R.L., 1990. Models for exceedances over high thresholds. J. R. Stat. Soc. Ser. B (Methodological) 52 (3), 393–442, https://www.jstor.org/stable/2345667.

Davison, A.C., Padoan, S.A., Ribatet, M., 2012. Statistical modeling of spatial extremes. Statist. Sci. 27 (2), 161–186. http://dx.doi.org/10.1214/11-STS376.

van de Bunt, E., Dekker, J., Scharnke, J., Jaouën, F., 2021. Applying force panels for wave impact measurements. Ocean Eng. 232, http://dx.doi.org/10.1016/j.oceaneng.2021.108857.

de Hauteclocque, G., Mackay, E., Vanem, E., 2022. Quantitative comparison of environmental contour approaches. Ocean Eng. 245, 110374. http://dx.doi.org/10.1016/j.oceaneng.2021.110374.

de Waal, D., van Gelder, P., 2005. Modelling of extreme wave heights and periods through copulas. Extremes 8, 345–356. http://dx.doi.org/10.1007/s10687-006-0006-y.

Dermatis, A., Bouscasse, B., Ducrozet, G., Bredmose, H., Bingham, H.B., 2025a. Prediction of the extreme slow-drift response of moored floating structures using design waves. Ocean Eng. 333, 121456. http://dx.doi.org/10.1016/j.oceaneng.2025.121456.

Dermatis, A., Lasbleis, M., Kim, S., Hauteclocque, G.D., Bouscasse, B., Ducrozet, G., 2025b. A multi-fidelity approach for the evaluation of extreme wave loads using nonlinear response-conditioned waves. Ocean Eng. 316, 119919. http://dx.doi.org/10.1016/j.oceaneng.2024.119919.

DNV, 2018. Class Guideline DNV-CG-0130: Wave loads. Det Norske Veritas, Oslo, Norway, https://rules.dnv.com/docs/pdf/DNV/CG/2018-01/DNV-CG-0130.pdf.

DNV, 2019a. Recommended Practice RP-C205: Environmental conditions and environmental loads. Det Norske Veritas, Oslo, Norway, https://www.dnv.com/energy/standards-guidelines/dnv-rp-c205-environmental-conditions-and-environmental-loads/.

DNV, 2019b. Offshore Technical Guidance DNV-OTG-14: Horizontal wave impact loads for column stabilised units. Det Norske Veritas, Oslo, Norway, https://www.dnv.com/maritime/Offshore/technical-guidance-otg.html.

Duan, X., Wang, S., Di Liu, Shi, J., Wu, Y., Zhou, X., 2024. A statistical analysis method for significant wave height and spectral peak frequency considering the random and time-varying effects based on copula function and Bayesian inference. Ocean. Model. 190, 102390. http://dx.doi.org/10.1016/j.ocemod.2024.102390.

Düz, B., Scharnke, J., Hallmann, R., Tukker, J., Khurana, S., Blanchard, K., 2020. Comparison of the CFD results to PIV measurements in kinematics of spilling and plunging breakers. In: 39th OMAE Conf.. ASME, Virtual, Online, http://dx.doi.org/10.1115/OMAE2020-19268.

Ersdal, G., Kvitrud, A., 2000. Green water on Norwegian production ships. In: 10th ISOPE Conf.. Int. Soc. of Offshore and Polar Eng., Seattle, USA.

Ezeta, R., Düz, B., 2025. Predicting the dynamics of a gas pocket during breaking wave impacts using machine learning. Ocean Eng. 321, 120321. http://dx.doi.org/10.1016/j.oceaneng.2025.120321.

Fang, C., Xu, Y.-L., Li, Y., Li, J., 2024. Serviceability analysis of sea-crossing bridges under correlated wind and wave loads. Reliab. Eng. Syst. Saf. 246, 110077. http://dx.doi.org/10.1016/j.ress.2024.110077.

Fernandez Castellon, D., Fenerci, A., Petersen, Ø.W., Øiseth, O., 2023. Full long-term buffeting analysis of suspension bridges using Gaussian process surrogate modelling and importance sampling Monte Carlo simulations. Reliab. Eng. Syst. Saf. 235, 109211. http://dx.doi.org/10.1016/j.ress.2023.109211.

Forrester, I.J., Sobester, A., Keane, A.J., 2007. Multi-fidelity optimization via surrogate modelling. Proc. R. Soc. A 463, 3251–3269. http://dx.doi.org/10.1098/rspa.2007.1900.

Fuhg, J.N., Fau, A., Nackenhorst, U., 2021. State–of–the–art and comparative review of adaptive sampling methods for kriging. Arch. Comput. Methods Eng. 28, 2689–2747. http://dx.doi.org/10.1007/s11831-020-09474-6.

Gaidai, O., Storhaug, G., Naess, A., 2016. Extreme large cargo ship panel stresses by bivariate ACER method. Ocean Eng. 123, 432–439. http://dx.doi.org/10.1016/j.oceaneng.2016.06.048.

Genest, C., Favre, A.-C., 2007. Everything you always wanted to know about copula modeling but were afraid to ask. J. Hydrol. Eng. 12 (4), 347–368. http://dx.doi.org/10.1061/(ASCE)1084-0699(2007)12:4(347).

Ghosh, S., Mallick, B.K., 2010. A hierarchical Bayesian spatio-temporal model for extreme precipitation events. Environmetrics 22, 192–204. http://dx.doi.org/10.1002/env.1043.

Gramstad, O., Agrell, C., Bitner-Gregersen, E., Guo, B., Ruth, E., Vanem, E., 2020. Sequential sampling method using Gaussian process regression for estimating extreme structural response. Mar. Struct. 72, http://dx.doi.org/10.1016/j.marstruc.2020.102780.

Gramstad, O., Johannessen, T.B., Lian, G., 2023. Long-term analysis of wave-induced loads using High Order Spectral Method and direct sampling of extreme wave events. Mar. Struc. 103473, http://dx.doi.org/10.1016/j.marstruc.2023.103473.

Güdendorf, G., Segers, J., 2010. Extreme-value copulas. In: Jaworski, P., Durante, F., Härdle, W.K., Rychlik, T. (Eds.), Copula Theory and Its Applications. In: Lecture Notes in Statistics, vol. 198, Springer, pp. 127–145.

Guth, S., Katsidoniotaki, E., Sapsis, T.P., 2023. Statistical modeling of fully nonlinear hydrodynamic loads on offshore wind turbine monopile foundations using wave episodes and targeted CFD simulations through active sampling. Wind. Energy 28, 1–26. http://dx.doi.org/10.1002/we.2880.

Guth, S., Sapsis, T.P., 2022. Wave episode based Gaussian process regression for extreme event statistics in ship dynamics: Between the Scylla of Karhunen–Loeve convergence and the Charybdis of transient features. Ocean Eng. 266, 112633. http://dx.doi.org/10.1016/j.oceaneng.2022.112633.

Halsne, M., N., O., Ersdal, G., Langøy, M., Andersen, T., Bjørheim, L.G., 2022. Semisubmersible in service experiences on the Norwegian Continental Shelf. In: 41th OMAE Conf.. ASME, Hamburg, Germany, http://dx.doi.org/10.1115/OMAE2022-81289.

Han, F., Wang, W., Zheng, X.-W., Han, X., Shi, W., Li, X., 2025. Investigation of essential parameters for the design of offshore wind turbine based on structural reliability. Reliab. Eng. Syst. Saf. 254, 110601. http://dx.doi.org/10.1016/j.ress.2024.110601.

Hansom, J.D., Switzer, A.D., Pile, J., 2015. Extreme waves: causes, characteristics, and impact on coastal environments and society (Ch.11). In: Shroder, J.F., Ellis, J.T., Sherman, D.J. (Eds.), Coastal and Marine Hazards, Risks, and Disasters. In: Hazards and Disasters Series, Elsevier, Boston, pp. 307–334. http://dx.doi.org/10.1016/B978-0-12-396483-0.00011-X.

Haselsteiner, A.F., Coe, R.G., Manuel, L., Chai, W., Leira, B., Clarindo, G., Guedes Soares, C., Hannesdóttir, Á., Dimitrov, N., Sander, A., Ohlendorf, J.-H., Thoben, K.-D., de Hauteclocque, G., Mackay, E., Jonathan, P., Qiao, C., Myers, A., Rode, A., Hildebrandt, A., Schmidt, B., Vanem, E., Huseby, A.B., 2021. A benchmarking exercise for environmental contours. Ocean Eng. 236, 109504. http://dx.doi.org/10.1016/j.oceaneng.2021.109504.

Haver, S., 2017. Metocean modelling and prediction of extremes. (May), Haver & havet, University in Stavanger, NTNU, Stavanger, Norway, pp. 1–256.

Havtil, N.O.I.A., 2025. Investigation of incident on 31 January 2024 involving a cabin window on Åsgard A being forced in by green water. In: Tech. rep.. https://www.havtil.no/contentassets/{6}e6aeac3d63f410a9845c36bae6da2af/{2024}_247-rapport-eng-gransking-equinor-asgard-a-innslatt-lugarvindu_oppdatert-mai_skjult-innhold.pdf.

Hermundstad, O.A., Moan, T., 2007. Efficient calculation of slamming pressures on ships in irregular seas. J. Mar. Sci. Technol. 12 (3), 160–182. http://dx.doi.org/10.1007/s00773-006-0238-1.

Hofert, M., Kojadinovic, I., Maechler, M., Yan, J., 2016. Copula: Multivariate dependence with copulas. URL http://cran.r-project.org/package=copula.

Hofert, M., Kojadinovic, I., Maechler, M., Yan, J., 2018. Elements of copula modeling with R. Springer Use R! Series, URL http://www.springer.com/de/book/9783319896342.

Hoffman, D., Maclean, W.M., 1970. Ship model study of incidence of shipping water forward. In: Mar. Technol..

Hong, S.Y., Lee, P.M., Gong, D.S., 1993. Experimental study on the deck wetting of a container ship in irregular head waves. Sel. Pap. Soc. Nav. Arch. Korea 1, 37–44.

Huang, D., Allen, T.T., Notz, W.I., Miller, R.A., 2006. Sequential kriging optimization using multiple-fidelity evaluations. Struct. Multidiscip. Optim. 32, 369–382. http://dx.doi.org/10.1007/s00158-005-0587-0.

Huseby, A.B., Vanem, E., Natvig, B., 2013. A new approach to environmental contours for ocean engineering applications based on direct Monte Carlo simulations. Ocean Eng. 60, 124–135. http://dx.doi.org/10.1016/j.oceaneng.2012.12.034.

ITTC, 2017a. Recommended Practice 7.5-02-07-02.3: Experiments on rarely occurring events. International Towing Tank Conf., https://www.ittc.info/media/8105/75-02-07-023.pdf.

ITTC, 2017b. Recommended Practice 7.5-02-07-02.6: Global loads seakeeping procedure. International Towing Tank Conference, https://www.ittc.info/media/8109/75-02-07-026.pdf.

Jensen, J.J., 2009. Extreme value predictions and critical wave episodes for marine structures by FORM. Ships Offshore Struct. 3, 4, 325–333. http://dx.doi.org/10.1080/17445300802370461.

Johannessen, T.B., Lande, Ø., 2018. Long term analysis of steep and breaking wave properties by event matching. In: 37th OMAE Conf.. ASME, Madrid, Spain, http://dx.doi.org/10.1115/OMAE2018-78283.

Jonathan, P., Randell, D., Wadsworth, J., Tawn, J., 2021. Uncertainties in return values from extreme value analysis of peaks over threshold using the generalised Pareto distribution. Ocean Eng. 220, 107725. http://dx.doi.org/10.1016/j.oceaneng.2020.107725.

Kapsenberg, G.K., 2018. On the Slamming of Ships (Ph.D. thesis). Delft University of Technology, Delft, The Netherlands, http://dx.doi.org/10.4233/uuid:14eac2bb-63ee-47e4-8218-1ba3830a97b4.

Kennedy, M.C., O'Hagan, A., 2000. Predicting the output from a complex computer code when fast approximations are available. Biometrika 87, 1, 1–13.

Kim, H., Jeong, S., Yang, C., Noblesse, F., 2011. Hull form design exploration based on response surface method. In: 21st ISOPE Conf.. Int. Society of Offshore and Polar Eng., pp. 816–825.

Kim, D.H., Lee, S.G., 2015. Reliability analysis of offshore wind turbine support structures under extreme ocean environmental loads. Renew. Energy 79, 161–166. http://dx.doi.org/10.1016/j.renene.2014.11.052.

Kim, M., O'Conner, K., Pipiras, V., Sapsis, T., 2025. Sampling low-fidelity outputs for estimation of high-fidelity density and its tails. SIAM/ASA J. Uncertain. Quantif. 13, 1, 30–62. http://dx.doi.org/10.1137/24M1639142.

Kingston, G.B., Rajabali Nejad, M., Gouldby, B.P., van Gelder, P.H.A.J.M., 2011. Computational intelligence methods for the efficient reliability analysis of complex flood defence structures. Struct. Saf. 33, 1. http://dx.doi.org/10.1016/j.strusafe.2010.08.002.

Kramer, B., Marques, A.N., Peherstorfer, B., Villa, U., Willcox, K., 2019. Multifidelity probability estimation via fusion of estimators. J. Comput. Phys. 392, 385–402. http://dx.doi.org/10.1016/j.jcp.2019.04.071.

Krupskii, P., Joe, H., 2019. Nonparametric estimation of multivariate tail probabilities and tail dependence coefficients. J. Multivariate Anal. 172, 147–161. http://dx.doi.org/10.1016/j.jmva.2019.02.013, Dependence Models.

Lavaud, S., 2018. Tail dependence in small samples: from theory to practice. J. Oper. Risk 13 (1), 15–49. http://dx.doi.org/10.21314/JOP.2017.196.

Li, Q., Ni, P., Du, X., Han, Q., 2025. Bayesian model updating with variational inference and Gaussian copula model. Comput. Methods Appl. Mech. Engrg. 438, 117842. http://dx.doi.org/10.1016/j.cma.2025.117842.

Li, M., Sadoughi, M., Hu, Z., Hu, C., 2020. A hybrid Gaussian process model for system reliability analysis. Reliab. Eng. Syst. Saf. 197, 106816. http://dx.doi.org/10.1016/j.ress.2020.106816.

Li, P., Wang, Y., 2022. An active learning reliability analysis method using adaptive Bayesian compressive sensing and Monte Carlo simulation (ABCS-MCS). Reliab. Eng. Syst. Saf. 221, 108377. http://dx.doi.org/10.1016/j.ress.2022.108377.

Lu, C., Feng, Y., Fei, C., Teng, D., 2025. Dimensional synchronous modeling-based enhanced Kriging algorithm and adaptive Copula method for multi-objective synthetical reliability analyses. Chin. J. Aeronaut. 38 (9), 103652. http://dx.doi.org/10.1016/j.cja.2025.103652.

Lucio, D., Lara, J., Tomas, A., Losada, I., 2024. Probabilistic assessment of climate-related impacts and risks in ports. Reliab. Eng. Syst. Saf. 251, 110333. http://dx.doi.org/10.1016/j.ress.2024.110333.

Ma, P., Zhang, Y., Cai, E., Luo, M., Guo, J., Guo, T., 2026. A copula-based transitional markov chain monte carlo method for bayesian model updating. Reliab. Eng. Syst. Saf. 265, 111572. http://dx.doi.org/10.1016/j.ress.2025.111572.

Mackay, E.B.L., Murphy-Barltrop, C., Jonathan, P., 2025. The SPAR model: a new paradigm for multivariate extremes: application to joint distributions of metocean variables (OMAE-24-1018). J. Offshore Mech. Arct. Eng. 147, 011205:1–10. http://dx.doi.org/10.1115/1.4065968.

Mares-Nasarre, P., van Gent, M.R., Morales-Nápoles, O., 2024. A copula-based model to describe the uncertainty of overtopping variables on mound breakwaters. Coast. Eng. 189, 104483. http://dx.doi.org/10.1016/j.coastaleng.2024.104483.

Marrel, A., Iooss, B., 2024a. Probabilistic surrogate modeling by Gaussian process: A new estimation algorithm for more robust prediction. Reliab. Eng. Syst. Saf. 247, 110120. http://dx.doi.org/10.1016/j.ress.2024.110120.

Marrel, A., Iooss, B., 2024b. Probabilistic surrogate modeling by Gaussian process: A review on recent insights in estimation and validation. Reliab. Eng. Syst. Saf. 247, 110094. http://dx.doi.org/10.1016/j.ress.2024.110094.

McCann, M., Ebrahimi, B., Cinar, G., Renteria, W., Stehno, A., Lynett, P., Kaihatu, J., 2024. Field observations of Hurricane Ian's wave and surge impact in the areas of Fort Myers Beach and Sanibel Island, USA. Coast. Eng. 188, 104450. http://dx.doi.org/10.1016/j.coastaleng.2023.104450.

Mohamad, M.A., Sapsis, T.P., 2018. Sequential sampling strategy for extreme event statistics in nonlinear dynamical systems. Proc. Natl. Acad. Sci. 115, 44, http://dx.doi.org/10.1073/pnas.1813263115.

Nelsen, R.B., 2006. An introduction to copulas, second ed. Springer Series in Statistics, Springer-Verlag, New York, NY, USA, http://dx.doi.org/10.1007/0-387-28678-0.

Nguyen, P.T.T., Manuel, L., 2024a. A bi-fidelity surrogate model for extreme loads on offshore structures. Ocean Eng. 307, 118175. http://dx.doi.org/10.1016/j.oceaneng.2024.118175.

Nguyen, P.T.T., Manuel, L., 2024b. Uncertainty quantification in low-probability response estimation using sliced inverse regression and polynomial chaos expansion. Reliab. Eng. Syst. Saf. 242, 109750. http://dx.doi.org/10.1016/j.ress.2023.109750.

Ochi, M.K., 1964. Prediction of occurrence and severity of ship slamming at sea ACR112. In: 5th Symp. Nav. Hydrodyn.. Office of Naval Research, Bergen, Norway.

Ochi, M.K., 1990. Applied probability and stochastic processes in engineering and physical sciences. John Wiley & Sons, Singapore.

Ogawa, Y., Minumi, M., Tunizuwu, K., Kumuno, A., Mutsunami, R., Huyushi, T., 2002. Shipping water load due to deck wetness. In: 12th ISOPE Conf.. Int. Soc. of Offshore and Polar Eng., Kitayushu, Japan.

Pákozdi, C., Califano, A., Akselsen, A., Croonenborghs, E., Kim, J., Peric, M., Loubeyre, S., Bouscasse, B., Ducrozet, G., Xu-Haihua, 2022. Joint-industry effort to develop and verify CFD modeling practice for predicting wave impact. In: 41st OMAE Conf.. ASME, Hamburg, Germany, http://dx.doi.org/10.1115/OMAE2022-79152.

Peherstorfer, B., Cui, T., Marzouk, Y., Willcox, K., 2016. Multifidelity importance sampling. Comput. Methods Appl. Mech. Engrg. 300, 490–509. http://dx.doi.org/10.1016/j.cma.2015.12.002.

Pereira, I., 2023. Coast Guard probing deaths, injuries of Americans on vessels in Antarctic waters. ABC News https://abcnews.go.com/International/coast-guard-probing-deaths-injuries-americans-vessels-antarctic/story?id=96864073.

Proppe, C., 2019. Reliability estimation with multi-fidelity simulation methods. In: Beer, M., Zio, E. (Eds.), 29th ESREL Conf.. Research Publishing, Singapore, http://dx.doi.org/10.3850/978-981-11-2724-30168-cd.

Qiao, X., Xie, Q., Shi, G., Hu, J., 2025. Seismic reliability assessment for valve hall in converter station based on Copula theory. Structures 77, 109002. http://dx.doi.org/10.1016/j.istruc.2025.109002.

Reuters, 2023. Norwegian cruise ship MS maud loses power in north sea during storm. Guardian https://www.theguardian.com/world/{2023}/dec/{22}/norwegian-cruise-ship-ms-maud-loses-power-navigation-system.

Roy, A., Manna, R., Chakraborty, S., 2018. Support vector regression based metamodeling for structural reliability analysis. Probabilistic Eng. Mech. 55, 78–89. http://dx.doi.org/10.1016/j.probengmech.2018.11.001.

Santjer, R., Agerwal, S., Colomés, O., Morales-Nápoles, O., 2025. Gaussian copula-based Bayesian networks for dynamic loads in mooring systems. Appl. Ocean Res. 165, 104809. http://dx.doi.org/10.1016/j.apor.2025.104809.

Saraygord Afshari, S., Enayatollahi, F., Xu, X., Liang, X., 2022. Machine learning-based methods in structural reliability analysis: A review. Reliab. Eng. Syst. Saf. 108223. http://dx.doi.org/10.1016/j.ress.2021.108223.

Scharnke, J., Helder, J.A., 2023. Scale effects and variability in wave-in-deck type of impact loading, more insights into the results of the Breakin JIP. In: 42nd OMAE Conf.. ASME, Melbourne, Australia, http://dx.doi.org/10.1115/OMAE2023-104288.

Scharnke, J., van Essen, S.M., Seyffert, H.C., 2023. Required test durations for converged short-term wave and impact extreme value statistics - Part 2: deck box dataset. Mar. Struc. 90, 103411. http://dx.doi.org/10.1016/j.marstruc.2023.103411.

Serinaldi, F., Cuomo, G., 2011. Characterizing impulsive wave-in-deck loads on coastal bridges by probabilistic models of impact maxima and rise times. Coast. Eng. 58 (9), 908–926. http://dx.doi.org/10.1016/j.coastaleng.2011.05.010.

Sharma, J.N., Dean, R.G., 1979. Development and evaluation of a procedure for simulating a random directional second order sea surface and associated wave forces. Ocean. Eng. Rep. 20.

Shen, J., Liu, H., 2024. On the structure dynamic response induced by the dam-break surge impact using multivariate copulas. Ocean Eng. 306, 118100. http://dx.doi.org/10.1016/j.oceaneng.2024.118100.

Sklar, A., 1959. Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de statistique de l'Université de Paris 8 (3), 229–231.

Speers, M., Randell, D., Tawn, J., Jonathan, P., 2024. Estimating metocean environments associated with extreme structural response to demonstrate the dangers of environmental contour methods. Ocean Eng. 311, 118754. http://dx.doi.org/10.1016/j.oceaneng.2024.118754.

Stansberg, C.T., 2008. A wave impact parameter. In: 27th OMAE Conf.. ASME, Estoril, Portugal, http://dx.doi.org/10.1115/OMAE2008-57801.

Stansberg, C.T., 2020. Wave front steepness and influence on horizontal deck impact loads. J. Mar. Sci. Eng. 8, 314, http://dx.doi.org/10.3390/jmse8050314.

Tabandeh, A., Jia, G., Gardoni, P., 2022. A review and assessment of importance sampling methods for reliability analysis. Struct. Saf. 97, 102216. http://dx.doi.org/10.1016/j.strusafe.2022.102216.

Takami, T., Fujimoto, W., Houtani, H., Matsui, S., 2023. Extreme wave and vertical bending moment predictions by higher-order spectrum method and FORM. In: 42nd OMAE Conf.. ASME, Melbourne, Australia, http://dx.doi.org/10.1115/OMAE2023-101876.

Tang, T., Ding, H., Dai, S., Chen, X., Taylor, P.H., Zang, J., Adcock, T.A.A., 2023. Data informed model test design with machine learning - an example in nonlinear wave load on a vertical cylinder. In: 42nd OMAE Conf.. ASME, Melbourne, Australia, http://dx.doi.org/10.1115/OMAE2023-102682.

Taylor, P.H., Jonathan, P., Harland, L.A., 1997. Time domain simulation of jack-up dynamics with the extremes of a Gaussian process. J. Vib. Acoust. 119, 624–628. http://dx.doi.org/10.1115/1.2889772.

Torhaug, R., Winterstein, S.R., Braathen, A., 1998. Nonlinear ship loads: stochastic models for extreme response. J. Ship Res. 42, 1, 46–55. http://dx.doi.org/10.5957/jsr.1998.42.1.46.

van Essen, S.M., 2021. Influence of wave variability on ship response during deterministically repeated seakeeping tests at forward speed. In: Okada, T., Suzuki, K., Kawamura, Y. (Eds.), 14th PRADS Conf. 2019. Lecture Notes in Civil Engineering, Vol 63. Springer, Singapore, Yokohama, Japan, http://dx.doi.org/10.1007/978-981-15-4624-2_54.

van Essen, S.M., Bunnik, T.H.J., Scharnke, J., 2024. Statistical uncertainty of ship response to waves as a function of test duration. In: 43rd OMAE Conf.. ASME, Singapore, http://dx.doi.org/10.1115/OMAE2024-122486.

van Essen, S.M., Monroy, C., Shen, Z., Helder, J.A., Kim, D.-H., Seng, S., Ge, Z., 2021. Screening wave conditions for the occurrence of green water events on sailing ships. Ocean Eng. 234, 109218. http://dx.doi.org/10.1016/j.oceaneng.2021.109218.

van Essen, S.M., Scharnke, J., Seyffert, H.C., 2023. Required test durations for converged short-term wave and impact extreme value statistics - Part 1: ferry dataset. Mar. Struc. 90, 103410. http://dx.doi.org/10.1016/j.marstruc.2023.103410.

van Essen, S.M., Seyffert, H.C., 2023. Finding dangerous waves – Review of methods to obtain wave impact design loads for marine structures (OMAE-22-1110). J. Offshore Mech. Arct. Eng. 145 (6), 060801. http://dx.doi.org/10.1115/1.4056888.

van Essen, S.M., Seyffert, H.C., 2025a. Designing for dangerous waves – a new 'Adaptive Screening' method to predict extreme values of non-linear marine and coastal structure responses to waves. Reliab. Eng. Syst. Saf. 264B, 111404. http://dx.doi.org/10.1016/j.ress.2025.111404.

van Essen, S.M., Seyffert, H.C., 2025b. Designing ships for extreme non-linear responses - the role of the acquisition function in the Adaptive Screening extreme value prediction method. In: 16th PRADS Conf.. Ann Arbor, Michigan, USA, http://dx.doi.org/10.5281/zenodo.17305196.

van Essen, S.M., Seyffert, H.C., 2025c. Scripts and data for PRADS publication that varies the acquisition function of the Adaptive Screening method. In: 4TU Repository. http://dx.doi.org/10.4121/12777259-c2f6-4b44-b71f-eec5557824d1.

van Essen, S.M., Seyffert, H.C., 2025d. Scripts and data underlying the paper that introduces and validates the Probabilistic Adaptive Screening method. In: 4TU Repository. http://dx.doi.org/10.4121/be3e7819-dabf-4a21-bb29-5b76179ff696.

van Essen, S.M., Seyffert, H.C., 2025e. Scripts and data underlying the publication that defines and applies the new Adaptive Screening method, for extreme value prediction of non-linear wave-induced responses. In: 4TU Repository.. http://dx.doi.org/10.4121/f1348609-c912-4d06-82b8-197c01f3437b.

van 't Veer, A.P., Vlasveld, E., 2014. Green water phenomena on a twin-hull FLNG concept. In: 33rd OMAE Conf.. ASME, San Francisco, USA, http://dx.doi.org/10.1115/OMAE2014-23915.

Vanem, E., Fekhari, E., Dimitrov, N., Kelly, M., Cousin, A., Guiton, M., 2024. A joint probability distribution for multivariate wind-wave conditions and discussions on uncertainties (OMAE-23-1131). J. Offshore Mech. Arct. Eng. 146 (6), 061701. http://dx.doi.org/10.1115/1.4064498.

Viste-Ollestad, I., Andersen, T.L., Oma, N., Zachariassen, S., 2016. Investigation report petroleumstilsynet - Investigation of an accident with fatal consequences on COSLInnovator, 30 December 2015. In: Tech. rep.. https://www.ptil.no/contentassets/{34}ba6b722c0c44a3a137240bae06f623/investigation-report—cosl-drilling—cosl-innovator.pdf.

Wang, H., Gramstad, O., Schär, S., Marelli, S., Vanem, E., 2024. Comparison of probabilistic structural reliability methods for ultimate limit state assessment of wind turbines. Struct. Saf. 111, 102502. http://dx.doi.org/10.1016/j.strusafe.2024.102502.

Wang, T., Wang, C., Zhou, Y., Cui, H., Ji, J., 2025. Reliability updating of copula-dependent spatially variable soil slopes based on data-augmented MGPR model along slip surfaces. Eng. Geol. 356, 108280. http://dx.doi.org/10.1016/j.enggeo.2025.108280.

Wei, K., Zhou, C., Xu, B., 2022. Spatial distribution models of horizontal and vertical wave impact pressure on the elevated box structure. Appl. Ocean Res. 125, 103245. http://dx.doi.org/10.1016/j.apor.2022.103245.

Wilkie, D., Galasso, C., 2021. Gaussian process regression for fatigue reliability analysis of offshore wind turbines. Struct. Saf. 88, 102020. http://dx.doi.org/10.1016/j.strusafe.2020.102020.

Winterstein, S.R., Ude, T.C., Cornell, C.A., Bjerager, P., Haver, S., 1993. Environmental parameters for extreme response: inverse FORM with omission factors. In: Int. Conf. Struct. Saf. Reliab.. Innsbruck, Austria.

Xu, Y.-B., Liu, L.-L., 2026. Copula-based conditional reliability analysis of slopes in spatially variable soils. Reliab. Eng. Syst. Saf. 265, 111522. http://dx.doi.org/10.1016/j.ress.2025.111522.

Yang, Z., Xu, Y., Shi, L., Zhu, C., Bao, Y., 2025. Investigation of methods for the localization and reconstruction of the wave impact on a floating wind turbine pontoon. J. Ocean. Eng. Sci. http://dx.doi.org/10.1016/j.joes.2025.08.006.

Yuan, X., Shu, Y., Qian, Y., Dong, Y., 2024. Adaptive importance sampling approach for structural time-variant reliability analysis. Struct. Saf. 111, 102500. http://dx.doi.org/10.1016/j.strusafe.2024.102500.

Zhang, Y., Dong, Y., Frangopol, D.M., 2024. An error-based stopping criterion for spherical decomposition-based adaptive Kriging model and rare event estimation. Reliab. Eng. Syst. Saf. 241, 109610. http://dx.doi.org/10.1016/j.ress.2023.109610.

Zhang, N., Xiao, L., Zou, Q., Cummins, C., 2025. Large-scale wave basin experimental study on the spatio-temporal distribution of wave impact loads on a semi-submersible platform. Ocean Eng. 327, 120991. http://dx.doi.org/10.1016/j.oceaneng.2025.120991.

Zhang, D., Zhang, J., Yang, M., Bai, S., Wang, F., 2026. Reliability-based design optimization method with correlated variables using adaptive conjugate gradient analysis and Copula. Reliab. Eng. Syst. Saf. 265, 111469. http://dx.doi.org/10.1016/j.ress.2025.111469.

Zhao, W., Fan, F., Wang, W., 2017. Non-linear partial least squares response surface method for structural reliability analysis. Reliab. Eng. Syst. Saf. 161, 69–77. http://dx.doi.org/10.1016/j.ress.2017.01.004.

Zhou, T., Yin, Y., Ma, Z., Chen, J., Zhai, G., 2023. Numerical investigation of breaking waves impact on vertical breakwater with impermeable and porous foundation. Ocean Eng. 280, 114477. http://dx.doi.org/10.1016/j.oceaneng.2023.114477.