Multi-fidelity design framework to support early-stage design exploration of the AXE frigates

The vertical bending moment case

Charisi, Nikoleta Dimitra; Defer, Emile; Hopman, Hans; Kana, Austin A.

DOI
10.1016/j.oceaneng.2025.123269

Publication date
2026

Document Version
Final published version

Published in
Ocean Engineering

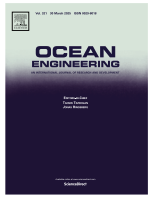
Citation (APA)
Charisi, N. D., Defer, E., Hopman, H., & Kana, A. A. (2026). Multi-fidelity design framework to support early-stage design exploration of the AXE frigates: The vertical bending moment case. *Ocean Engineering*, *343*, Article 123269. https://doi.org/10.1016/j.oceaneng.2025.123269

Ocean Engineering 343 (2026) 123269

Contents lists available at ScienceDirect

Ocean Engineering

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



Research paper

Multi-fidelity design framework to support early-stage design exploration of the AXE frigates: The vertical bending moment case

Nikoleta Dimitra Charisi [a,*], Emile Defer [b], Hans Hopman [a], Austin A. Kana [a]

[a] Maritime and Transport Technology, Delft University of Technology, Delft, The Netherlands
[b] Damen Naval, Vlissingen, The Netherlands

ARTICLE INFO

Keywords:
Early-stage design framework
Novel vessels
Wave-induced loads
Vertical bending moment
Multi-fidelity models
Gaussian processes
Compositional kernels

ABSTRACT

Early-stage design assessment of loads such as vertical bending moments can be a critical quantity of interest for design exploration. Traditionally, classification societies' rules are used to calculate such loads. However, relying solely on these rules for designing new vessels may be insufficient, and conducting direct analyses of a large number of designs to support design exploration is computationally infeasible. Currently, key factors such as wave-induced loads are typically evaluated only in later design stages, where a limited number of promising designs are under consideration. This research explores the potential of harnessing multi-fidelity models for early-stage predictions of wave-induced loads, with a specific focus on wave-induced vertical bending moments. As an initial step in this direction, the vertical bending moment analysis was simplified to consider responses in a regular sea state, where the wavelength matches the vessel's length. The assessed multi-fidelity models include the application of both linear and nonlinear Gaussian processes and compositional kernels to improve predictions of wave-induced loads, specifically focusing on wave-induced vertical bending moments. The case study focuses on the early-stage exploration of the AXE frigates. Multi-fidelity models were constructed using both frequency- and time-domain methods to evaluate the vertical bending moments experienced by the hull. Finally, a critical reflection is provided on how traditional early-stage design processes can be enhanced by integrating multi-fidelity models.

1. Introduction

Currently, engineers face inherent technological challenges that motivate them to strive for the design of vessels with improved capabilities. This is especially true with the push towards higher performing, more sustainable, safer, and autonomous vessels. This research is closely linked to the development of higher-performance and safer ships. From the standpoint of designers, this is where the concept of novelty becomes a central aspect in the early-stage design process. However, the introduction of novel design features introduces a trade-off of increased uncertainty, primarily stemming from limited knowledge of how these innovative features will impact the vessel's performance. This research investigates early-stage design of such a novel concept, the AXE frigate via the use of multi-fidelity (MF) models.

Early-stage design of novel vessels requires revising traditional design processes. Traditionally, early-stage design of complex vessels starts by exploring a broad design space, where various concepts are examined using low-fidelity (LF) analysis (Van Oers et al., 2017). For example, common methods for predicting global loads include using rules from classification societies (Li et al., 2016) and employing linear potential flow (Shigunov et al., 2015). As the design process progresses, the fidelity of our tools incrementally increases. High-fidelity (HF) methods offer greater accuracy but come with higher costs, both computationally and financially. HF analysis is introduced towards the end of the process when only a few promising solutions are further analyzed for detailing and verifying. Examples of HF analysis for predicting global loads include computational fluid dynamics (CFD) (Temarel et al., 2016) and model tests (Bennett et al., 2013).

For novel designs, the goal is to introduce HF analysis earlier on, as LF models often fall short in capturing the complex physics associated with the performance of such concepts. As illustrated by Sapsis (2021), the introduced nonlinearities affect the seakeeping and vertical bending moment (VBM) of the tumblehome hull. However, assessing numerous designs using HF tools is impractical due to high computational costs. Therefore, a promising approach involves employing MF models that integrate various fidelity levels to achieve accurate results with lower

* Corresponding author.
E-mail addresses: N.D.Charisi@tudelft.nl (N.D. Charisi), E.Defer@damennaval.com (E. Defer), J.J.Hopman@tudelft.nl (H. Hopman), A.A.Kana@tudelft.nl (A.A. Kana).

https://doi.org/10.1016/j.oceaneng.2025.123269
Received 1 November 2024; Received in revised form 25 September 2025; Accepted 20 October 2025
Available online 18 November 2025

Acronyms

BEM	boundary element method
BIC	Bayesian Information Criterion
CFD	computational fluid dynamics
DoE	design of experiments
GPs	Gaussian Processes
HF	high-fidelity
LF	low-fidelity
MF	multi-fidelity
NARGP	Nonlinear AutoRegressive Gaussian Process
NURBS	non-uniform rational B-splines
SF	single-fidelity
URANS	unsteady Reynolds-Averaged Navier–Stokes
VBM	vertical bending moment

computational expenses (Peherstorfer et al., 2018; Fernández-Godino, 2023).

Accurately predicting wave-induced loads, such as bending moments and shear forces, is important for the comprehensive evaluation of various design alternatives, facilitating informed decision-making. For example, VBM is important for structural stability and fatigue lifetimes (Guth et al., 2022). Traditionally, engineers rely on established rules and guidelines prescribed by classification societies to guide their design practices (Hirdaris et al., 2014). However, when dealing with novel ships featuring unconventional shapes and sizes, blindly following the class society formulations proves insufficient (Shigunov et al., 2015; Seyffert and Kana, 2019; Parunov et al., 2022). The coupling of unique hull shapes with high service speed can lead to significantly higher bending moments than those determined through classification rules (Shigunov et al., 2015). Thus, it is necessary to adopt direct analysis techniques using HF methods for more accurate predictions.

The VBM emerges as a significant load with substantial implications for ship structural design. It results from the uneven distribution of water pressure and gravity, resulting in the bending of the elastic hull structure (Molland, 2008). Notably, instances of marine accidents, such as the MOL Comfort incident in 2013, have been attributed to extreme bending moments (Jiang, 2015). The vessel broke in two after encountering adverse weather in the Indian Ocean. The severity of such incidents highlights the criticality of comprehending and effectively addressing the challenges associated with the VBM to ensure the safety and structural integrity of vessels.

Despite extensive research on predicting VBM, a gap exists in integrating these methods into early-stage design. This paper focuses on design exploration by considering short-term responses, rather than relying on wave statistics for a single design variation, as is common in most of the relevant literature. This paper addresses the challenge. To introduce HF analysis earlier in the design process for predicting wave-induced VBM in AXE frigates, three different models based on MF Gaussian Processes (GPs) were built and compared. The models are based on the linear autoregressive AR1 scheme, the linear AR1 scheme with the integration of compositional kernels, and the Nonlinear AutoRegressive Gaussian Process (NARGP) scheme. In general, different MF schemes come with their own advantages and disadvantages, making the choice dependent on the specific problem being addressed. In this research, GPs were selected for two key reasons. First, they provide accurate predictions even with small training datasets (Nitzler et al., 2022), which is essential for early-stage design problems where only limited analysis data is available. Second, GPs quantify the uncertainty in their predictions (Rasmussen and Williams, 2005), which supports informed decision-making, a fundamental aspect of early-stage design. However, a common limitation of GPs is their scalability. That said, since early-stage design typically involves small data regimes, this constraint was not relevant for this research problem. The case study involves the assessment of wave-induced VBM for early-stage design exploration of the AXE frigates.

Relevant work regarding the topics of VBM and the design exploration of novel vessels is discussed in Section 2. The mathematical formulation of the underlying methods is given in Section 3. The technical details of the case study are provided in Section 4. The results are presented and discussed in Section 5, and finally, Section 7 is where the conclusions are presented, and recommendations for further research are given.

2. Relevant work

This section provides additional information on the state-of-the-art use of MF models for early-stage design exploration, wave-induced VBM prediction, the applicability of MF models for such predictions, and the AXE frigate concept.

2.1. Design exploration of novel vessels using MF analysis

The initial phase of design is considered the pivotal stage in the design process, as it is during this period that numerous key decisions are formulated (Mavris et al., 1998; Andrews, 2018). A significant challenge is that these critical decisions need to be made while there is a limited knowledge regarding the design (Mavris et al., 1998). To address this challenge, designers ideally want to conduct a broad exploration of the design space. Design exploration requires numerous evaluations, making it common practice to rely on LF methods and tools for these applications. As discussed, early-stage design exploration of novel designs requires the integration of HF analysis data. However, evaluating a large number of designs with HF tools is practically infeasible. A promising avenue is to use MF models. MF models are models combining LF fidelity models with a HF model to enhance accuracy. Currently, MF models are considered state-of-the-art for advanced design applications. This section examines the applicability of MF methods in ship design optimization and exploration.

So far, a substantial portion of relevant research has concentrated on hull optimization, primarily focusing on quantities of interest like resistance or seakeeping. For example, the design optimization of SWATH hull forms was explored by Bonfiglio et al. (2018a, 2020). Bonfiglio et al. (2018a) combined two methods, namely strip theory and a boundary element method (BEM) based on potential flow, to predict the seakeeping performance of hull forms. Bonfiglio et al. (2020) assessed the calm water resistance using a BEM formulated approach, assuming a potential flow-like behavior as the LF model, and a solver based on the unsteady Reynolds-Averaged Navier-Stokes (URANS) equation, serving as the HF model. In both studies, the analysis data were integrated into an MF framework using MF-GPs and Bayesian optimization. The optimization was performed in a 35-dimensional design space for former study and an 8-dimensional design space for the latter. Serani et al. (2022) addressed the design problem of optimizing the DTMB 5415 hull form for seakeeping and resistance. The researchers employed various analysis models, ranging from potential flow to Reynolds-Averaged Navier-Stokes equations, to solve the physical problem. The researchers developed and compared several frameworks based on the following MF approaches: stochastic radial basis functions, kriging partial least squares, augmented expected improvement kriging, and mixed fidelity neural networks. Gaggero et al. (2022) tackled the problem optimizing a marine propeller through two methods-utilizing an inviscid potential flow-based BEM approach as the LF method and employing an inviscid finite volume RANS solver as the HF method. The MF framework was developed based on the coKriging model.

There is a limited number of studies available in the field of structural design. Bonfiglio et al. (2018b) developed a MF design framework relying on GPs and Bayesian optimization for the hydrostructural optimization of supercavitating foils. In the hydrodynamic aspect, the framework integrated data from both coarse and fine resolution URANS methods,

while for the structural component, it incorporated data from both 1D and 3D Finite Element Method simulations. Regarding aerospace design, Bryson et al. (2022) conducted a comparison of various MF approaches, including AR1 GPs with partial least squares, MF Broyden-Fletcher-Goldfarb-Shanno, MF sparse polynomial chaos expansions with Kriging, and Co-Kriging, in the context of an aero-structural design problem. The results showed that the four models exhibited similar performance in terms of reducing the objective value within the allocated computational budget. However, notable distinctions arose due to variations in the initial sample size and distribution.

The aforementioned studies have primarily focused on hydro-structural optimization applications, typically related to later design stages. The authors envision the possibility of advancing such frameworks to earlier stages in the design process to enhance decision-making effectiveness. In ship design, early-stage decision-making involves determining the main design variables and exploring different design concepts.

2.2. Prediction of wave-induced vertical bending moments

The estimation of wave-induced motions and loads is crucial in the design of marine structures (Hirdaris et al., 2016). From the designer's perspective, it is essential to ensure that the global structural strength can withstand operational and environmental loads over the ship's lifetime while considering economic and environmental constraints that drive the development of lighter and more efficient ship structures (Temarel et al., 2016).

In traditional practice, the estimation of global loads in the initial design stages relies on adhering to the regulations of classification societies. While effective for conventional designs, this approach may not properly account for nonlinear effects arising from complex hull geometry and wave interactions, which are critical for unconventional hull forms. Time domain simulations conducted on a tumblehome hull, accounting for nonlinearities in the Froude Krylov force, and hydrostatic restoring pressure forces revealed a complex vertical bending moment curve characterized by asymmetry, attributed to the inherent hull asymmetry (Sapsis, 2021). In addition, Seyffert and Kana (2019) argue that the Equivalent Design Wave method proposed by classification societies is not suitable for novel hull forms like the trimaran when defining lifetime combined loading scenarios.

Numerous methods are available for predicting VBM loads. LF methods encompass classification rules and empirical formulations, such as the method proposed in Murray (1947). A more precise approach involves employing linear potential flow analysis, which takes into account the hull's geometry (e.g., Rajendran et al., 2016). However, in terms of accuracy, the most advanced methods include fully nonlinear URANS models (e.g., Ley and el Moctar, 2021), model tests (e.g., Bouscasse et al., 2022), or onboard measurements during real-world operations (e.g., Waskito et al., 2020). These advanced approaches, while highly accurate, are often impractical for early-stage design processes that may require evaluating a broad design space spanning possibly thousands of design concepts or more.

In general, the analysis methods relevant to predicting VBM loads can be classified into linear and nonlinear methods. Linear methods can yield reliable results for certain problems, but when dealing with a ship operating in stochastic weather conditions, various nonlinearities come into play, as noted in (ITTC, 2014). These nonlinearities, as outlined by Hirdaris et al. (2016), include those associated with the wetted surface of the body, the free surface, the seabed, and the remaining surfaces bounding the fluid domain. Linear theories inherently suggest that sagging and hogging moments are equal. However, empirical evidence from experiments and full-scale measurements, e.g., Fonseca and Soares (2002), has demonstrated that sagging exceeds hogging. The transition from straightforward methods like class rules to more sophisticated and computationally costly techniques is primarily propelled by shifts in ship types, sizes, complexities, the availability of additional data regarding real ship responses in waves, and enhanced computational capabilities (Temarel et al., 2016).

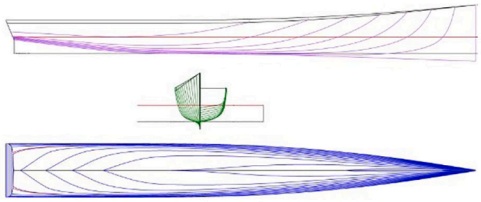
Numerous recent studies aim to develop MF models that facilitate the analysis of a specific design under various loading conditions the vessel may encounter throughout its lifetime. These predictions, encompassing various loading scenarios, undergo statistical treatment to establish a design load. For example, Guth et al. (2022) introduced a MF active search method applied to a ship-wavegroup problem using the ONR Topside series flared variant with constant velocity through long-crested head seas. The research demonstrated that, in specific cases, incorporating LF data enhances the prediction of the probability density functions of the responses. Another study by Drummen et al. (2022) proposed a MF approach, integrating model tests with results from linear potential flow, to predict design loads linked to the extreme VBM. The findings underscored the significance of considering nonlinearities in such problems, revealing substantial corrections to the linear results. For instance, a 40 % increase in the linear long-term sagging occurred due to weak nonlinearities. To the best of the author's knowledge, there is a current gap in research concerning the assessment of wave-induced VBM from a design exploration perspective. Traditionally, such analyses occur later in the design process, once the primary design variables are already defined. However, we argue that VBM is a crucial quantity of interest that should be considered earlier in the design process.

To determine the design load associated with VBM, the vessel must be analyzed across various sea states. Additionally, the vessel's operational profile should be incorporated to calculate long-term responses over the vessel's lifetime. This process is computationally intensive and can only be performed for a very limited number of design variations. This research focuses on early-stage design, where the goal is to explore a broad design space. As an initial step toward incorporating VBM analysis in this phase, the analysis is simplified to a single regular sea state, with the vessel's length matching the wavelength. This decision was made to: (1) enhance the interpretability of load prediction results for different design variations and (2) shift the research focus toward developing MF models for assessing the design space rather than refining analysis methods for load prediction.

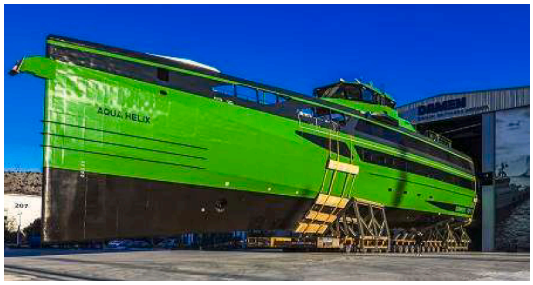
2.3. AXE frigate concept

As previously noted, the primary objective of this study is to facilitate early-stage design of novel vessels. The implementation of the AXE bow (Fig. 1) to large vessels such as frigates represents an innovative concept. The AXE bow was originally developed by Keuning et al. (2002). The AXE bow has the potential to enhance the vessel's performance in terms of seakeeping capabilities, making it an attractive choice for frigates, as they need to execute their missions effectively even in adverse weather conditions. The AXE bow has been implemented in high-speed and relatively small vessels like yachts and offshore patrol vessels, such as the Aqua Helix Fig. 1(b). However, its application has not extended to larger vessels like frigates as of now.

Recent studies have demonstrated the hydrodynamic advantages of incorporating the AXE bow design on a frigate. One of the initial studies conducted by Eefsen et al. (2004) explored seven alternatives for hull forms and concepts derived from the parent hull of a 120 m frigate. The evaluation of these alternatives involved an assessment of their seakeeping performance, operability, and resistance, relying on experimental data. A key finding from the study is that the incorporation of the AXE bow does not compromise dynamic stability or course-keeping capabilities. Keuning and Van Wiilree (2006) conducted a comparative analysis of the hydrodynamic characteristics of three offshore patrol vessel concepts: the enlarged ship concept, the AXE bow concept, and the Wave Piercer concept. These concepts were evaluated as part of the conceptual design process for a 55 m fast monohull patrol vessel capable of reaching speeds of 50 knots. The findings revealed that the AXE bow concept outperformed the other vessels. The results were derived from model tests assessing calm water resistance, as well as the vessels' behavior in both



(a) AXE 4100 lines plan developed Keuning et al. [2015]



(b) Aqua Helix [Allen, 2021]

Fig. 1. AXE bow.

head and following waves. In the recent study by Rijkens and Mikelic (2022), a hydrodynamic comparison was conducted between two frigate designs: a conventional frigate and an AXE frigate. Model tests were carried out in calm water, regular waves, and irregular waves to discern differences in seakeeping performance. The experimental findings indicate that the calm water resistance of the AXE frigate is approximately 9 % lower at the design speed, primarily attributed to the longer waterline length of the AXE frigate. For fast vessels like frigates, residuary resistance makes up a significant portion of the total resistance, and it decreases as the ship's length increases (Papanikolaou, 2014). Tests in regular and irregular head waves reveal reduced heave and pitch motions for the AXE frigate. Additionally, in large waves, the AXE frigate exhibits reduced deck wetness due to its higher freeboard height in the bow area.

Hence, the AXE bow concept holds promise for its integration into a frigate design. This study will further investigate its impact on wave-induced VBM from the perspective of design exploration.

3. Methods

This section provides the mathematical formulation of the methods used in this research. Section 3.1 offers a comprehensive overview of the design framework itself. Section 3.2.1 provides the mathematical formulation of MF-GPs, including compositional kernels. Section 3.3 discusses the technical details of the tools used for predicting the wave-induced VBM.

3.1. Overview of the design framework

The design framework is structured around three primary components: the generation, analysis, and optimization engines. This framework structure is commonly employed in design optimization frameworks, such as the design engineering engine (La Rocca, 2012). In the context of the generation engine, a parametric model was set up to facilitate the exploration of the key design variables. These include the length, beam, depth, deadrise, and flare angles, along with three variables linked to bow design: bow length, depth, and height. In addition, this component should produce a mesh of the hull, which will be further analyzed using the dedicated solvers. Further technical details are given in Section 4.1.

The analysis engine includes both analysis tools for assessing wave-induced VBM and methods for constructing the surrogate model of the objective landscape. The included analysis tools consist of two solvers with distinct fidelities: the frequency domain solver PRECAL (Section 3.3.1), used as the LF tool, and the time domain solver PRETTI (Section 3.3.2), used as the HF tool. One of the most notable differences is that the time domain solver is capable of capturing the nonlinearities introduced by the hull shape. For creating the surrogate model of the objective landscape, three different methods will be employed and compared. First, the initial approach involves the AR1 scheme of MF-GPs, introduced by Kennedy and O'Hagan (2000) using the squared exponential kernel. The second model consists of the AR1 scheme of MF-GPs with compositional kernels, a method introduced and investigated by the authors (Charisi et al., 2022, 2024b). The compositional kernels were initially proposed by Duvenaud et al. (2013). The third model is the nonlinear scheme of the MF-GPs proposed by Perdikaris et al. (2017). The mathematical formulation of the aforementioned methods is given in Sections 3.2.1 and 3.2.2. The final component is the optimization engine. The optimization engine utilizes the surrogate model generated by the analysis engine and employs a search strategy to find the optimum solution. The technical details of the optimization engine implementation are beyond the scope of this paper; instead, the focus is on how effectively the different methods approximate the objective landscape. An overview of how all the individual blocks of the design framework fit together is shown in the flowchart in Fig. 2.

3.2. Multi-fidelity Gaussian processes and compositional kernels

This section gives the mathematical formulation of GPs, MF-GPs (the linear scheme AR1 and the nonlinear scheme NARGP), and compositional kernels. In general, frameworks based on the AR1 scheme assume a linear correlation between fidelity levels, represented by a scaling function and an additive function. In contrast, the NARGP scheme captures a nonlinear relationship using a function alongside the additive term. The objective of the paper is to evaluate which scheme more effectively models the inter-fidelity relationship in the case of predicting the VBM for the AXE frigate.

3.2.1. Multi-fidelity Gaussian processes

The mathematical formulation for GPs follows the work presented in Rasmussen and Williams (2005). GPs are used to build approximations of real-world processes $f(x)$, which can be fully defined by a mean $\mu(x)$ and a covariance function $k(x, x')$ according to Eqs. (1)–(3). The mean function $m(x)$ defines the expected value for individual random variables within the function $f(x)$. Additionally, it can integrate domain knowledge when such information is accessible. In situations where this knowledge is not available, we can make the assumption that $m(x) = 0$ (Rasmussen and Williams, 2005), since data can be normalized to have a zero mean. Regarding the covariance matrix, each element represents the covariance between two random variables, namely, $f(x)$ and $f(x')$. These elements are determined by the kernel function $k(x_i, x_j; \theta)$, where the kernel function relies on the model's parameters denoted as θ, also referred to as hyperparameters.

$$f(x) \sim \mathcal{GP}(m(x), k(x, x')) \tag{1}$$

$$m(x) = \mathbb{E}[f(x)] \tag{2}$$

$$k(x, x') = \mathbb{E}[f(x) - m(x)][f(x') - m(x')] \tag{3}$$

The available analysis or experimental data can be described according to Eq. (4):

$$y = f(x) + \epsilon, \epsilon \sim \mathcal{N}(0, \sigma_n^2 \mathcal{I}) \tag{4}$$

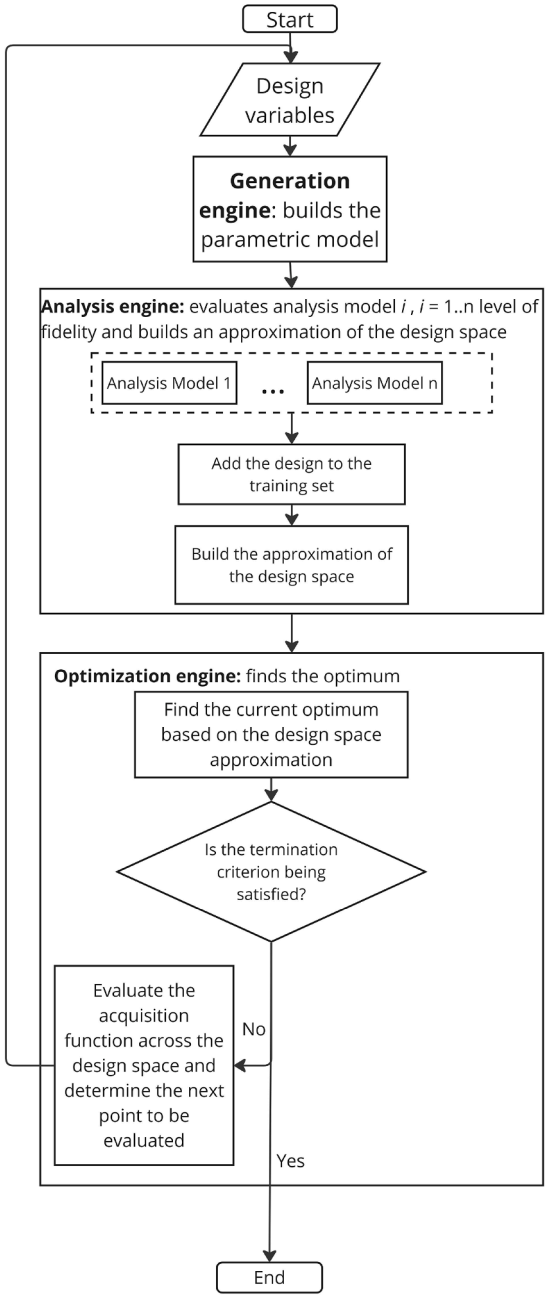


Fig. 2. Flowchart framework for the early-stage design assessment of the wave-induced VBM of AXE frigates.

where f represents the function to be approximated, and ϵ stands for the error term. In Bayesian methods like GPs, a pivotal aspect of the analysis is the prior distribution. This distribution encapsulates our pre-existing knowledge or assumptions about the unknown function f. The prior distribution for both the observed data X and the test data X_* is established as per Eq. (5):

$$\begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{f}_* \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} K(X,X)+\sigma_n^2 I & K(X,X_*) \\ K(X_*,X) & K(X_*,X_*) \end{bmatrix} \right) \tag{5}$$

where f_* are the function values evaluated at the test locations X_*. The GP prior forms a distribution over functions since the GP is a multivariate Gaussian distribution over the random variable vector $\begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{f}_* \end{bmatrix}$ which leads to different probabilities for different functions.

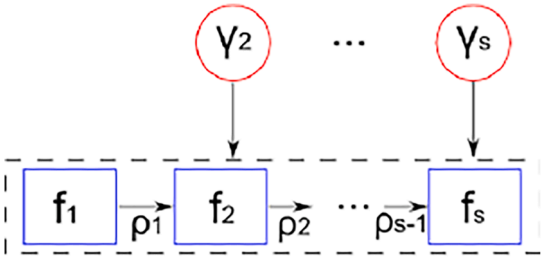


Fig. 3. Schematic representation AR1 (Brevault et al., 2020).

Within Bayesian learning, the prior distribution undergoes modification by assimilating the observed data, leading to the emergence of the posterior distribution. Essentially, the probabilities assigned to different functions according to the prior are adjusted based on the observed data through the likelihood. In mathematical terms, the prior distribution is conditioned on the observed data to yield the posterior distribution, as described by Eqs. (6)–(8):

$$\boldsymbol{f}_*|X, X_*, \boldsymbol{y} \sim \mathcal{N}(\bar{\boldsymbol{f}}_*, \mathrm{cov}(\boldsymbol{f}_*)) \tag{6}$$

$$\bar{\boldsymbol{f}}_* = \boldsymbol{k}_*^T [K + \sigma_n^2 I]^{-1} \boldsymbol{y} \tag{7}$$

$$\mathrm{cov}(\boldsymbol{f}_*) = k(x_*, x_*) - \boldsymbol{k}_*^T [K + \sigma_n^2 I]^{-1} \boldsymbol{k}_* \tag{8}$$

where $K = K(X, X)$, and $k_* = k(x_*)$. Numerous approaches exist for optimizing kernel hyperparameters, including techniques like cross-validation and maximum likelihood estimation (e.g. Bachoc, 2013). In the present work, the maximization of the marginal log-likelihood, Eq. (9), was implemented. The single-fidelity (SF) model examined was built using GPs.

$$\log p(\boldsymbol{y}|X) = -\frac{1}{2}\log|K + \sigma_n^2 I| \\ -\frac{1}{2}\boldsymbol{y}^T[K + \sigma_n^2 I]^{-1}\boldsymbol{y} - \frac{n}{2}\log 2\pi \tag{9}$$

The autoregressive scheme AR1 illustrated in Fig. 3, as introduced by Kennedy and O'Hagan (2000), assumes a linear dependency among various fidelity models. The mathematical formulation follows the description in Gratiet and Garnier (2014). Due to the bifidelity nature of the case study, the mathematical description of the bifidelity model is given. The interconnection of sub-models is governed by Eqs. (10) and 11. Specifically, the HF function connects to the LF function through a scaling function ρ and an additive function δ. It is assumed that f_2 corresponds to the HF function, while f_1 corresponds to the LF function. The function δ is a GP independent of f_1 as shown in Eq. (11).

$$f_2(\boldsymbol{x}) = \rho(\boldsymbol{x}) f_1(\boldsymbol{x}) + \delta(\boldsymbol{x}) \tag{10}$$

$$f_1(\boldsymbol{x}) \perp \delta(\boldsymbol{x}) \tag{11}$$

The predictive model is a multivariate normal distribution described by Eq. (12), with a mean function according to Eq. (13) and a variance according to Eq. (14).

$$f_{2*}|(\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \rho), (\sigma_1^2, \sigma_2^2), (\boldsymbol{\theta_1}, \boldsymbol{\theta_2}) \sim \mathcal{N}(\bar{\boldsymbol{f}}_{2*}, \mathrm{cov}(\boldsymbol{f}_{2*})) \tag{12}$$

$$\bar{\boldsymbol{f}}_{2*} = \mathbf{h}(\boldsymbol{x})^T \boldsymbol{\beta} + \boldsymbol{k}_*^T \boldsymbol{V}^{-1}(\boldsymbol{f} - \boldsymbol{H}\boldsymbol{\beta}) \tag{13}$$

$$\mathrm{cov}(\boldsymbol{f}_{2*}) = \rho^2\sigma_1^2 + \sigma_2^2 - \boldsymbol{k}_*^T \boldsymbol{V}^{-1} \boldsymbol{k}_* \tag{14}$$

where the trend parameters $\boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta_1} \\ \boldsymbol{\beta_2} \end{pmatrix}$, and $\boldsymbol{f} = \begin{pmatrix} \boldsymbol{f_1} \\ \boldsymbol{f_2} \end{pmatrix}$. The variance parameters σ_1^2, σ_2^2 and the parameters $\boldsymbol{\theta_1}, \boldsymbol{\theta_2}$ are the model's hyperparame-

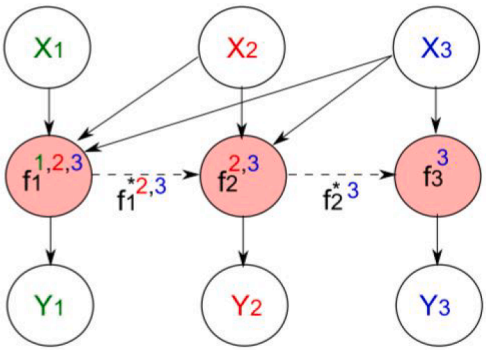


Fig. 4. Schematic representation NARGP (Brevault et al., 2020).

ters.

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{f}_1'(x_1^{(1)}) & 0 \\ \vdots & \\ \boldsymbol{f}_1'(x_{n_1}^{(1)}) & 0 \\ \rho \boldsymbol{f}_1'(x_1^{(2)}) & \boldsymbol{f}_2'(x_1^{(2)}) \\ \vdots & \vdots \\ \rho \boldsymbol{f}_1'(x_{n_2}^{(2)}) & \boldsymbol{f}_2'(x_{n_2}^{(2)}) \end{pmatrix} \tag{15}$$

$$\boldsymbol{h}(x)^T = (\rho \boldsymbol{f}_1'(x), \boldsymbol{f}_2'(x)) \tag{16}$$

The covariance matrix is calculated as described in Eq. (17),

$$\boldsymbol{V} = \begin{pmatrix} \sigma_1^2 R_1(\boldsymbol{D_1}) & \rho\sigma_1^2 R_1(\boldsymbol{D_1}, \boldsymbol{D_2}) \\ \rho\sigma_1^2 R_1(\boldsymbol{D_2}, \boldsymbol{D_1}) & \rho^2\sigma_1^2 R_1(\boldsymbol{D_2}) + \sigma_2^2 R_2(\boldsymbol{D_2}) \end{pmatrix} \tag{17}$$

where $\boldsymbol{D_1} = \{x_1^{(1)} \dots x_{n_1}^{(1)}\}$, and $\boldsymbol{D_2} = \{x_1^{(2)} \dots x_{n_2}^{(2)}\}$.

The nonlinear scheme for GPs, NARGP, as shown in Fig. 4, was proposed by Perdikaris et al. (2017). The authors state that the scheme 'enables the construction of flexible and inherently nonlinear and non-Gaussian multi-fldelity information fusion algorithms'. The mathematical formulation allows for a nonlinear mapping between the fidelities, as described in Eq. (18).

$$f_2(\boldsymbol{x}) = z_1(f_1(\boldsymbol{x})) + \delta(\boldsymbol{x}) \tag{18}$$

where $z(\cdot)$ is a function that maps the low fidelity data to the higher fidelity data. A GP prior distribution is assigned to both terms z_1 and f_1, thereby classifying the term $z_1(f_1(\boldsymbol{x}))$ as a deep GP. Therefore, the posterior distribution $f_2(\boldsymbol{x})$ is not gaussian. The GP prior f_1 is being replaced by the GP posterior f_{*1}. Thus, the scheme can be written as shown in Eq. (19). The function g_2 follows a GP as shown in Eq. (20).

$$f_2(\boldsymbol{x}) = g_2(\boldsymbol{x}, f_{*1}(\boldsymbol{x})) \tag{19}$$

$$g_2 \sim \mathcal{GP}(f_2|0, k_2((\boldsymbol{x}, f_{*1}(\boldsymbol{x})), (\boldsymbol{x}', f_{*1}(\boldsymbol{x}'))) \tag{20}$$

Since the posterior distribution of f_2 is not Gaussian, it should be computed according to Eq. (21).

$$p(f_{*2}(\boldsymbol{x})) = \int p(f_{*2}(\boldsymbol{x}, f_{*1}(\boldsymbol{x})))p(f_{*1}(\boldsymbol{x}))\, d\boldsymbol{x} \tag{21}$$

where $p(f_{*1}(\boldsymbol{x}))$ is the posterior distribution of the lower fidelity level. Monte Carlo techniques are employed to approximate the posterior predictive mean and variance. For the NARGP implementation, quasi-Monte Carlo sampling was applied with the default setting of 100 samples to estimate the predictive distribution.

3.2.2. Compositional kernels

The covariance matrix expresses the degree of similarity between data points (Rasmussen and Williams, 2005) and incorporates prior knowledge about the function f. Previous research has explored basis functions serving as kernel functions, which generate valid covariance matrices. Ensuring the covariance matrix's validity necessitates both symmetry and positive semi-definiteness. Compositional kernels, as outlined by Duvenaud et al. (2013), refer to combinations of a finite set of basis kernel functions via addition or multiplication. The selection of these basis kernels aims to mathematically encapsulate the key features of the function f or, in the context of this specific research problem, the objective landscape.

In this study, the framework introduced in Charisi et al. (2022) was employed. The fundamental concept involved integrating compositional kernels into the AR1 scheme, with the aim of enhancing early-stage design analysis and optimization by uncovering the structure of the underlying objective landscape. For a problem involving models with fidelities ranging from 1 to s, where fidelity 1 represents the lowest fidelity model, the first step is to build a compositional kernel for the lowest fidelity data (fidelity 1) using a single fidelity GP model. For each fidelity i ranging from 2 to s, a compositional kernel is constructed based on the bi-fidelity GP model, utilizing fidelity i data and fidelity $i+1$ data. The Bayesian Information Criterion (BIC) serves as the objective function, following the approach outlined by Duvenaud et al. (2013). The BIC is formally defined by Eq. (22).

$$BIC = k_{\text{hyp}} \ln n - 2 \ln L \tag{22}$$

where n is the number of training data, k_{hyp} is the number of hyperparameters, and L is the maximized likelihood value. BIC comprises two elements: a penalty term determined by the count of model parameters and a term derived from the likelihood function. The advantage of opting for BIC over maximizing the marginal log-likelihood is its attention to the complexity of the kernel function. By favoring functions with fewer hyperparameters, BIC aids in preventing overfitting. The integration of compositional kernels to the AR1 scheme to facilitate design analysis can be found in (Charisi et al., 2024b).

3.3. Vertical bending moment analysis during early design stages

As mentioned, the VBM arises from the disparity in load distribution between the weight and wave pressure along the ship's length. The physical mechanism can be explained as follows: the distribution of load, denoted as $q(x)$, induces the generation of shear force $Q(x)$ (Eq. (23)), which subsequently results in bending moments $M(x)$ (Eq. (24)), where x defines the position along the length of the vessel.

$$Q(x) = -\int q(x)\, dx \tag{23}$$

$$M(x) = -\int q(x) \cdot x\, dx \tag{24}$$

To predict the wave-induced VBM, a frequency domain method is used as the LF model, and a time domain method is used as the HF model.

3.3.1. Frequency domain analysis

To compute the vertical bending moment in the frequency domain, PRECAL software was used. The software was developed by MARIN (MARIN, 2019). PRECAL is a specialized tool designed for predicting linear responses through potential flow calculations. The tool functions in the following manner: (1) it divides the wetted hull into multiple quadrilateral panels, (2) it calculates hydrodynamic coefficients by solving the linearized boundary value problem, and (3) it determines ship motions and loads using linearized potential flow theory, incorporating adjustments for viscous damping through empirical corrections.

3.3.2. Time domain analysis

For time domain, PRETTI_R (Register, 2019) was used, which is a 3D time-domain nonlinear seakeeping and The present study did not investigate slamming and whipping in depth, as this lies beyond the scope of the paper.hydroelasticity tool. Unlike PRECAL, PRETTI_R is specifically designed to predict motions in high sea states, encompassing rigid-body motion, and hydrodynamic loads. Additionally, it can take into account

Table 1
Design variables ranges.

Design parameters	Value	Units
Length (L)	98–120	m
Beam (B)	13–16	m
Depth (D)	12–16	m
Coefficient length forehull (α_L)	0.4-0.8	–
Coefficient height AXE bow (α_H)	0.0-0.5	–
Coefficient depth AXE bow (α_D)	0.0-0.5	–
Deadrise angle ($\phi_{deadrise}$)	5–45	deg
Flare angle (ϕ_{flare})	0–25	deg

Table 2
Parameters for DS.A.

Design parameters	Value	Units
α_H	0.3	–
α_D	0.3	–
α_L	0.6	–
u_{test}	[0,15]	knots

slamming and whipping loads. The present study did not investigate slamming and whipping in depth, as this lies beyond the scope of the paper. This software was developed as part of the Cooperative Research Ships initiative.

The Froude Krylov force is determined by integrating incident wave hydrodynamics and hydrostatic pressure across the vessel's hull surface. The diffraction force is estimated by scaling the frequency domain diffraction force response amplitude operator with the incident wave amplitude. The radiation force is computed through a convolution integral involving an impulse function. PRETTI_R leverages frequency domain results to derive the necessary impulse functions.

4. Case study

To showcase the results, three design case studies were employed. The initial case study focused on exploring a two-dimensional design space, encompassing the hull's length L and beam B. This study serves as a proof of concept as it enables the visualization of objective landscape and the identification of design trends. The second case study focuses on the design of the AXE bow. In this case, a 3D problem is tackled, addressing the length L_{fore}, height H_{axe}, and depth D_{axe} of the AXE bow. The third case study focuses on a larger-scale problem and involves eight design variables. The design variables and their corresponding value ranges can be found in Table 1. The ranges were determined based on industrial experience with realistic frigate design. The primary design parameter is the vessel's displacement, which remains constant at 6000 tons across all design scenarios.

Sensitivity analysis, which is particularly useful when dealing with a large number of design variables, was not performed to identify which variables most significantly influence the wave-induced VBM. In this study, the design exploration problem was limited to a few key variables related to the vessel's main dimensions and the shape of the AXE bow, keeping the number of design variables relatively small. Additionally, the objective was to increase the dimensionality of the design problem to assess the accuracy of various MF models as the problem's dimensionality grows.

For the development of the MF models, the Python packages used included GPy (GPy, 2012) and Emukit (Paleyes et al., 2019, 2023). The accuracy of the four MF models was assessed based on two statistical measures, namely the R-Squared (R^2) and the normalized Root Mean Square Error, (RMSE), which are expressed as follows according to Eqs. (25) and (26). The comparison was intended to assess the relative accuracy between the developed MF models, rather than to benchmark them against other MF methods reported in the literature. As noted earlier, the analysis of this paper for the assessment of the VBM focused on a simplified case involving a single wave loading condition. In contrast, to the best of the authors' knowledge, existing literature focuses on evaluating a single design on multiple sea states. Consequently, a direct comparison with existing studies is not applicable. The Latin hypercube sampling technique was used to select analysis points, and a total of 20 different design of experiments (DoE) were employed to compute statistics related to prediction errors.

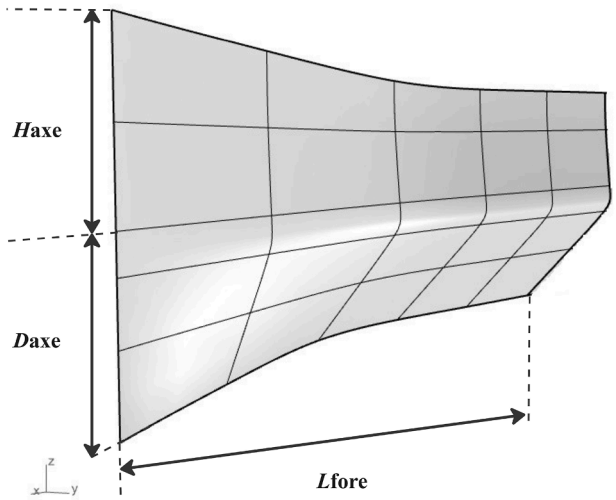


Fig. 5. Visualization of the design parameters for the AXE frigate design.

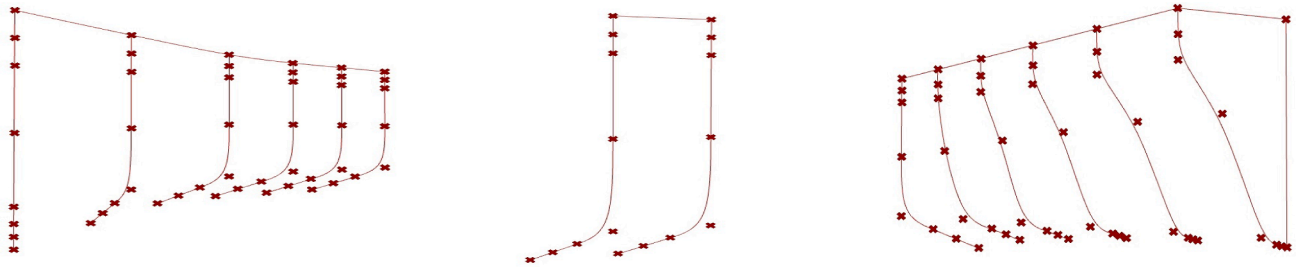
$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{S}(y_i - \overline{y})^2} \tag{25}$$

$$\text{RMSE} = \frac{1}{y_{max} - y_{min}} \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{26}$$

where y_i refers to the observed value of each data point i, $\hat{y}_i$ refers to the predicted value of each data point i, $\overline{y}$ refers to the mean of the observed values for all the data points, and N refers to the total number of the samples. Both errors are used to provide a more comprehensive assessment of the prediction quality. More specifically, R^2 quantifies the proportion of variability in the dependent variable that can be explained by the independent variable (James et al., 2014). In linear models, it provides an intuitive measure since its value ranges from 0 to 1 (Spiess and Neumeyer, 2010). However, for nonlinear models like GPs, R^2 is not restricted to this range and can take negative values (Cameron and Windmeijer, 1997). A negative R^2 indicates that the model performs worse than simply using the mean of the predicted values. Since R^2 alone is insufficient for a comprehensive performance assessment, it was used alongside RMSE, which quantifies model accuracy based on residual errors. Although various other metrics could have been considered, the combination of these two provides a sufficient understanding of the model's performance. R^2 offers insight into the model's explanatory capabilities, while RMSE captures the magnitude of prediction errors. The combination of these two metrics measures both interpretability and precision.

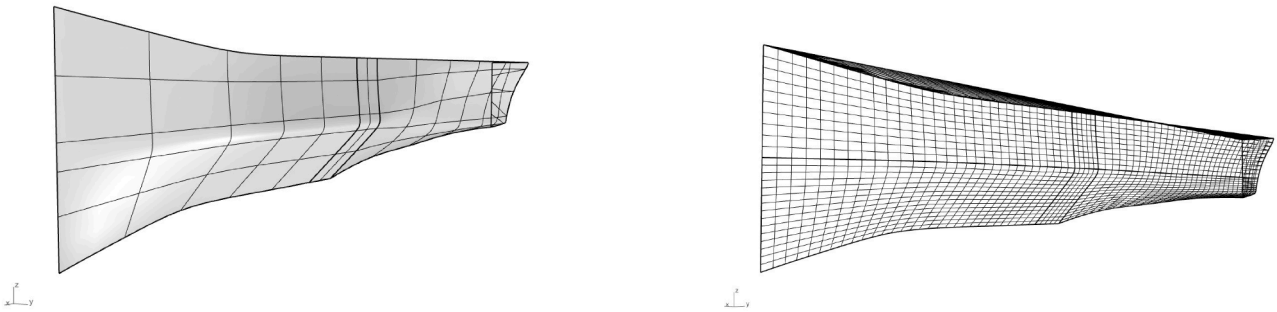
4.1. Parametric models

The parametric model for the AXE frigates was created to capture the primary design variables that hold significance for exploration during the early design stages. These variables, summarized in Table 1, include the main particulars of the vessel, the deadrise and flare angle



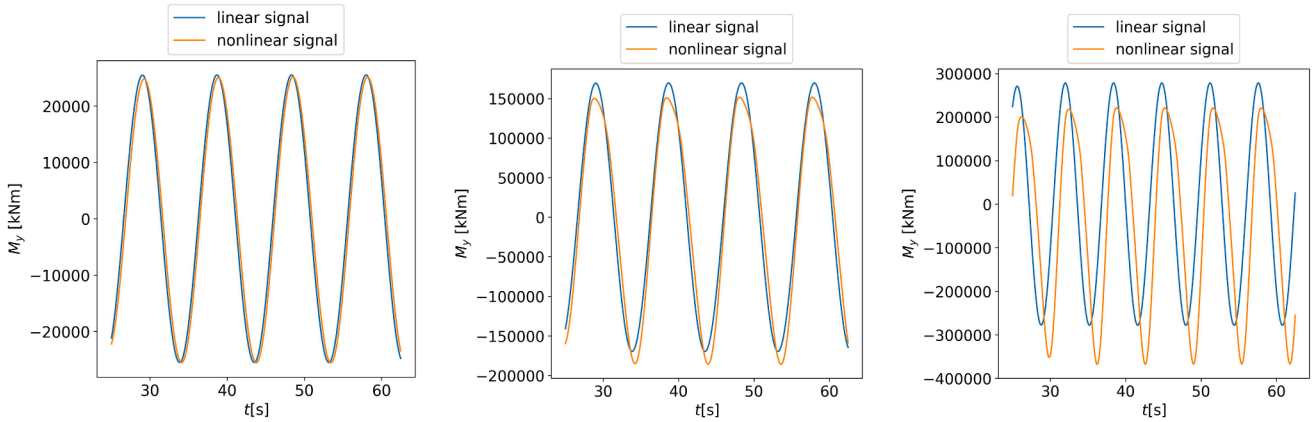
(a) Forehull cross sections (b) Midhull cross sections (c) Afthull cross sections

Fig. 6. Hull cross sections.



(a) Visualizing the hull surfaces (b) Visualizing the hull's mesh

Fig. 7. Parametric model example.



(a) v=0 knots, H_w=1m (b) v=0 knots, H_w=6.65m (c) v=15 knots, H_w=6.65m

Fig. 8. Loading cases.

of the hull, and the design variables of the AXE bow. The design variables related to the AXE bow are described by Eqs. (27)–(29) and are illustrated in Fig. 5.

$$L_{\text{fore}} = \alpha_{\text{L}} \cdot L \tag{27}$$

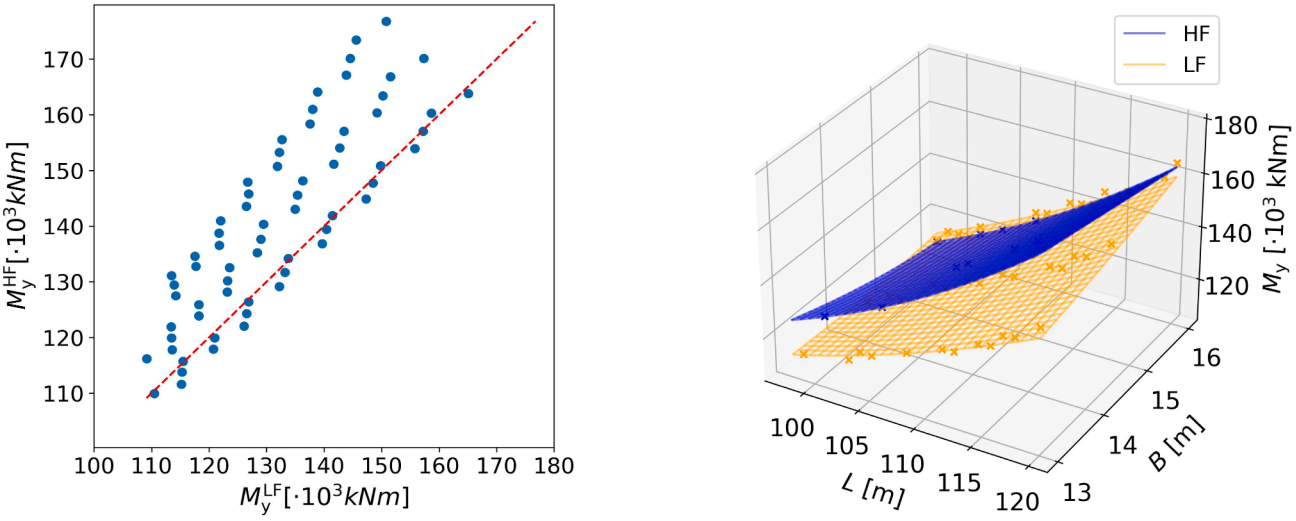
$$H_{\text{axe}} = \alpha_{\text{H}} \cdot D \tag{28}$$

$$D_{\text{axe}} = \alpha_{\text{D}} \cdot D \tag{29}$$

The parametric model was developed using Rhino (Associates, 2023b) and Grasshopper (Associates, 2023a). The initial step in constructing the parametric model is to determine the values of the design variables that guide the formation of the hull sections. The primary midhull cross sections were shaped following the mathematical formulation adapted from Filipa Marques Sanches (2016). For a visual representation, refer to the examples of the cross sections in Fig. 6(a) (forehull), Fig. 6(b) (midhull), and Fig. 6(c) (afthull). After defining the design variables, the control points are established. The cross-sections are built by fitting a non-uniform rational B-splines (NURBS) curve through them. Each cross section is characterized by 8 control points. Subsequently, the hull surfaces are generated based on the NURBS curves. The final step involves properly meshing the hull Fig. 7(b) for further analysis by numerical solvers.

4.2. Wave loading conditions

The wave loading conditions are determined separately for each design variation. The aim is to analyze each vessel in a sea state that in-



(a) Correlation between fidelities

(b) Visualization of the objective landscape

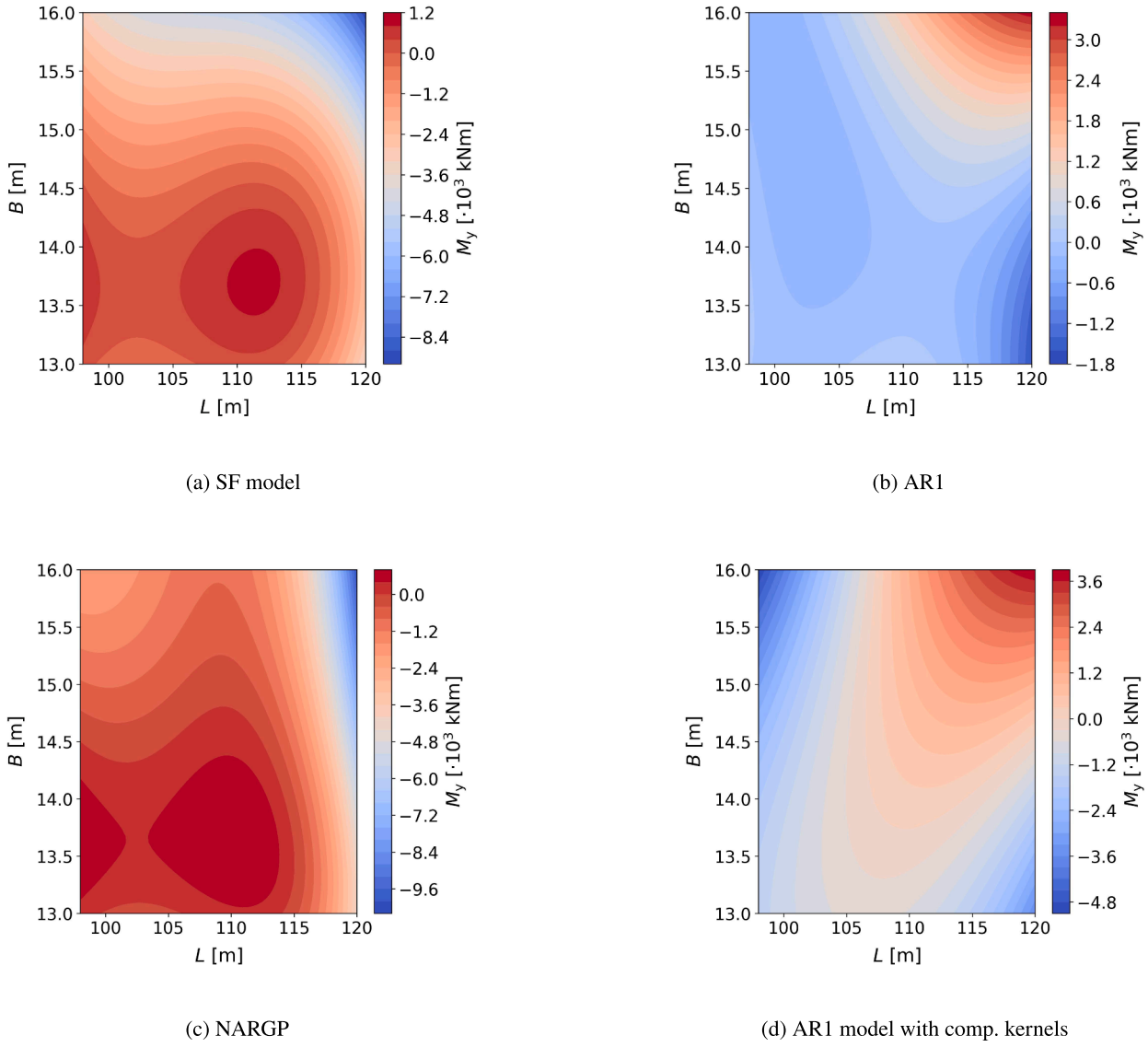
Fig. 9. $2D$ design case, $v = 0$ knots.

Table 3
Error metrics for the $2D$ design case, $u = 0$ knots.

DoE	GP HF R2 (std)	GP HF RMSE (std)	AR1 R2 (std)	AR1 RMSE (std)	NARGP R2 (std)	NARGP RMSE (std)	AR1 with comp kernels R2 (std)	AR1 with comp kernels RMSE (std)
	−0.6944	0.3226	−0.4070	0.2299	−0.6225	0.3142	−13.2166	0.3953
(2,40)	(0.8290)	(0.0735)	(3.7136)	(0.1942)	(0.8528)	(0.0780)	(59.2206)	(0.8682)
	0.6240	0.1246	0.7653	0.0960	0.7087	0.1105	0.5969	0.1516
(4,40)	(0.4989)	(0.0976)	(0.3089)	(0.0767)	(0.3573)	(0.0850)	(0.2471)	(0.0560)
	0.9869	0.0259	0.9602	0.0339	0.9905	0.0229	0.8006	0.0837
(6,40)	(0.0140)	(0.0143)	(0.0018)	(0.0387)	(0.0070)	(0.0101)	(0.3600)	(0.0786)
	0.9962	0.0128	0.9946	0.0147	0.9966	0.0135	0.9829	0.0265
(8,40)	(0.0059)	(0.0094)	(0.0094)	(0.0119)	(0.0030)	(0.0065)	(0.0245)	(0.0213)
	0.9944	0.0130	0.9978	0.0103	0.9966	0.0131	0.9826	0.0175
(10,40)	(0.0146)	(0.0135)	(0.0025)	(0.0065)	(0.0035)	(0.0076)	(0.0642)	(0.0288)
	0.9956	0.0088	0.9945	0.0098	0.9988	0.0078	0.9987	0.0079
(12,40)	(0.0165)	(0.0138)	(0.0208)	(0.0155)	(0.0015)	(0.0041)	(0.0020)	(0.0048)
	0.9993	0.0059	0.9993	0.0064	0.9991	0.0073	0.9992	0.0066
(14,40)	(0.0010)	(0.0034)	(0.0008)	(0.0032)	(0.0007)	(0.0030)	(0.0010)	(0.0036)
	0.9997	0.0042	0.9996	0.0045	0.9997	0.0046	0.9996	0.0050
(16,40)	(0.0002)	(0.0012)	(0.0004)	(0.0021)	(0.0002)	(0.0012)	(0.0003)	(0.0018)
	0.9997	0.0040	0.9998	0.0038	0.9997	0.0042	0.9997	0.0041
(18,40)	(0.0001)	(0.0011)	(0.0002)	(0.0012)	(0.0003)	(0.0018)	(0.0002)	(0.0015)
	0.9998	0.0040	0.9998	0.0038	0.9997	0.0043	0.9997	0.0044
(20,40)	(0.0001)	(0.0011)	(0.0001)	(0.0011)	(0.0002)	(0.0013)	(0.0002)	(0.0014)

Table 4
Error metrics for the $2D$ design case, $u = 15$ knots.

DoE	GP HF R2 (std)	GP HF RMSE (std)	AR1 R2 (std)	AR1 RMSE (std)	NARGP R2 (std)	NARGP RMSE (std)	AR1 with comp kernels R2 (std)	AR1 with comp kernels RMSE (std)
	−0.6100	0.2796	−3.0363	0.3620	−0.5796	0.2776	−2,5180	0.3380
(2,40)	(1.1858)	(0.0959)	(7.4229)	(0.3240)	(1.1556)	(0.0943)	(6.7976)	(0.3009)
	0.8009	0.1013	0.4485	0.1675	0.7010	0.1222	0.4626	0.1678
(4,40)	(0.1091)	(0.0257)	(0.3003)	(0.0481)	(0.1951)	(0.0377)	(0.2441)	(0.0349)
	0.8089	0.1003	0.6488	0.1338	0.7688	0.1099	0.4506	0.1701
(6,40)	(1.1219)	(0.0335)	(0.2268)	(0.0455)	(0.1385)	(0.0355)	(0.2935)	(0.0486)
	0.7754	0.1044	0.7875	0.1044	0.7276	0.1118	0.5070	0.1539
(8,40)	(0.2432)	(0.0454)	(0.1389)	(0.0333)	(0.2990)	(0.0577)	(0.4847)	(0.0698)
	0.8315	0.0937	0.7800	0.1080	0.8007	0.1016	0.6265	0.1410
(10,40)	(0.1152)	(0.0251)	(0.1205)	(0.0290)	(0.1317)	(0.0326)	(0.1793)	(0.0350)
	0.8220	0.0980	0.7327	0.1149	0.8241	0.0979	0.6341	0.1347
(12,40)	(0.1129)	(0.0283)	(0.2407)	(0.0456)	(0.0975)	(0.0278)	(0.3125)	(0.0577)
	0.8461	0.0933	0.8417	0.0951	0.8341	0.0967	0.7159	0.1235
(14,40)	(0.0936)	(0.0278)	(0.0782)	(0.0262)	(0.1047)	(0.0317)	(0.2082)	(0.0534)
	0.8832	0.0812	0.8705	0.0847	0.8815	0.0812	0.7627	0.1044
(16,40)	(0.0403)	(0.0152)	(0.0568)	(0.0192)	(0.0600)	(0.0225)	(0.2950)	(0.0570)
	0.8771	0.0836	0.8335	0.0951	0.8760	0.0844	0.8021	0.1061
(18,40)	(0.0480)	(0.0178)	(0.1153)	(0.0342)	(0.0465)	(0.0186)	(0.0811)	(0.0246)
	0.8693	0.0845	0.8523	0.0893	0.8872	0.0790	0.8326	0.0932
(20,40)	(0.0490)	(0.0178)	(0.0703)	(0.0201)	(0.0382)	(0.0162)	(0.1161)	(0.0274)



(a) SF model

(b) AR1

(c) NARGP

(d) AR1 model with comp. kernels

Fig. 10. Visualization of the models' error for the zero-speed $2D$ case with 6 HF points.

Table 5
Parameters for DS.B.

Design parameters	Value	Units
L	120.0	m
B	16.0	m
D	17.0	m
ϕ_{deadrise}	45.0	deg
ϕ_{flare}	5.0	deg
v_{test}	[0,15]	knots

duces maximum wave-induced VBM, occurring when the wavelength is equal to the ship's length. Thus, a regular sea state is selected and described by Eqs. (30) and (31) (Tupper, 2004).

$$\lambda_{\text{w}} = L \tag{30}$$

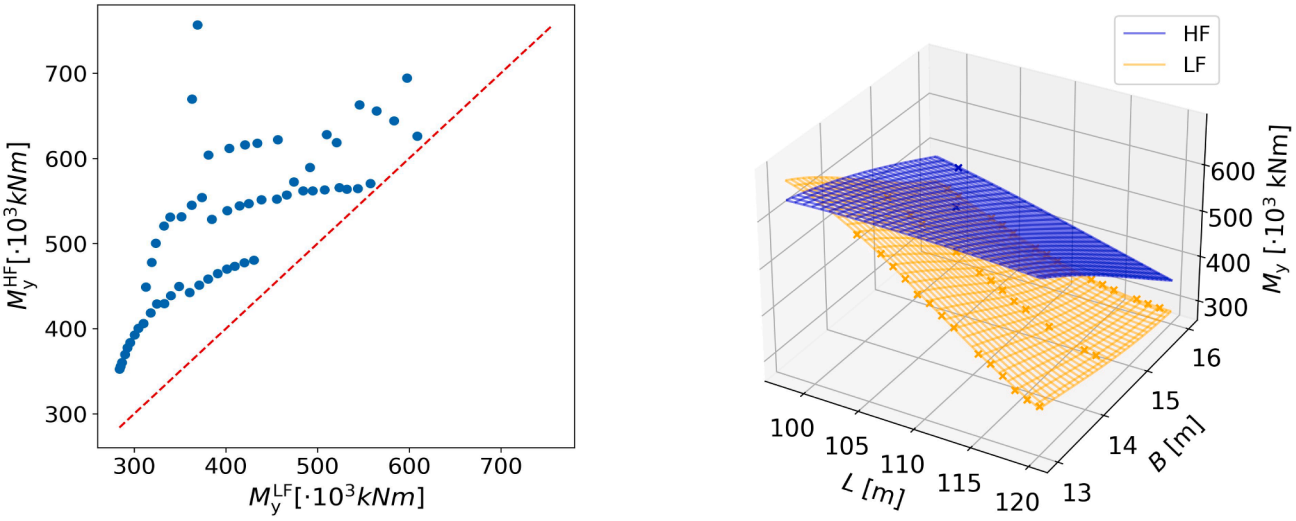
$$H_{\text{w}} = 0.607\sqrt{L} \tag{31}$$

where λ_w is the wavelength, H_w is the wave height, and L is the length between perpendicular of the vessel. In addition, the analysis assumes head seas, thus $\mu = 180^o$. Regarding the vessel's speed, the scenario of sailing at 15 knots was examined.

4.3. Load assessment

This section seeks to provide an overview of the analysis results obtained from both frequency- and time-domain tools. To achieve this, a hull will undergo analysis under various wave-loading conditions. The outcomes will be compared to evaluate the significance of nonlinearities captured by the time-domain tool. Although nonlinear phenomena were not explicitly analyzed in the context of the present study, their effects are inherently embedded in the high-fidelity data used to train the models. Therefore, the frameworks implicitly learn to account for such nonlinearities through the discrepancy between fidelity levels.

To compare the outcomes obtained from the frequency- and time-domain methods, three cases are examined. For the first case, the forward speed v_{test} was set to zero and the wave height H_{w} was set to 1 m, both the linear and non-linear tools were expected to yield similar results. This expectation was confirmed by the data presented in Fig. 8(a). The second case involved a loading case where the forward speed v_{test} is set to zero and the wave height H_{w} is set to 6.65 m (based



(a) Correlation between fidelities

(b) Visualization of the objective landscape

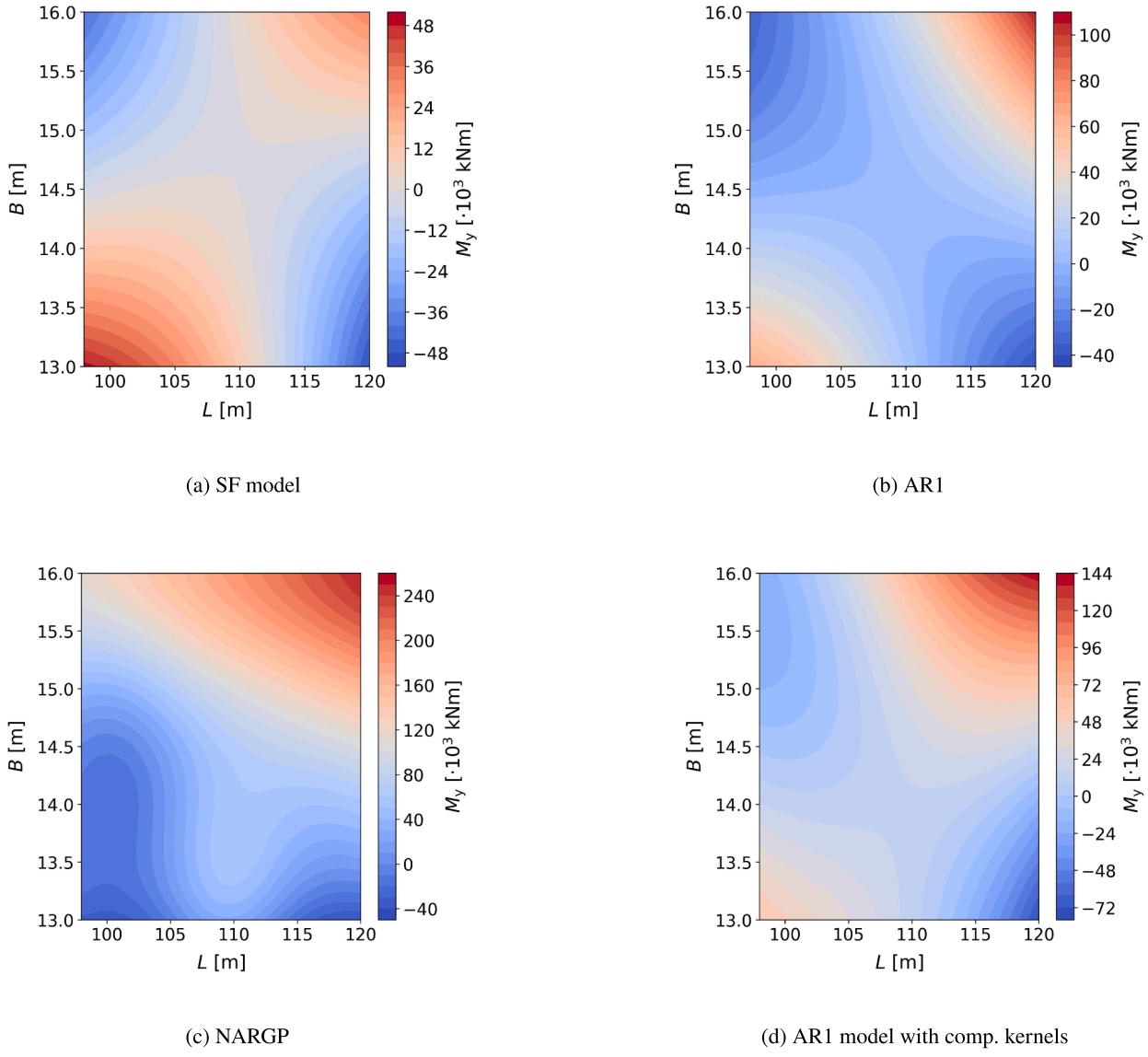
Fig. 11. $2D$ design case, $v = 15$ knots.

Table 6
Error metrics for the $3D$ design case, $u = 0$ knots.

DoE	GP HF R2 (std)	GP HF RMSE (std)	AR1 R2 (std)	AR1 RMSE (std)	NARGP R2 (std)	NARGP RMSE (std)	AR1 with comp kernels R2 (std)	AR1 with comp kernels RMSE (std)
	−0.1153	0.2673	0.3040	0.2072	−0.1034	0.2585	0.4977	0.1701
(5,60)	(0.4519)	(0.0517)	(0.3525)	(0.0507)	(0.7689)	(0.0817)	(0.4637)	(0.0600)
	0.4493	0.1857	0.6038	0.1536	0.5427	0.1656	0.4979	0.1687
(10,60)	(0.2762)	(0.0510)	(0.2893)	(0.0506)	(0.3009)	(0.0569)	(0.5403)	(0.0704)
	0.8458	0.0878	0.8539	0.0944	0.76210	0.1155	0.8064	0.1028
(15,60)	(0.2246)	(0.0551)	(0.0793)	(0.0267)	(0.2337)	(0.0538)	(0.2046)	(0.0441)
	0.9549	0.0501	0.9255	0.0620	0.8800	0.0794	0.9125	0.0685
(20,60)	(0.0373)	(0.0212)	(0.0851)	(0.0275)	(0.1024)	(0.0356)	(0.0854)	(0.0273)
	0.9824	0.0315	0.9708	0.0401	0.9536	0.0498	0.9610	0.0437
(25,60)	(0.0121)	(0.0091)	(0.0234)	(0.0136)	(0.0398)	(0.0218)	(0.0401)	(0.0215)
	0.9899	0.0248	0.9875	0.0276	0.9760	0.0342	0.9829	0.0281
(30,60)	(0.0049)	(0.0063)	(0.0051)	(0.0063)	(0.0335)	(0.0198)	(0.0272)	(0.0155)
	0.9932	0.0202	0.9911	0.02302	0.9944	0.0178	0.9900	0.0209
(35,60)	(0.0033)	(0.0044)	(0.0046)	(0.0049)	(0.0040)	(0.0055)	(0.0151)	(0.0097)
	0.9951	0.0180	0.9938	0.0200	0.9934	0.0196	0.9957	0.0162
(40,60)	(0.0020)	(0.0042)	(0.0030)	(0.0051)	(0.0077)	(0.0105)	(0.0031)	(0.0066)
	0.9961	0.0154	0.9959	0.0159	0.9960	0.0155	0.9969	0.0134
(45,60)	(0.0026)	(0.0047)	(0.0026)	(0.0047)	(0.0036)	(0.0076)	(0.0025)	(0.0048)
	0.9970	0.0135	0.9969	0.0139	0.9973	0.0131	0.9943	0.0154
(50,60)	(0.0016)	(0.0033)	(0.0017)	(0.0035)	(0.0016)	(0.0039)	(0.0136)	(0.0116)

Table 7
Error metrics for the $3D$ design case, $u = 15$ knots.

DoE	GP HF R2 (std)	GP HF RMSE (std)	AR1 R2 (std)	AR1 RMSE (std)	NARGP R2 (std)	NARGP RMSE (std)	AR1 with comp kernels R2 (std)	AR1 with comp kernels RMSE (std)
	−0.2111	0.2533	−0.4961	0.2681	−0.1168	0.2431	−1.4526	0.3431
5,60	(0.4843)	(0.0448)	(1.2012)	(0.1010)	(0.4547)	(0.0457)	(2.0622)	(0.1313)
	0.0517	0.2265	0.5057	0.1649	0.2894	0.1969	−0.1438	0.2256
10,60	(0.4125)	(0.0403)	(0.1412)	(0.0222)	(0.2274)	(0.0330)	(1.3436)	(0.1096)
	0.1782	0.2160	0.5124	0.1625	0.4771	0.1703	0.3493	0.1821
15,60	(0.1427)	(0.0191)	(0.2380)	(0.0389)	(0.1887)	(0.0311)	(0.4976)	(0.0646)
	0.2841	0.1997	0.6606	0.1382	0.5880	0.1511	0.6394	0.1423
20,60	(0.2879)	(0.0377)	(0.1000)	(0.0210)	(0.1711)	(0.0302)	(0.1197)	(0.0237)
	0.3294	0.1957	0.6804	0.1352	0.6295	0.1450	0.5766	0.1498
25,60	(0.1713)	(0.0277)	(0.0749)	(0.0168)	(0.1216)	(0.0241)	(0.3698)	(0.0578)
	0.4286	0.1806	0.7415	0.1224	0.6900	0.1340	0.7261	0.1259
30,60	(0.2308)	(0.0361)	(0.0735)	(0.0170)	(0.0959)	(0.0210)	(0.0823)	(0.0189)
	0.6014	0.1544	0.7375	0.1265	0.6922	0.1367	0.7361	0.1267
35,60	(0.1357)	(0.0228)	(0.0570)	(0.0144)	(0.0841)	(0.0219)	(0.0612)	(0.0155)
	0.5769	0.1547	0.7359	0.1221	0.7202	0.1261	0.7293	0.1232
40,60	(0.1217)	(0.0230)	(0.0810)	(0.0204)	(0.0833)	(0.0223)	(0.1030)	(0.0251)
	0.6985	0.1342	0.7897	0.1111	0.7658	0.1169	0.7732	0.1150
45,60	(0.0817)	(0.0190)	(0.0742)	(0.0183)	(0.0949)	(0.0243)	(0.0912)	(0.0227)
	0.7057	0.1379	0.7895	0.1166	0.7752	0.1200	0.8002	0.1135
50,60	(0.0769)	(0.0268)	(0.0640)	(0.0261)	(0.0778)	(0.0266)	(0.0620)	(0.0249)



(a) SF model

(b) AR1

(c) NARGP

(d) AR1 model with comp. kernels

Fig. 12. Visualization of the models' error for the 15-knot $2D$ case with 6 HF points.

on Eq. (31)). In this case, it was expected that nonlinearities will be more pronounced, resulting in different predicted loads from the linear and nonlinear codes. This expectation was confirmed by the results shown in Fig. 8(b). For the third case, the forward speed v_{test} was set to 15 knots, and the wave height H_{w} was set to 6.65 m (based on Eq. (31)). As shown in Fig. 8(c), the influence of nonlinearities becomes even more pronounced when forward speed is included. This loading case will be used for the design exploration case studies.

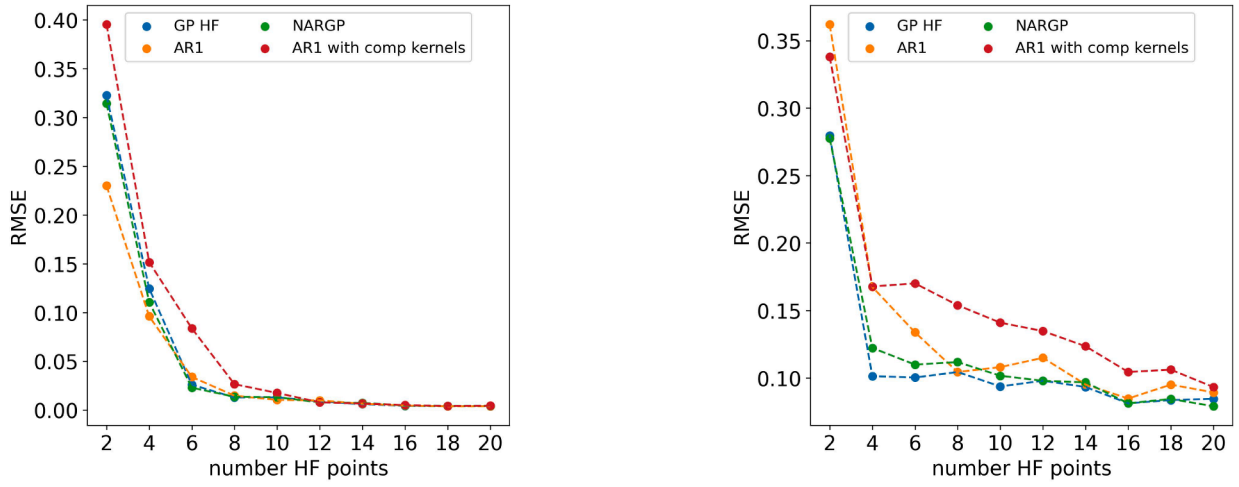
5. Results

In this section, the design scenarios are presented and discussed. The analysis covers three scenarios with increasing problem dimensionality: a $2D$, a $3D$, and an $8D$ problem. The $2D$ case acts as a simplified version of the realistic design problem, but it serves the purpose of understanding and visualizing the design surfaces. In addition, the $3D$ case aims to identify the design trends stemming from the AXE bow. The $8D$ case study includes the full set of design variables as described in Section 4.1. The different designs are assessed based on the maximum sagging value, since, as previously shown in Section 4.3, sagging generally exceeds hogging. Finally, the influence of speed is discussed across the different design scenarios.

5.1. Design scenario A (DS.A)

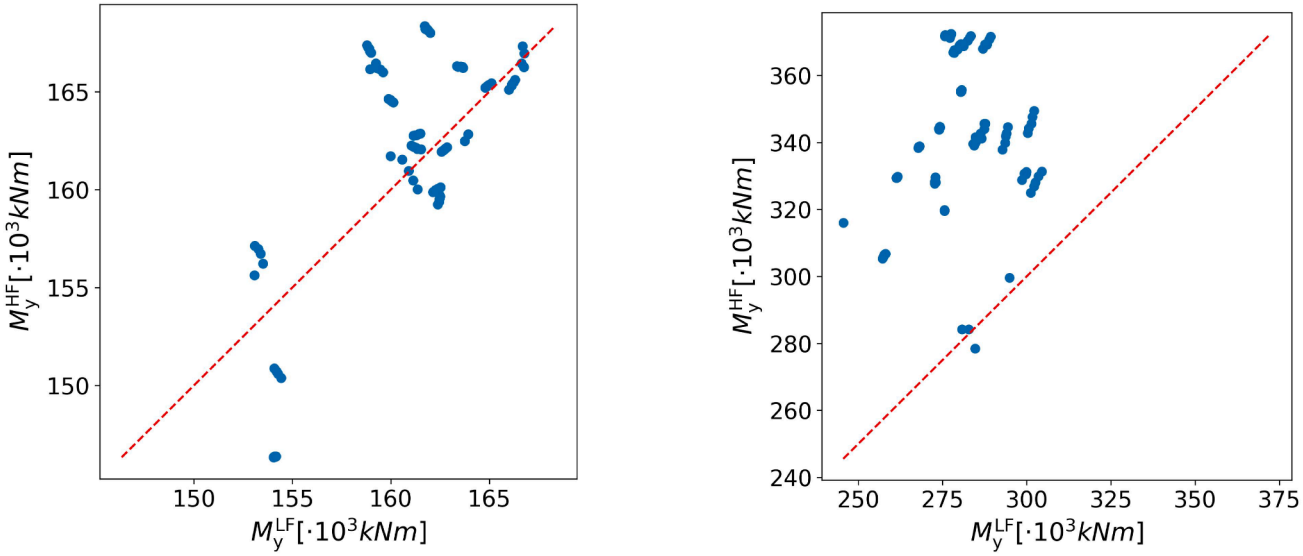
In this $2D$ design scenario, the length L and the beam B of the vessel are taken as the independent variables. The design parameters are detailed in Table 2. To better understand the design problem for $v = 0$ knots, the bifidelity objective landscape is illustrated in Fig. 9(b), while the correlation between the low- and high-fidelity data is depicted in Fig. 9(a). The objective landscape (Fig. 9(b) shows that vessels with greater L are subject to higher loads. Additionally, the loads decrease as the vessel's B increases. As a general trend, as discussed in Section 4.3, the HF TD data predict higher values for the wave-induced sagging compared to the LF FD data. Finally, similar trends are discerned in the LF and HF surfaces of the objective landscape. The complete dataset consists of 64 designs.

For the zero-speed case, the results are presented in Table 3, and the evolution of RMSE with an increasing number of HF points is shown in Fig. 13(a). The methods exhibit larger discrepancies in data regimes



(a) v=0 knots (b) v=15 knots

Fig. 13. Evolution of the RMSE as the number of HF points increases in the $2D$ design case.

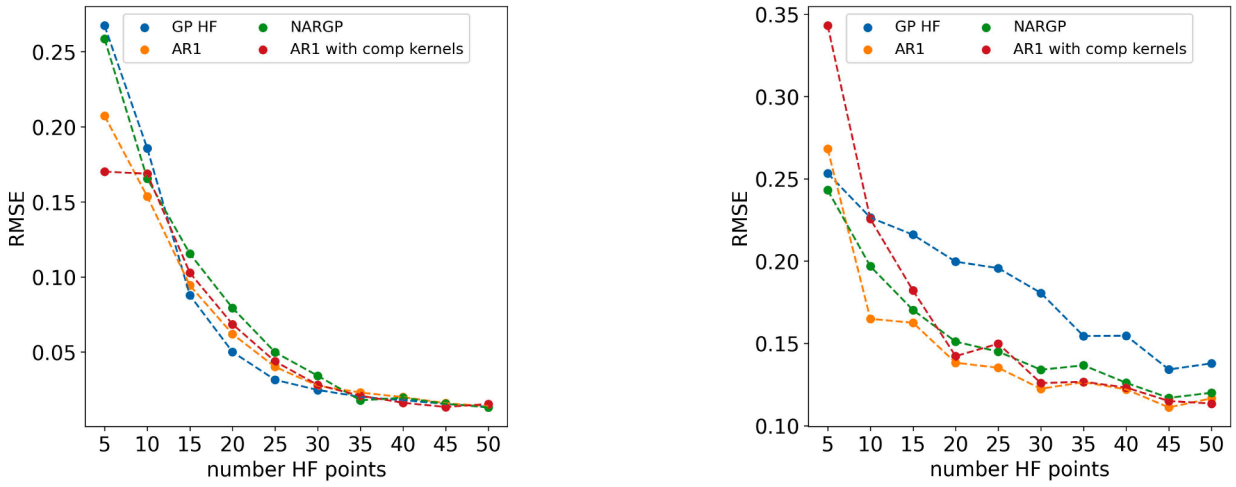


(a) v=0 knots (b) v=15 knots

Fig. 14. Correlation between fidelities $3D$ design case.

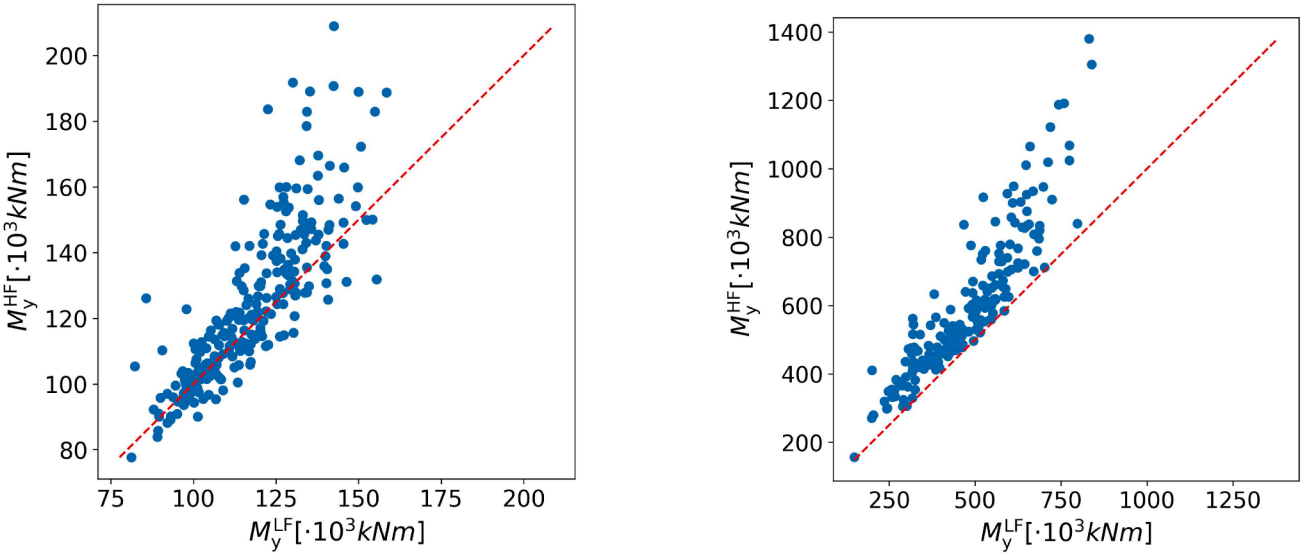
Table 8
Error metrics for the $8D$ design case, $u = 0$ knots.

DoE	GP HF R2 (std)	GP HF RMSE (std)	AR1 R2 (std)	AR1 RMSE (std)	NARGP R2 (std)	NARGP RMSE (std)	AR1 with comp kernels R2 (std)	AR1 with comp kernels RMSE (std)
	0.6896	0.1026	0.7849	0.0851	0.7383	0.0934	0.7527	0.0912
(20,100)	(0.0642)	(0.0121)	(0.0611)	(0.0123)	(0.0973)	(0.0168)	(0.0686)	(0.0135)
	0.7839	0.0878	0.8193	0.0795	0.7932	0.0851	0.7933	0.0844
(30,100)	(0.0386)	(0.0106)	(0.0606)	(0.0135)	(0.0760)	(0.0171)	(0.0853)	(0.0171)
	0.8085	0.0818	0.8470	0.0733	0.8338	0.0764	0.8352	0.0762
(40,100)	(0.0475)	(0.0094)	(0.0332)	(0.0080)	(0.0393)	(0.0085)	(0.0315)	(0.0075)
	0.8386	0.0773	0.8662	0.0705	0.8648	0.0708	0.8569	0.0728
(50,100)	(0.0353)	(0.0099)	(0.0186)	(0.0058)	(0.0260)	(0.0075)	(0.0286)	(0.0086)
	0.8585	0.0718	0.8753	0.0672	0.8697	0.0686	0.8707	0.0683
(60,100)	(0.0263)	(0.0070)	(0.0308)	(0.0075)	(0.0357)	(0.0095)	(0.0343)	(0.0081)
	0.8728	0.0682	0.8836	0.0652	0.8836	0.0652	0.8768	0.0664
(70,100)	(0.0173)	(0.0060)	(0.0159)	(0.0064)	(0.0171)	(0.0070)	(0.0207)	(0.0070)
	0.8796	0.0671	0.8894	0.0644	0.8943	0.0628	0.8885	0.0645
(80,100)	(0.0217)	(0.0073)	(0.0172)	(0.0067)	(0.0224)	(0.0077)	(0.0203)	(0.0069)



(a) v=0 knots (b) v=15 knots

Fig. 15. Evolution of the RMSE as the number of HF points increases in the $3D$ design case.

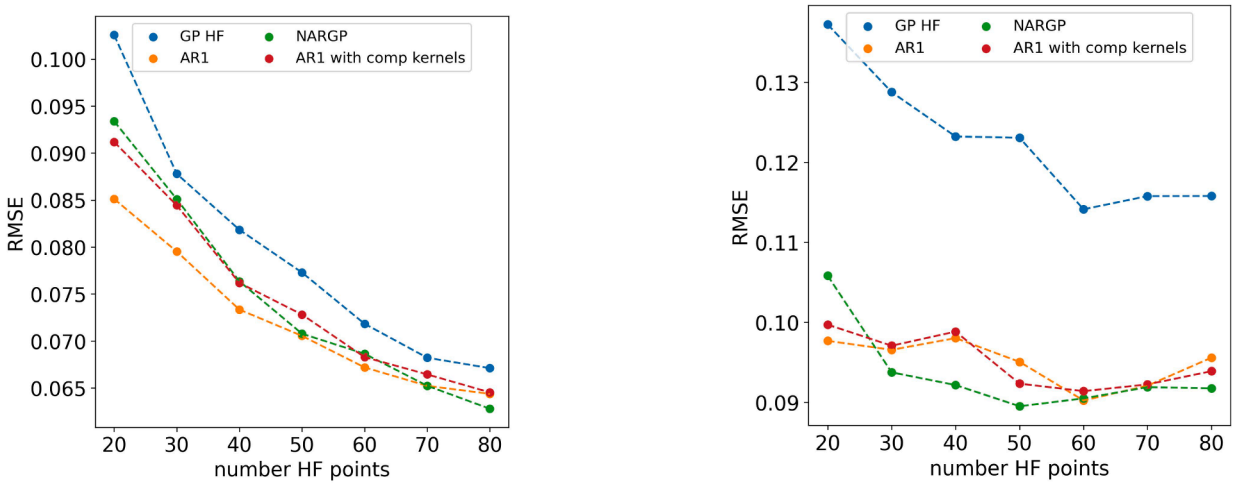


(a) v=0 knots (b) v=15 knots

Fig. 16. Correlation between fidelities $8D$ design case.

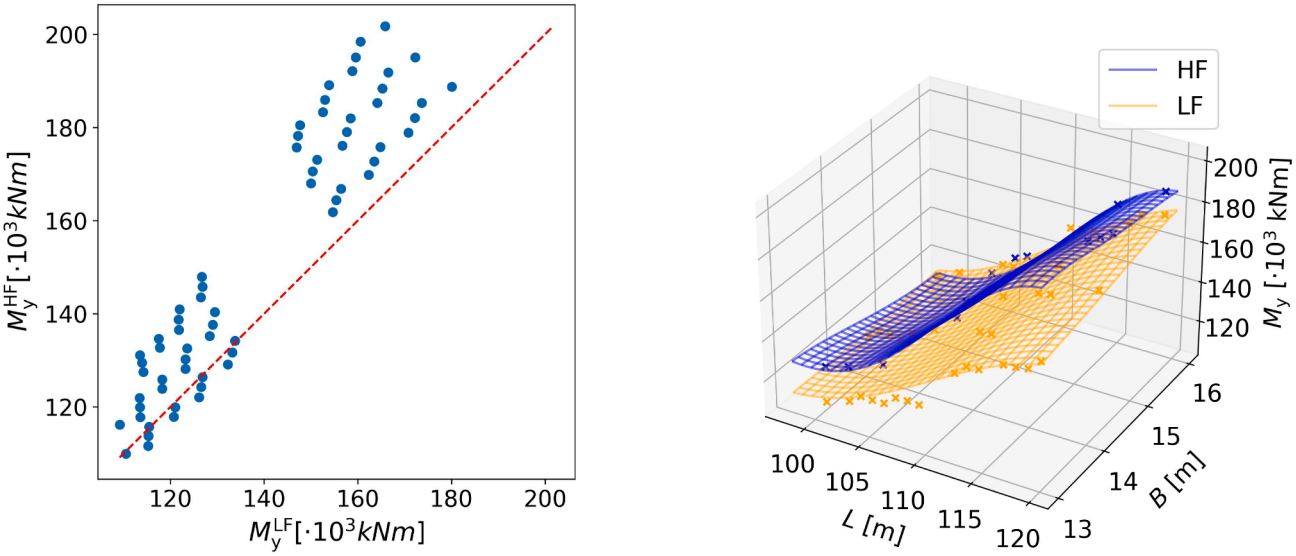
Table 9
Error metrics for the $8D$ design case, $u = 15$ knots.

DoE	GP HF R2 (std)	GP HF RMSE (std)	AR1 R2(std)	AR1 RMSE(std)	NARGP R2(std)	NARGP RMSE (std)	ARl with comp kernels R2 (std)	AR1 with comp kernels RMSE (std)
	0.3265	0.1372	0.6583	0.0977	0.5891	0.1058	0.6423	0.0997
(20,100)	(0.0986)	(0.0101)	(0.0604)	(0.0088)	(0.1529)	(0.0189)	(0.0950)	(0.0143)
	0.4260	0.1288	0.6813	0.0966	0.6967	0.0938	0.6760	0.0971
(30,100)	(0.1585)	(0.0169)	(0.0273)	(0.0052)	(0.0627)	(0.0095)	(0.0504)	(0.0070)
	0.4918	0.1232	0.6788	0.0980	0.7142	0.0922	0.6724	0.0988
(40,100)	(0.0857)	(0.0141)	(0.0322)	(0.0069)	(0.0510)	(0.0082)	(0.0470)	(0.0085)
	0.5028	0.1231	0.7028	0.0951	0.7360	0.0895	0.7195	0.0923
(50,100)	(0.0805)	(0.0139)	(0.0482)	(0.0099)	(0.0495)	(0.0106)	(0.0410)	(0.0088)
	0.5762	0.1141	0.7352	0.0902	0.7341	0.0905	0.7276	0.0914
(60,100)	(0.0485)	(0.0085)	(0.0346)	(0.0080)	(0.0345)	(0.0088)	(0.0384)	(0.0071)
	0.5796	0.1158	0.7340	0.0921	0.7348	0.0919	0.7323	0.0922
(70,100)	(0.0390)	(0.0076)	(0.0351)	(0.0083)	(0.0466)	(0.0108)	(0.0367)	(0.0073)
	0.5651	0.1158	0,7044	0.0956	0.7282	0.0917	0.7151	0.0939
(80,100)	(0.1330)	(0.0214)	(0.0775)	(0.0169)	(0.0717)	(0.0165)	(0.0694)	(0.0162)



(a) v=0 knots (b) v=15 knots

Fig. 17. Evolution of the RMSE as the number of HF points increases in the $8D$ design case.



(a) Correlation between fidelities (b) Visualization of the objective landscape

Fig. 18. $2D$ **synthetic** design case, $v = 0$ knots.

with fewer than 12 HF points. Overall, AR1 performs best in low data regimes, particularly with up to 4 HF points. The performance of the SF and NARGP models is similar, and they converge more quickly than both the AR1 and AR1 with compositional kernels. Regarding the statistics of the error metrics, the AR1 with compositional kernels is outperformed by all other models in this case. An example of a case with 6 HF points is shown in Fig. 10. The contour plots aim to visualize the accuracy of the different models by displaying the error across the design space, defined as the difference between predicted and true values ($y_{\text{predicted}} - y_{\text{true}}$). As demonstrated, the AR1 model's predictions align more closely with the HF objective landscape.

Notably, sailing speed has a significant impact on the wave-induced VBM acting on the hull. The key effect of interest is the altered correlation between the LF and HF analysis data compared to the zero-speed condition, as illustrated in Fig. 11(a). Additionally, the design trends shift, as shown in Fig. 11(b). The clear trend observed at zero speed, where longer vessels experience higher loads Fig. 9(b), no longer holds true in this scenario. The complete dataset consists of 60 designs.

For the 15-knot case, the results are summarized in Table 4 and the evolution of the RMSE is illustrated in Fig. 13(b). It is important to highlight that the SF model, the GP HF, consistently outperforms the others in most of the tested cases. The most likely explanation for this is that, in this case, the LF and HF models do not capture the same trends, as previously mentioned. An example where the SF model more accurately captures the objective landscape compared to other models is shown in Fig. 12. Similar to the zero-knot case shown in Fig. 10, 6 HF points were used to train the different models. However, in this case, the predictions deviate further from the HF objective landscape compared to the zero-knot case. This underscores the fact that altering a design parameter can significantly increase the complexity of the problem, thereby affecting the predictive accuracy of the models.

Table 10
Error metrics for the 2D synthetic case study.

DoE	GP HF R2 (std)	GP HF RMSE (std)	AR1 R2 (std)	AR1 RMSE (std)	NARGP R2 (std)	NARGP RMSE (std)	AR1 with comp. kernels R2(std)	AR1 with comp. kernels RMSE (std)
(2,40)	−0.5220 (0.7963)	0.3665 (0.0937)	0.1216 (2.3334)	0.2136 (0.1893)	−0,4172 (0.8640)	0.3493 (0.1059)	−1.9599 (11.6172)	0.2636 (0.4501)
(4,40)	0.5762 (0.3356)	0.1874 (0.0697)	0.7306 (0.1786)	0.1516 (0.0507)	0.4863 (0.3809)	0.2037 (0.0852)	0.7840 (0.1416)	0.1365 (0.0433)
(6,40)	0.7731 (0.1582)	0.1381 (0.0476)	0.8532 (0.1435)	0.1080 (0.0465)	0.7879 (0.1631)	0.1312 (0.0520)	0.8462 (0.1528)	0.1073 (0.0538)
(8,40)	0.7940 (0.1631)	0.1317 (0.0494)	0.8551 (0.0855)	0.1123 (0.0345)	0.8878 (0.0628)	0.0999 (0.0264)	0.9328 (0.0601)	0.0724 (0.0346)
(10,40)	0.8247 (0.1154)	0.1225 (0.0392)	0.8983 (0.0637)	0.0939 (0.0287)	0.8803 (0.0945)	0.0996 (0.0381)	0.9683 (0.0462)	0.0476 (0.0270)
(12,40)	0.8674 (0.1118)	0.1060 (0.0413)	0.9065 (0.0646)	0.0908 (0.0286)	0.9175 (0.0646)	0.0831 (0.0336)	0.9789 (0.0163)	0.0426 (0.0157)
(14,40)	0.9026 (0.0617)	0.0920 (0.0271)	0.9322 (0.0412)	0.0769 (0.0215)	0.9467 (0.0285)	0.0686 (0.0180)	0.9770 (0.0305)	0.0414 (0.0198)
(16,40)	0.9324 (0.0158)	0.0808 (0.0087)	0.9484 (0.0191)	0.0701 (0.0115)	0.9490 (0.0291)	0.0675 (0.0202)	0.9738 (0.0354)	0.0430 (0.0265)
(18,40)	0.9336 (0.0257)	0.0790 (0.0128)	0.9465 (0.0185)	0.0710 (0.0120)	0.9581 (0.0250)	0.0603 (0.0181)	0.9841 (0.0145)	0.0361 (0.0153)
(20,40)	0.9312 (0.0202)	0.0807 (0.0118)	0.9531 (0.0159)	0,0665 (0.0098)	0.9532 (0.0221)	0.0653 (0.0157)	0.9875 (0.0103)	0.0321 (0.0135)

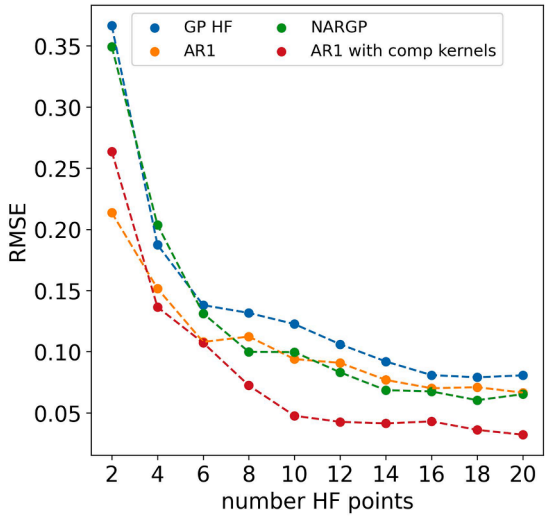


Fig. 19. Evolution of the RMSE as the number of HF points increases in the 2D synthetic design case.

5.2. Design scenario B (DS.B)

This design scenario examines the variation of three independent variables related to the design of the AXE bow, L_{fore}, D_{axe}, and H_{axe}. The design parameters are listed in Table 5. This scenario aims to investigate how the bow design affects the wave-induced VBM experienced by the vessel. The correlation between the HF and the LF data can be seen in Fig. 14 for both the 0 and 15 knots cases. It is evident that the correlation is not strong, unlike in design scenario D.S.A. However, similar to D.S.A, the 15-knot case shows greater complexity compared to the zero-speed scenario.

For the zero-speed case, the results are summarized in Table 6 and visualized in Fig. 15(a). In the case of 5 HF points, the AR1 model with compositional kernels performs the best. For the range of 10 to 35 HF points, the SF model delivers the best results, while the AR1 and AR1 with compositional kernels perform similarly, with AR1 slightly outperforming the other. The NARGP model shows the weakest performance. For HF points exceeding 35, all models converge to a similarly low error value.

For the 15-knot case, the results are summarized in Table 7 and illustrated in Fig. 15(b). Similarly to the previous cases, no single model consistently outperforms the others across all scenarios. The NARGP model performs best with 5 HF points. Unlike the zero-speed case, the SF model has the weakest overall performance. Both the AR1 and AR1 with compositional kernels converge to similar low error values when the number of HF points exceeds 30. However, when the HF points are fewer than 30, the AR1 model demonstrates superior performance.

5.3. Design scenario C (DS.C)

This design scenario explores the variation of eight independent variables related to the hull design (length L, beam B, depth D, deadrise angle ϕ_{deadrise}, and flare angle ϕ_{flare}), as well as the AXE bow design parameters (L_{fore}, D_{axe}, and H_{axe}). The only design parameter considered is the speed, $u \in [0, 15]$ knots. This scenario aims to assess the scalability and predictive accuracy of the evaluated methods in a realistic early-stage ship design problem. The correlation between the HF and the LF data can be seen in Fig. 16 for both the 0 and 15 knots cases.

For the zero-speed case, the results are summarized in Table 8 and illustrated in Fig. 17(a). It is worth noting that the SF model consistently exhibits the weakest performance across all tested cases. In the range of 20 to 60 HF points, the AR1 model shows the best performance, while beyond 60 HF points, the NARGP model delivers more accurate predictions. The performance of the AR1 model with compositional kernels is comparable to that of the NARGP model.

For the 15-knot, the results are summarized in Table 9 and visualized in Fig. 17(b). In this case, similar to the zero-speed case, the GP HF model is consistently outperformed by all three MF models across all tested cases. It is noteworthy that the predictions from the GP HF model exhibit significantly larger errors compared to those from the MF models. The NARGP model demonstrates superior accuracy across most of the tested cases. The AR1 model and the AR1 with compositional kernels perform similarly, with AR1 slightly outperforming the latter.

6. Reflection on the performance of the AR1 with the compositional kernels

In the case studies presented in this paper, the AR1 with compositional kernels model did not yield improved predictions compared to its base model, AR1. In contrast, the case studies from Charisi et al. (2024b) demonstrated that the proposed model can significantly enhance predictions over AR1. This section will further analyze and provide insights into the performance of this model.

The primary motivation for integrating compositional kernels was to identify and leverage the underlying patterns governing the objective landscape. The analytical case studies presented in Charisi et al. (2024b) were designed to model distinct patterns tied to specific attributes of the objective landscape, such as multi-modality. Thus, such case studies are

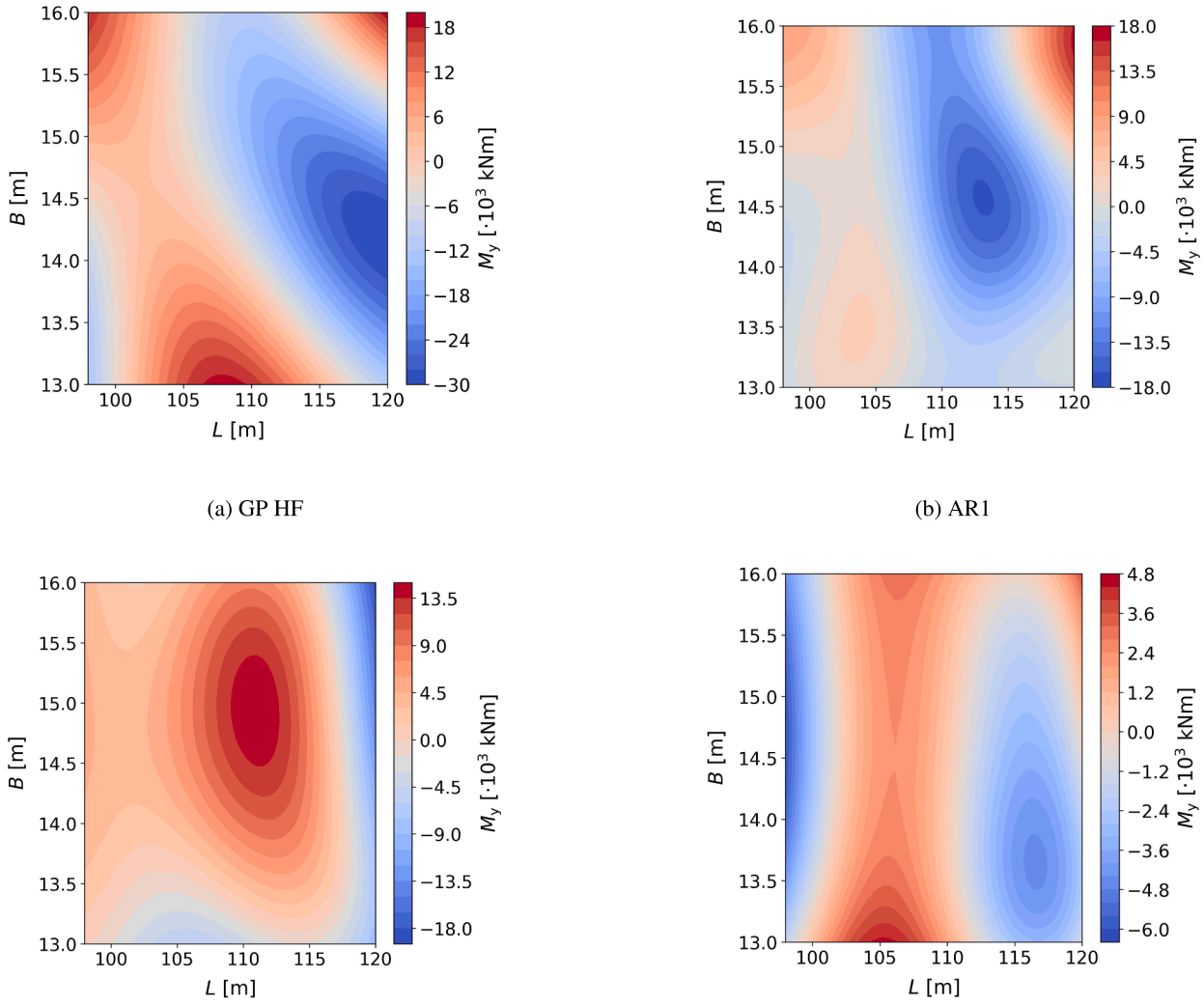


Fig. 20. Visualization of the models' error for the 0-knot $2D$ synthetic case with 10 HF points.

ideal to showcase the performance of the proposed model. Therefore, these case studies serve as the perfect testing ground for demonstrating the effectiveness of the AR1 with the compositional kernels model. A similar behavior was observed in a cantilever beam case. However, this observation does not hold true for the case studies of this paper.

Thus, a synthetic case study has been introduced to test the hypothesis that the proposed model can improve predictions in a objective landscape with distinct features. In this case, a discontinuity was intentionally added to the $2D$ design case of the previously examined AXE frigates. More specifically, the original VBM datasets were modified according to Eqs. (32) and (33). Fig. 18(a), illustrates the correlation between the fidelities, while Fig. 18(b) shows the resulting discontinuous objective landscape.

$$\tilde{M}_y^{LF} = \begin{cases} M_y^{LF}, & 98 \leq L < 110 \\ M_y^{LF} + 15000, & 110 \leq L < 120 \end{cases} \tag{32}$$

$$\tilde{M}_y^{HF} = \begin{cases} M_y^{HF}, & 98 \leq L < 110 \\ M_y^{HF} + 25000, & 110 \leq L < 120 \end{cases} \tag{33}$$

The results of the synthetic case study are summarized in Table 10, and Fig. 19 illustrates the evolution of the RMSE with the number of HF points. The proposed model consistently outperforms the other models in nearly all tested cases, converging to a lower RMSE value. In Fig. 20, shows an example where the training set consists of 10 HF and 40 LF points. The proposed model effectively captures the objective landscape's discontinuity, unlike the other models. Regarding the performance of the other models, the MF models, including AR1 and NARGP, generally outperform the SF model.

6.1. Computational efficiency

Computational efficiency is an important consideration in surrogate modeling. A well-known limitation of GPs is their cubic training complexity, $O(n^3)$, with respect to the number of data points n (Liu et al., 2019). This makes them less suitable for large datasets or high-dimensional design spaces. However, early-stage ship design problems typically involve moderate data sizes, where the dominant computational cost lies in the HF load analysis rather than the model training itself. Therefore, the authors argue that the key consideration is select-

ing a method capable of accurately predicting the objective landscape using as little HF data as possible.

Indicatively, the mean time to build the SF model is 0.18 s, compared to 5.13 s for the MF linear model in the $2D$ design scenario using 2 HF points. This increases to 0.48 s for the SF model and 5.96 s for the MF model in the $8D$ design scenario with 80 HF points. However, the development of the AR1 model with compositional kernels is significantly more computationally demanding. As discussed in (Charisi et al., 2024b), this method involves a higher computational cost; therefore, the problem's dimensionality and the number of fidelities become critical factors for the applicability of the method. In the present case study, the time required to find the optimal compositional kernel function ranges from 966.04 s in the $2D$ scenario with 2 HF points to 1650.15 s in the $8D$ case with 80 HF points. The reported computational times were measured on a system with an Intel(R) Xeon(R) E5-1620 v3 processor (3.50 GHz, 4 cores) and 40 GB of RAM. No GPU acceleration was used.

7. Conclusions and recommendations for future research

7.1. Conclusions

The paper explored the potential of harnessing MF models for early-stage predictions of wave-induced loads. The case study focused on the early design assessment of wave-induced VBM for the AXE frigates. The bi-fidelity setup was developed by combining frequency- and time-domain analysis data. The predictive accuracy of four models was evaluated: a model based on SF-GPs, the AR1 scheme with the squared exponential kernel, the AR1 with the integration of compositional kernels and NARGP. Three design scenarios were tested, involving a $2D$ variation of length and beam, a $3D$ variation of AXE parameters, and an 8D variation of the main hull parameters.

An important observation is that, in most cases, the MF models demonstrated superior accuracy compared to the SF model, which was based solely on HF data. Based on the evidence from the case study, MF models have indeed proven capable of enabling the early incorporation of HF analysis into the design process. There was only one instance where the SF model outperformed the MF models, specifically in scenario DS.A at a speed of 15 knots. In this case, the HF and LF models captured slightly different trends, which likely explains why the SF model proved to be more accurate.

Another key observation is that changes in design parameters, such as speed (u), can significantly impact the shape of the objective landscape, the complexity of the problem, and the correlation between LF and HF analysis data, ultimately affecting the predictive accuracy of the models. As observed across all the presented cases, the zero-speed variation consistently converged to lower predictive error values with fewer HF analysis data compared to the 15-knot variation.

Even within the same design problem of assessing the wave-induced VBM for the AXE frigates, different models outperformed others depending on the specific variations. In scenario DS.A ($u = 0$ knots), the AR1 model tends to outperform the other MF models, providing a better representation of the smooth objective landscape corresponding to the $2D$ design space. Additionally, as the dimensionality and complexity of the problem increase, the SF models begin to perform significantly worse compared to the MF models, particularly in scenarios such as DS.B ($u = 15$ knots) and DS.C. Additionally, the linear AR1 scheme outperforms the nonlinear NARGP scheme in the lower data regime; however, NARGP ultimately converges to a lower error value in the $8D$ problem. Finally, the integration of compositional kernels does not enhance the predictive accuracy of the AR1 scheme. Instead, for the examined case study, the squared exponential kernel appears to be more effective in encoding the covariance matrix of the design space data in most cases.

Further investigation into the AR1 model with compositional kernels was carried out to explore the contradictory results. Its scalability was demonstrated through its successful application to a realistic ship design problem. The AR1 model with compositional kernels was designed to identify distinct features within the objective landscape. However, the case study presented did not exhibit such characteristics. To test this hypothesis, a discontinuity was intentionally introduced into the objective landscape, similar to the Jump Forrester analytical benchmark. The proposed model successfully captured this distinct feature and outperformed the other models.

Regarding the scalability of the presented MF models, GP-based schemes face increased computational costs as dimensionality grows; however, the tested range in this study remains relatively small. Therefore, the increase in dimensionality is not a limitation in this specific problem for the multi-fidelity schemes. However, as dimensionality increases, the amount of required training data also inevitably grows. Each training point represents one design variation. In this case study, computational time for the VBM analysis remained reasonable since only a single wave condition was tested. However, if multiple irregular sea states were assessed for each design variation, the increase in training data would become a significant limitation.

7.2. Reflection on the ship design problem

Overall, the integration of MF models into an early-stage design framework can be highly effective for certain design cases. As demonstrated in the case studies, the MF models outperformed the SF model in most design scenarios. However, the main challenge lies in the fact that the effectiveness of different MF models depends heavily on the characteristics of the design problem. As demonstrated in the presented case study, even within the same physical problem, altering design parameters or varying the amount of HF data can significantly impact the performance of the various MF models. Some general trends can be extracted from the current research findings (including the case studies results of this paper); however, the primary challenge lies in the fact that, in practical applications, the features of the objective landscape are often unknown, making the selection of an appropriate MF model particularly challenging.

Another important consideration is that the ultimate goal is to extend the proposed approach to calculate the design load by accounting for the various sea states a vessel may encounter throughout its lifetime. As a first step, however, the current study focuses on vessel response in a regular sea state, specifically a critical condition for the VBM problem, where the vessel's length matches the wavelength. Scaling the framework to irregular sea states introduces several key challenges that must be addressed to ensure its applicability in realistic design problems. The authors argue that three key aspects must be considered when extending the current framework to operational sea spectra:

- Limited availability of analysis data: Evaluating each design across a wide range of sea states drastically increases computational cost, reducing the number of feasible design variations that can be simulated. This limitation has been anticipated in the present study, where models are developed and assessed under low-data regimes.
- Complexity of the objective landscape: Introducing spectral variability is expected to reshape the objective landscape, potentially introducing stronger nonlinearities or multimodal features. This added complexity may affect the predictive performance of surrogate models and should be explicitly addressed in future work.
- Scalability of the modeling framework: In a real sea spectrum setting, each design must be evaluated across multiple sea states, making the overall computational cost prohibitive. While prior research (e.g., Guth et al., 2022) has considered MF modeling for predicting responses across different sea conditions, the authors propose that a combined approach–employing MF surrogates across both the design and loading condition spaces–offers a promising direction to scale the framework effectively.

7.3. Recommendations for future research

To further advance the development of the DAF, it would be beneficial to integrate a Bayesian optimization approach. This would enable the assessment of active search techniques in the development of the surrogate models for the objective landscape by the various MF models.

Additionally, evaluating and comparing these GP-based models in other early-stage design analysis problems, such as resistance and seakeeping, will be beneficial. This would help provide a comprehensive understanding of how MF models can contribute to early-stage design optimization and exploration. Similarly, integrating compositional kernels into the AR1 scheme has proven beneficial for other design problems. However, in this research, the MG-GP model with compositional kernels outperformed the others only in the synthetic case study. Therefore, further testing on different design problems is recommended to assess the effectiveness of this approach in real naval architecture applications.

Furthermore, increasing the case study's complexity by incorporating higher vessel speeds would be valuable. Additionally, it would be valuable to explore how this approach could be extended to account for nonlinear loads, such as green water and slamming. In such analyses, each design variation needs to be tested in a large amount of wave conditions to accurately assess the probabilities of occurrence or failure. As a result, the computational cost can become prohibitive, making it crucial to train our MF models with fewer design variations to reduce the computational burden.

From an application perspective, several directions can be explored. First, this research did not compare different analysis solvers for calculating the wave-induced VBM in terms of accuracy and computational cost. Investigating this could enhance the accuracy of the analysis data used for training without significantly increasing computational expenses. Since higher-quality training data ultimately leads to more accurate surrogate models for predicting the objective landscape, this could significantly improve overall model performance. Additionally, future research should focus on increasing the fidelity of the HF tool to CFD, further enhancing the accuracy of the analysis.

CRediT authorship contribution statement

Nikoleta Dimitra Charisi: Writing – original draft, Methodology, Conceptualization; **Emile Defer:** Writing – review & editing, Conceptualization; **Hans Hopman:** Writing – review & editing, Supervision; **Austin A. Kana:** Writing – review & editing, Supervision.

Data availability

The data used for this research paper can be found in Charisi et al. (2024a).

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Nikoleta Dimitra Charisi reports financial support was provided by Dutch Research Council. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This publication is part of the project "Multi-fidelity Probabilistic Design Framework for Complex Marine Structures" (project number TWM.BL.019.007) of the research program "Topsector Water & Maritime: the Blue route" which is (partly) financed by the Dutch Research Council (NWO). The authors thank DAMEN, the Netherlands Commit (Commando Materieel en IT), and MARIN for their contribution to the research.

References

Andrews, D., 2018. The sophistication of early stage design for complex vessels. Int. J. Marit. Eng. Vol 160 (SE 18). https://doi.org/10.3940/RINA.IJME.2018.SE.472

Associates, R. M., 2023a. Grasshopper (version 1.0). Accessed: 2023-07-02. https://www.grasshopper3d.com.

Associates, R. M., 2023b. Rhinoceros (version 7.0). Accessed: 2023-07-02. https://www.rhino3d.com.

Bachoc, F., 2013. Cross validation and maximum likelihood estimations of hyper-parameters of Gaussian processes with model misspecification. Comput. Stat. Data Anal. 66, 55–69. https://www.sciencedirect.com/science/article/pii/S0167947313001187. https://doi.org/https://doi.org/10.1016/j.csda.2013.03.016

Bennett, S.S., Hudson, D.A., Temarel, P., 2013. The influence of forward speed on ship motions in abnormal waves: experimental measurements and numerical predictions. J. Fluids Struct. 39, 154–172. https://www.sciencedirect.com/science/article/pii/S088997461300025X. https://doi.org/https://doi.org/10.1016/j.jfluidstructs.2013.01.006

Bonfiglio, L., Perdikaris, P., Brizzolara, S., 2020. Multi-fidelity Bayesian optimization of SWATH hull forms. J. Ship. Res. 64 (02), 154–170. https://onepetro.org/JSR/article-pdf/64/02/154/2257847/sname-jsr-2020-64-2-154.pdf. https://doi.org/10.5957/jsr.2020.64.2.154. https://doi.org/10.5957/jsr.2020.64.2.154

Bonfiglio, L., Perdikaris, P., Vernengo, G., de Medeiros, J.S., Karniadakis, G., 2018a. Improving SWATH seakeeping performance using multi-fidelity Gaussian process and Bayesian optimization. J. Ship. Res. 62 (04), 223–240. https://onepetro.org/JSR/article-pdf/62/04/223/2232316/sname-jsr-2018-62-4-223.pdf. https://doi.org/10.5957/JOSR.11170069. https://doi.org/10.5957/JOSR.11170069

Bonfiglio, L., Perdikaris, P., del Águila, J., Karniadakis, G.E., 2018b. A probabilistic framework for multidisciplinary design: application to the hydrostructural optimization of supercavitating hydrofoils. Int. J. Numer. Methods Eng. 116 (4), 246–269. https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.5923. https://doi.org/https://doi.org/10.1002/nme.5923

Bouscasse, B., Merrien, A., Horel, B., De Hauteclocque, G., 2022. Experimental analysis of wave-induced vertical bending moment in steep regular waves. J. Fluids Struct. 111, 103547. https://www.sciencedirect.com/science/article/pii/S0889974622000287. https://doi.org/https://doi.org/10.1016/j.jfluidstructs.2022.103547

Brevault, L., Balesdent, M., Hebbal, A., 2020. Overview of Gaussian process based multi-fidelity techniques with variable relationship between fidelities, application to aerospace systems. Aerosp. Sci. Technol. 107, 106339. https://www.sciencedirect.com/science/article/pii/S127096382031021X. https://doi.org/https://doi.org/10.1016/j.ast.2020.106339

Bryson, D.E., Beran, P.S., Thelen, A.S., Rumpfkeil, M.P., Nikbay, M., Çakmak, E., Yildiz, S., 2022. Comparison of multi-fidelity optimization methods using an aero-structural benchmark problem. In: AVT-354 Research Workshop on "Multi-Fidelity Methods for Military Vehicle Design". Varna.

Cameron, A.C., Windmeijer, F. A.G., 1997. An R-squared measure of goodness of fit for some common nonlinear regression models. J. Econom. 77 (2), 329–342. https://doi.org/10.1016/S0304-4076(96)01818-0

Charisi, N.D., Hopman, H., Kana, A., 2024a. Data underlying chapter 5 of the PhD dissertation: Multi-fidelity probabilistic design framework for early-stage design of novel vessels. Dataset. https://doi.org/10.4121/2cc2e2a8-8d04-49ca-9f68-8c1363ec03c3.v2. https://doi.org/10.4121/2cc2e2a8-8d04-49ca-9f68-8c1363ec03c3.v2

Charisi, N.D., Hopman, J.J., Kana, A., 2024b. Multi-fidelity design framework integrating compositional kernels to facilitate early-stage design exploration of complex systems. J. Mech. Des. , 1–22https://asmedigitalcollection.asme.org/mechanicaldesign/article-pdf/doi/10.1115/1.4065890/7352094/md-24-1071.pdf. https://doi.org/10.1115/1.4065890

Charisi, N.D., Kana, A., Hopman, H., 2022. Compositional kernels to facilitate multi-fidelity design analysis: applications for early-stage design. In: AVT-354 Multi-Fidelity Methods for Military Vehicle Design.

Drummen, I., Hageman, R.B., Stambaugh, K., 2022. Multifidelity approach for predicting extreme global bending load effects. In: 9th International Conference on HYDROELASTICITY in MARINE TECHNOLOGY. Rome.

Duvenaud, D., Lloyd, J., Grosse, R., Tenenbaum, J., Ghahramani, Z., 2013. Proceedings of the 30th international conference on machine learning. In: ICML. PMLR, pp. 1166–1174.

Eefsen, T., Van Walree, F., Peri, D., Van Terwisga, P., Kristensen, H.O., Dattola, R., Visser, M., 2004. Development of frigate designs with good seakeeping characteristics. In: 9th Symposium on Practical Design of Ships and Other Floating Structures. Luebeck-Travemuende.

Fernández-Godino, M.G., 2023. Review of multi-fidelity models. Adv. Comput. Sci. Eng. 1 (4), 351–400. https://www.aimsciences.org/article/id/656460fbeca2737fc4ef5257. https://doi.org/10.3934/acse.2023015

Filipa Marques Sanches, 2016. Parametric modelling of hull form for ship optimization. https://fenix.tecnico.ulisboa.pt/downloadFile/281870113703278/Thesis.pdf.

Fonseca, N., Soares, C.G., 2002. Comparison of numerical and experimental results of nonlinear wave-induced vertical ship motions and loads. J. Mar. Sci. Technol. 6, 193–204.

Gaggero, S., Vernengo, G., Villa, D., 2022. A marine propeller design method based on two-fidelity data levels. Appl. Ocean Res. 123, 103156. https://www.sciencedirect.com/science/article/pii/S0141118722001031. https://doi.org/https://doi.org/10.1016/j.apor.2022.103156

GPy, 2012. GPy: a Gaussian process framework in python. http://github.com/SheffieldML/GPy.

Gratiet, L.L., Garnier, J., 2014. Recursive co-kriging model for design of computer experiments with multiple levels of fidelity. Int. J. Uncertain. Quantif. 4 (5), 365–386. https://doi.org/10.1615/Int.J.UncertaintyQuantification.2014006914

Guth, S., Champenois, B., Sapsis, T.P., 2022. Application of Gaussian process multi-fidelity optimal sampling to ship structural modeling. In: 34th Symposium on Naval Hydrodynamics. Washington, DC, USA.

Hirdaris, S.E., Bai, W., Dessi, D., Ergin, A., Gu, X., Hermundstad, O.A., Huijsmans, R., Iijima, K., Nielsen, U.D., Parunov, J., Fonseca, N., Papanikolaou, A., Argyriadis, K., Incecik, A., 2014. Loads for use in the design of ships and offshore structures. Ocean Eng. 78, 131–174. https://www.sciencedirect.com/science/article/pii/S0029801813003557. https://doi.org/https://doi.org/10.1016/j.oceaneng.2013.09.012

Hirdaris, S.E., Lee, Y., Mortola, G., Incecik, A., Turan, O., Hong, S.Y., Kim, B.W., Kim, K.H., Bennett, S., Miao, S.H., Temarel, P., 2016. The influence of nonlinearities on the symmetric hydrodynamic response of a 10,000 TEU container ship. Ocean Eng. 111, 166–178. https://www.sciencedirect.com/science/article/pii/S002980181500596X. https://doi.org/https://doi.org/10.1016/j.oceaneng.2015.10.049

ITTC, 2014. ITTC's 2014 seakeeping committee report and recommendations. In: 26th ITTC. Copenhagen.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2014. An Introduction to Statistical Learning: with Applications in R. Springer Publishing Company, Incorporated.

Jiang, X., 2015. What happened to MOL comfort? SWZ Marit. , 13–16.

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. http://www.jstor.org/stable/2673557.

Keuning, J.A., Pinkster, J., van Wairee, F., 2002. Further investigation into the hydrodynamic performance of the AXE bow concept. In: Proceedings of the WEMT-HSMV 2002 Conference, 6th Symposium on High Speed Marine Vehicles. Castello di Baia, Italy.

Keuning, J.A., Van Wiilree, F., 2006. The comparison of the hydrodynamic behaviour of three fast patrol boats with special hull geometries. In: 5th International Conference on High-Performance Marine Vehicles (HIPER). Launceston. https://repository.tudelft.nl/islandora/object/uuid%3Af394ac33-d748-46ce-9db6-d24712b7d2a9.

La Rocca, G., 2012. Knowledge based engineering: between AI and CAD. review of a language based technology to support engineering design. Adv. Eng. Inf. 26 (2), 159–179. Knowledge based engineering to support complex product design. https://www.sciencedirect.com/science/article/pii/S1474034612000092. https://doi.org/https://doi.org/10.1016/j.aei.2012.02.002

Ley, J., el Moctar, O., 2021. A comparative study of computational methods for wave-induced motions and loads. J. Mar. Sci. Eng. 9 (1). https://www.mdpi.com/2077-1312/9/1/83. https://doi.org/10.3390/jmse9010083

Li, M., Boulougouris, E., Lazakis, I., Theotokatos, G., 2016. Analysis of the wave-induced vertical bending moment and comparison with the class imposed design loads for 4250 TEU container ship. In: International Conference on Maritime Safety and Operations. Glasgow, pp. 137–143.

Liu, H., Ong, Y.-S., Shen, X., Cai, J., 2019. When Gaussian process meets big data: a review of scalable GPs. IEEE Trans. Neural Netw. Learn. Syst. 31 (11), 4405–4423. https://doi.org/10.1109/TNNLS.2019.2957109

MARIN, 2019. PRECAL_R, Version 18.1.3 Theory Manual. Technical report. MARIN.

Mavris, D., DeLaurentis, D., Bandte, O., Hale, M., 1998. A stochastic approach to multidisciplinary aircraft analysis and design. In: 36th AIAA Aerospace Sciences Meeting and Exhibit. Reno,NV,U.S.A. https://arc.aiaa.org/doi/abs/10.2514/6.1998-912. https://doi.org/10.2514/6.1998-912

Molland, A.F., 2008. Chapter 4 - ship structures. In: The Maritime Engineering Reference Book. Butterworth-Heinemann, Oxford, pp. 116–180. https://www.sciencedirect.com/science/article/pii/B9780750689878000044. https://doi.org/https://doi.org/10.1016/B978-0-7506-8987-8.00004-4

Murray, J.M., 1947. Longitudinal bending moments. [Trans. Inst. Eng. Shipbuilders Scotland 90.

Nitzler, J., Biehler, J., Fehn, N., Koutsourelakis, P.-S., Wall, W.A., 2022. A generalized probabilistic learning approach for multi-fidelity uncertainty quantification in complex physical simulations. Comput. Methods Appl. Mech. Eng. 400, 115600. https://www.sciencedirect.com/science/article/pii/S0045782522005643. https://doi.org/https://doi.org/10.1016/j.cma.2022.115600

Paleyes, A., Mahsereci, M., Lawrence, N.D., 2023. Emukit: a python toolkit for decision making under uncertainty. In: The Python in Science Conference (SciPy).

Paleyes, A., Pullin, M., Mahsereci, M., McCollum, C., Lawrence, N., González, J., 2019. Emulation of physical processes with Emukit. In: Second Workshop on Machine Learning and the Physical Sciences, NeurIPS.

Papanikolaou, A., 2014. Ship Design: Methodologies of Preliminary Design. Springer Netherlands. https://doi.org/10.1007/978-94-017-8751-2

Parunov, J., Guedes Soares, C., Hirdaris, S., Iijima, K., Wang, X., Brizzolara, S., Qiu, W., Mikulić, A., Wang, S., Abdelwahab, H.S., 2022. Benchmark study of global linear wave loads on a container ship with forward speed. Mar. struct. 84, 103162. https://www.sciencedirect.com/science/article/pii/S0951833922000053. https://doi.org/https://doi.org/10.1016/j.marstruc.2022.103162

Peherstorfer, B., Willcox, K., Gunzburger, M., 2018. Survey of multifidelity methods in uncertainty propagation, inference, and optimization. SIAM Rev. 60 (3), 550–591. https://doi.org/10.1137/16M1082469

Perdikaris, P., Raissi, M., Damianou, A., Lawrence, N.D., Karniadakis, G.E., 2017. Nonlinear information fusion algorithms for data-efficient multi-fidelity modelling. Proc. R. Soc. A Math. Phys. Eng. Sci. 473 (2198) 20160751.

Rajendran, S., Fonseca, N., Soares, C.G., 2016. Prediction of extreme motions and vertical bending moments on a cruise ship and comparison with experimental data. Ocean Eng. 127, 368–386. https://www.sciencedirect.com/science/article/pii/S0029801816304528. https://doi.org/https://doi.org/10.1016/j.oceaneng.2016.10.021

Rasmussen, C.E., Williams, C. K.I., 2005. Gaussian Processes for Machine Learning. The MIT Press. https://doi.org/10.7551/mitpress/3206.001.0001. https://doi.org/10.7551/mitpress/3206.001.0001

Register, L., 2019. User Manual of PRETTI_R Version 19.0.1. Technical report. Lloyd's Register.

Rijkens, A.A.K., Mikelic, A., et al., 2022. The hydrodynamic comparison between a conventional and an axe bow frigate hull. Proc. Int. Nav. Eng. Conf. 16. 16th International Naval Engineering Conference and Exhibition incorporating the International Ship Control Systems Symposium, INEC/iSCSS 2022, INEC/iSCSS 2022 ; Conference date: 08-11-2022 Through 10-11-2022. https://www.imarest.org/events/category/categories/imarest-event/international-naval-engineering-conference-and-exhibition-2022. https://doi.org/10.24868/10651

Sapsis, T.P., 2021. Statistics of extreme events in fluid flows and waves. Annu. Rev. Fluid Mech. 53 (Volume 53, 2021), 85–111. https://www.annualreviews.org/content/journals/10.1146/annurev-fluid-030420-032810. https://doi.org/https://doi.org/10.1146/annurev-fluid-030420-032810

Serani, A., Ficini, S., Grigoropoulos, G., Bakirtzogou, C., Broglia, R., Diez, M., Papadakis, G., Goren, O., Danisman, D.B., Scholcz, T., Klinkenberg, H.J., Pehlivan, S., Y1ld1z, S., 2022. Resistance and seakeeping optimization of a naval destroyer by multi-fidelity methods. In: AVT-354 Research Workshop on Multi-Fidelity Methods for Military Vehicle Design. Varna.

Seyffert, H.C., Kana, A.A., 2019. Evaluation of an equivalent design wave method to define lifetime combined loading scenarios for trimarans. In: Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering - OMAE. American Society of Mechanical Engineers Digital Collection. https://doi.org/10.1115/OMAE2019-95497

Shigunov, V., Schellin, T., Troesch, A., Kim, D.-H., Maki, K., 2015. Prediction of loads for ship structural design. Nav. Eng. J. 127, 103–134.

Spiess, A.N., Neumeyer, N., 2010. An evaluation of R2as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach. BMC Pharmacol. 10 (1), 1–11. https://bmcpharma.biomedcentral.com/articles/10.1186/1471-2210-10-6. https://doi.org/10.1186/1471-2210-10-6/FIGURES/4

Temarel, P., Bai, W., Bruns, A., Derbanne, Q., Dessi, D., Dhavalikar, S., Fonseca, N., Fukasawa, T., Gu, X., Nestegård, A., Papanikolaou, A., Parunov, J., Song, K.H., Wang, S., 2016. Prediction of wave-induced loads on ships: progress and challenges. Ocean Eng. 119, 274–308. https://www.sciencedirect.com/science/article/pii/S0029801816300099. https://doi.org/https://doi.org/10.1016/j.oceaneng.2016.03.030

Tupper, E.C., 2004. Main hull strength. In: Tupper, E.C. (Ed.), Introduction to Naval Architecture (Fourth Edition). Butterworth-Heinemann, Oxford, Fourth edition edition. pp. 276–303. https://www.sciencedirect.com/science/article/pii/B978075066554450015X. https://doi.org/https://doi.org/10.1016/B978-075066554-4/50015-X

Van Oers, B., Takken, E., Duchateau, E., Zandstra, R., Cieraad, S., Van Den Broek De Bruijn, W., Janssen, M., 2017. Warship concept exploration and definition at the Netherlands defence materiel organisation introduction: the Netherlands defence materiel organisation. In: Set-Based Design Workshop of the American Society of Naval Engineers. Washington D.C.

Waskito, K.T., Kashiwagi, M., Iwashita, H., Hinatsu, M., 2020. Prediction of nonlinear vertical bending moment using measured pressure distribution on ship hull. Appl. Ocean Res. 101, 102261. https://www.sciencedirect.com/science/article/pii/S0141118720303266. https://doi.org/https://doi.org/10.1016/j.apor.2020.102261