BI-FIDELITY SURROGATE MODELING WITH SUBSET SIMULATION FOR STRUCTURAL RELIABILITY ASSESSMENT

JOÃO PAULO SILVA LIMA

**DOCTORAL THESIS IN CIVIL AND STRUCTURAL ENGINEERING
DEPARTMENT OF CIVIL AND ENVIRONMENTAL ENGINEERING**

SCHOOL DE TECHNOLOGY

UNIVERSITY OF BRASÍLIA

UNIVERSITY OF BRASÍLIA

SCHOOL DE TECHNOLOGY

DEPARTMENT OF CIVIL AND ENVIRONMENTAL ENGINEERING

BI-FIDELITY SURROGATE MODELING WITH SUBSET SIMULATION FOR STRUCTURAL RELIABILITY ASSESSMENT

JOÃO PAULO SILVA LIMA

ADVISOR: FRANCISCO EVANGELISTA JUNIOR

DOCTORAL THESIS IN CIVIL AND STRUCTURAL ENGINEERING

UNIVERSITY OF BRASÍLIA

SCHOOL OF TECHNOLOGY

DEPARTMENT OF CIVIL AND ENVIRONMENTAL ENGINEERING

BI-FIDELITY SURROGATE MODELING WITH SUBSET SIMULATION FOR STRUCTURAL RELIABILITY ASSESSMENT

JOÃO PAULO SILVA LIMA

THESIS SUBMITTED IN PART FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF CIVIL AND STRUCTURAL ENGINEERING

ASSESSMENT COMMITTEE:

_____
Prof. Francisco Evangelista Junior, PhD. (ENC/UnB)
(Advisor)

_____
Prof. Carlos António Pancada Guedes Soares, PhD. (ENC/UnB)
(External committee member)

_____
Prof. Mauro de Vasconcellos Real, PhD. (ENC/UnB)
(External committee member)

_____
Prof. Marcio Muniz de Farias, PhD. (ENC/UnB)
(Internal committee member)

BRASÍLIA/DF, 7 JULY 2023

FICHA CATALOGRÁFICA

LIMA, JOÃO PAULO SILVA
Bi-fidelity surrogate modeling with Subset Simulation for structural reliability assessment [Distrito Federal] 2023.

xxix, 118p., 210×297 mm (ENC/FT/UnB, Doutor, Estruturas e Construção Civil, 2023). Tese de Doutorado – Universidade de Brasília. Faculdade de Tecnologia.
Departamento de Engenharia Civil e Ambiental.

1. Confiabilidade	2. Redes Neurais Artificiais
3. Multitarefa	4. Multifidelidade
5. Análise não linear de Elementos Finitos	
I. ENC/FT/UnB	II. Título (Doutor)

REFERÊNCIA BIBLIOGRÁFICA

LIMA, J. P. S. (2023). Bi-fidelity surrogate modeling with Subset Simulation for structural reliability assessment. Tese de Doutorado em Estruturas e Construção Civil, Publicação TD 06A/23, Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, Brasília, DF, 118p.

CESSÃO DE DIREITOS

AUTOR: João Paulo Silva Lima

TÍTULO: Bi-fidelity surrogate modeling with Subset Simulation for structural reliability assessment.

GRAU: Doutorado ANO: 2023

_____

João Paulo Silva Lima
Faculdade de Ciência e Tecnologia (UFG) – UFG
Rua Mucuri, S/N, Parque Itatiaia.
74968755, Aparecida de Goiânia, GO, Brasil.
joaoplima@ufg.br

ACKNOWLEDGEMENTS

ABSTRACT

BI-FIDELITY SURROGATE MODELING WITH SUBSET SIMULATION FOR STRUCTURAL RELIABILITY ASSESSMENT.

João Paulo Silva Lima

Advisor: Francisco Evangelista Junior

Graduate Program in Civil and Structural Engineering

Brasília, 2023.

This thesis presents a novel Bi-Fidelity Multi-task Learning Model based on a Deep Neural Network (BFMT-DNN) to address the computational challenge of structural reliability analysis applied to complex structures. The main contribution is the development of a novel hyperparameter-optimized BFMT-DNN, that considers the advantages of Bayesian Optimization, focusing on prediction accuracy, stability, and computational efficiency to accesses the reliability of high nonlinear problems. For constructing this surrogate model, the study proceeded throughout two stages before. Firstly, is presented a method based on a Bi-fidelity Kriging surrogate model associated with Subset Simulation for structural reliability analysis. The efficiency of the bi-fidelity Kriging model is evaluated using a stiffened panel reliability problem that demands high computational costs, such as non-linear finite element analysis structural models. The next step proposed a two-stage Bi-Fidelity Deep Neural Network surrogate model in association with Subset Simulation to quantify the uncertainty of structural analysis and assess the probability of failure of high dimensional rare events. In the two steps, the surrogate models can reproduce the non-linear behaviour in the variable's uncertainty analysis, reducing the high computational demand of these problems. Furthermore, the BF-DNN surrogate model used Bayesian optimization to fine-tunning the hyperparameters. The multi-fidelity models used low-fidelity data samples added to the model to predict high-fidelity responses, and, when presenting a good correlation between the fidelities, the assessment of the proposed method showed that the proposed Multi-fidelity Method is a good strategy because it can provide an accurate probability of failure estimation with a lower computational cost. A hyperparameter-optimized BFMT-DNN using low-fidelity data samples added to the model to predict high-fidelity responses for structural collapse behaviour framework is presented in the final analysis. The assessment is realized in an

offshore wind turbine in extreme conditions and described using non-linear Finite Element analysis to obtain multiple outputs. The results show that the proposed multi-fidelity methods can give a precise failure probability estimation with less computational cost.

Keywords: Reliability; Artificial Neural Networks; Multi-task; Multi-fidelity; Non-linear Finite Element Analysis

RESUMO

MODELAGEM SUBSTITUTA BI-FIDELIDADE COM SIMULAÇÃO DE SUBCONJUNTOS PARA AVALIAÇÃO DE CONFIABILIDADE ESTRUTURAL

João Paulo Silva Lima

Orientador: Francisco Evangelista Junior

Programa de Pós-Graduação em Estruturas e Construção Civil

Brasília, 2023.

Esta tese apresenta um novo Modelo de Aprendizagem Multitarefa Bi-Fidelidade baseado em uma Rede Neural Profunda (BFMT-DNN) para abordar o desafio computacional da análise de confiabilidade estrutural aplicada a estruturas complexas. A principal contribuição é o desenvolvimento de um novo BFMT-DNN com hiperparâmetros otimizados, que considera as vantagens da Otimização Bayesiana, focando na precisão da previsão, estabilidade e eficiência computacional para acessar a confiabilidade de problemas não lineares elevados. Para a construção deste modelo substituto, o estudo procedeu ao longo de duas etapas preliminares. Primeiramente, é apresentado um método baseado em um modelo substituto de Krigagem Bi-fidelidade associado à Simulação de Subconjuntos para análise de confiabilidade estrutural. A eficiência do modelo Krigagem Bi-fidelidade é avaliada usando um problema de confiabilidade de painel enrijecido que demanda altos custos computacionais, como modelos estruturais não lineares de análise de Elementos Finitos. A próxima etapa propôs um modelo substituto de Rede Neural Profunda de Bi-fidelidade de dois estágios em associação com Simulação de Subconjunto para quantificar a incerteza da análise estrutural e avaliar a probabilidade de falha de eventos raros de alta dimensão. Nas duas etapas, os modelos substitutos podem reproduzir o comportamento não linear na análise da incerteza da variável, reduzindo a alta demanda computacional desses problemas. Além disso, o modelo substituto BF-DNN usou a Otimização Bayesiana para ajustar os hiperparâmetros. Os modelos de multifidelidade utilizaram amostras de dados de baixa fidelidade adicionadas ao modelo para prever respostas de alta fidelidade e, ao apresentar uma boa correlação entre as fidelidades, a avaliação do método proposto mostrou que o método multifidelidade proposto é uma boa estratégia por fornecer uma probabilidade precisa de estimativa de falha com um custo reduzido computacional. Um modelo BFMT-DNN com

hiperparâmetros otimizados usando amostras de dados de baixa fidelidade adicionadas ao modelo para prever respostas de alta fidelidade para análise de colapso estrutural é apresentado na análise final. A avaliação é realizada em uma turbina eólica offshore em condições extremas e descrita usando análise não linear de Elementos Finitos para obter múltiplas saídas. Os resultados mostram que os métodos de multifidelidade propostos podem fornecer uma estimativa precisa da probabilidade de falha com menor custo computacional.

Keywords: Confiabilidade; Redes Neurais Artificiais; Multitarefa; Multifidelidade; Análise não linear de Elementos Finitos

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS AND SYMBOLS

ADAM	Adaptive Moment Estimation
AK	Adaptive Kriging
ANN	Artificial Neural Networks
BF	Bi-fidelity
BF-DNN	Bi-fidelity Model based on a Deep Neural Network
BFMT-DNN	Bi-fidelity Multi-task Learning Model based on a Deep Neural Network
BISO	Bilinear Isotropic Hardening
BO	Bayesian Optimization
CDF	Cumulative Distribution Function
CFSR	Climate Forecast System Reanalysis
CG	Center of Gravity
CI	Confidence Interval
CM	Center of Mass
CNN	Convolutional Neural Networks
COV	Coefficient of Variation
CSR	Common Structural Rules
DIC	Distance Correlation
DL	Deep Learning
DLC	Design Load Case
DNN	Deep Neural Network
DNV	Det Norske Veritas
E	East
ELU	Exponential Linear Unit
FE	Finite Element
FEA	Finite Element Analysis
FORM	Fist Order Reliability Method
GEV	Generalized Extreme Value
GP	Gaussian Process
HF	High-fidelity
IACS	International Association of Classification Societies
IS	Importance Sampling
LF	Low-fidelity
MCMC	Markov Chain Monte Carlo
MCS	Monte Carlo Simulation
MF	Multi-fidelity
ML	Machine Learning
MLE	Maximum Likelihood Estimation
MLP	Multi-layer Feed-forward Perceptron
MRE	Mean Relative Error
MT	Multi-task
N	North
NE	Northeast
NN_{BF}	Deep Neural Network bi-fidelity stage
NN_{LF}	Deep neural network low-fidelity stage

NOAA	National Oceanic and Atmospheric Administration
NREL	National Renewable Energy Laboratory
OC4	Offshore Code Comparison Collaboration Continuation
OWT	Offshore Wind Turbine
PDF	Probability Density Function
PF	Performance Functions
PCE	Polynomial Chaos Expansion
RBF	Radial Basis Functions
ReLU	Rectified Linear Units
RMSE	Root Mean Squared Error
RNA	Rotor Nacelle Assembly
S	South
SORM	Second Order Reliability Method
SUS	Subset Simulation
SW	Southwest
Tanh	Hyperbolic Tangent Function
W	West
WW3	Wavewatch III
a	Plate length
A	Machine Learning model
a_m	Activation function in the Neural Network
b	Plate breadth
b_f	Flange breadth
bm	Bias in the Neural Network
c	Number of layers in the Neural Network
C_d	Drag coefficient of the piles
c_{HF}	High-fidelity cost
c_{LF}	Low-fidelity cost
C_m	Coefficient of inertia of the piles
COV_{Pf}	Overall Coefficient of Variation in Subset simulation
C_w	Drag coefficient of the tower
d	Dimension of the model
D	External diameter of the structure
dcov	Sample distance variance in Distance Correlation
D_{test}	Test dataset
$D_{tr,HF}$	High-fidelity training dataset
$D_{tr,LF}$	Low-fidelity training dataset
D_{val}	Validation dataset
dvar	Sample distance variance in Distance Correlation
d_w	Web height
E	Young's modulus
F(.)	Function that maps the LF data to the HF level
F_{LF}(.)	Function that maps the LF level
F_{thrust}	Thrust load
G(.)	Performance function
H_s	Significant height of combined wind waves and swell

k	GEV distribution shape parameter
L	Subset simulation threshold level
m	Number of buckling half-waves in the plate longitudinal direction
ms	Mesh size
M_{sw}	Random vertical still water bending moments
M_{tilt}	Aerodynamic tilt moment
M_{wv}	Wave-induced bending moments
M_{yaw}	Aerodynamic yaw moment
n	Number of input neurons in the Neural Network
N_{call}	Number of equivalent high-fidelity samples
n_{HF}	Number of LF observations
n_{LF}	Number of HF observations
N_n	Subset simulation samples in each subset
N_s	Subset simulation samples retained as seed for the MCMC sampling
n_{test}	Number of test observations
N_{ts}	Number of mesh division in the stiffener flange
n_{val}	Number of validation observations
N_x	Number of mesh divisions in x
N_y	Number of mesh divisions in y
N_z	Number of mesh divisions in z
p	Kriging hyperparameter
p_0	Percentile of the samples retained as seed for the MCMC sampling
P_f	Failure probability
r	Pearson Correlation between two variables
r_x	Rotational boundary condition in x
r_y	Rotational boundary condition in y
r_z	Rotational boundary condition in z
S	Subset
std	Standard deviation
t	Number of tasks
t_f	Flange thickness
T_{HF}	High-fidelity task
T_{LF}	Low-fidelity task
t_p	Plate thickness
T_p	Primary wave mean period
T_r	Return period
t_w	Web thickness
u	Horizontal component of the velocity
$\mathbf{u}^*$	Local failure design point
u_x	Translational boundary condition in x
u_y	Translational boundary condition in y
u_z	Translational boundary condition in z
v	Vertical component of the velocity
V_w	Wind velocity
w	Connection weights in the Neural Network
w_{oc}	Column-type imperfection amplitude

w_{op}	Plate imperfection amplitude
w_{os}	Sideways imperfection amplitude
x	X-direction
$\mathbf{x}^{G}$	Vector of basic geometric variables
$\mathbf{x}^{M}$	Vector of basic material variables
$\mathbf{x}_{test}$	Test dataset inputs
$\mathbf{x}_{tr,HF}$	High-fidelity training dataset inputs
$\mathbf{x}_{tr,LF}$	Low-fidelity training dataset input
$\mathbf{x}_{val}$	Validation dataset inputs
y	Y-direction
$\mathbf{y}_{test,HF}$	High-fidelity Test dataset outputs
$\mathbf{y}_{test,LF}$	Low-fidelity test dataset outputs
$\mathbf{y}_{tr,HF}$	High-fidelity training dataset outputs
$\mathbf{y}_{tr,LF}$	Low-fidelity training dataset output
$\mathbf{y}_{val,HF}$	High-fidelity validation dataset outputs
$\mathbf{y}_{val,LF}$	Low-fidelity validation dataset outputs
Z_{v}	Midship cross-section modulus
z	Z-direction
$H_{s,50}$	50-year Wave Significant height
$T_{p,50}$	50-year Wave period
$V_{w,50}$	50-year Wind velocity
α	Gumbel distribution parameter
$\boldsymbol{\alpha}$	Reliability sensitivity vector
α_{r}	Roughness coefficient
β	Reliability index
γ_{d}	Displacement safety factor
γ_{m}	Stress safety factor
δ_{allow}	Allowable jacket tower displacement
δ_{max}	Maximum jacket tower displacement
ε_{avg}	Average relative error
ε_{β}	Mean relative error for the reliability index
$\varepsilon_{,\beta avg}$	Average relative error for the reliability index
ε_{σ}	Mean relative error for the ultimate strength
$\varepsilon_{,\sigma avg}$	Average relative error for the ultimate strength
ζ	Gumbel distribution parameter
η	GEV location parameter
θ	Kriging hyperparameter
Λ	Space of the overall hyperparameters of the Neural Network
λ^{*}	Optimal hyperparameters of the Neural Network
λ_{BF}	Sample size ratio between HF and LF inputs
μ	Mean
μ^{*}	Mean value of the material and geometric variables
ξ	GEV distribution scale parameter
ρ	Scaling factor between High and Low-fidelity models
ρ_{a}	Density of the air
ρ_{c}	Density of the concrete

ρ_s	Density of the steel
σ	Variance
σ_a	Uniaxial compressive strength induced by the ship hull girder bending moments
σ_{allow}	Allowable stress
σ_c	Compressive strength of the concrete
σ_{max}	Maximum von-Misses Stress
σ_t	Tensile strength of the concrete
σ_y	Yield stress
σ_z	Compressive strength along z axis
σ_{zu}	Ultimate compressive strength of the plate
τ	Penalty parameter in the loss function
υ	Poisson's ratio
Φ	Cumulative distribution function
φ	Gumbel distribution parameter
ϕ	Probability density function
χ_u	Ultimate capacity calculation model uncertainty
χ_s	Stress capacity uncertainty factor
χ_d	Displacement capacity uncertainty factor
χ_{site}	Uncertainty related to the site and atmospheric conditions
χ_{aero}	Uncertainty in the aerodynamic properties
χ_{dyn}	Model uncertainty related to the structural dynamics
χ_{mat}	Uncertainty due to variations in material and geometrical properties
χ_{wind}	Model uncertainty in the wind model
χ_{sim}	Statistical uncertainty associated with the design process of sampling wind conditions with a limited number of simulations
ψ	Correlation between two variables in Kriging and Gaussian Process
Ψ	Correlation matrices in kriging model
ω	Exponential Linear Unit parameter

1   INTRODUCTION

As the complexity of facing problems in every field of engineering increases, the solutions get equally complex. One of these problems is the structural reliability analysis, which in many applications is concerned with estimating the probability of a distinct event of interest, often the failure event under prescribed conditions and considering the uncertainties involved, that influences some event concerning the performance criterion [1-3].

Engineers and researchers have well-recognized the importance of reliability in the past few decades, and its improvements have concentrated on making the performance functions more realistic, using proper strength and loading assessment methods based on numerical methods [4]. Furthermore, the limit state functions became implicit for most practical engineering problems, requiring different approaches to render them explicit and to calculate the reliability index [5, 6].

Simulation methods, such as Monte Carlo Simulation (MCS), are an alternative technique to assess reliability. However, MCS often requires numerous high-fidelity model evaluations to ensure a small coefficient of variation of failure probability, which in complex structures, is computationally demanding [3]. The approximate methods, such as the First-Order or Second-Order Reliability Methods (FORM/SORM) (e.g., [7-9]), utilize the Taylor series expansion of the Performance Function to approximate the failure probability. These methods reduce the computational demand compared to simulation-based methods. However, their results may need significant approximation errors when dealing with highly non-linear problems.

Considering rare events with very small failure probabilities (i.e., on the order 1×10^{-4} or less) and complex failure boundaries, advanced simulation methods based on Monte Carlo simulation (MCS) with variance reduction techniques can be used to estimate the probability of the failure. The MCS method with Importance Sampling (MCS-IS) (e.g. [10-12]) and the Subset Simulation (SUS) (e.g. [13-17]) are examples.

Due to the high cost of structural reliability analysis, surrogate models have been widely applied to replace physical experiments or expensive simulations. The surrogate models aim to reproduce a similar Performance Function based on limited calculations to obtain samples, reducing the computational cost. However, there might be a massive gap between the accuracy scale and efficiency to determine the structure's behavior. Many surrogate models have been used for reliability analysis, including Bayesian Networks (e.g., [18, 19]), Support Vector Machine Models (e.g., [12, 16, 20]), Polynomial Chaos Expansion

(PCE) [21], and Radial Basis Functions (RBF) [9, 22]. The Kriging surrogate model has gained significant popularity in the reliability analysis community due to its ability to provide uncertain information (e.g., [8, 23-25]). Artificial Neural Networks (ANN) are amongst the most popular Machine Learning (ML) methods used to generate the response surface in structural reliability problems allowing to assess high dimensional problems (e.g., [2, 26, 27]).

Based on the previous description, finding an effective balance between accurate simulation and computational cost is necessary. Consequently, Multi-fidelity surrogate approaches have attracted significant attention recently for data regression which maximizes the use of cheaper low-fidelity functions to predict high-fidelity outputs. In many scientific domains, models with multiple fidelities are available for analyzing the same phenomena of interest. The MF-Kriging (e.g. [28-30]) and MF-RBF (e.g. [9, 31]) are examples of MF approaches. In complex engineering problems, as an alternative, the different fidelities functions can be obtained by varying the mesh resolution (e.g., [9, 32]) or the boundary conditions (e.g., [33, 34]) of the Finite Element model.

In a multiple-output scenario, most surrogates obtain various outputs approximating one by one. In this context, multi-tasking surrogates solve tasks simultaneously, sharing partially or totally the surrogate structure [35-37]. In addition, an ML model contains model parameters and hyperparameters. The model parameters are obtained by fitting the training data and defining the hyperparameters. Selecting an optimized model that archives the best performance on the data in a reasonable amount of time becomes a problem of optimizing these hyperparameters.

1.1 AIMS AND OBJECTIVES

The main goal of this thesis is to propose a hyperparameter-optimized Bi-Fidelity Multi-task Learning Model based on a Deep Neural Network (BFMT-DNN) for the reliability assessment of complex structures with multiple state limits. Particularly, this thesis will:

- Evaluate the mesh dimension and element type influence in different fidelities functions construction using Finite Elements Models.
- Measure the correlation between the high-fidelity and low-fidelities functions and the influence of the different fidelities' datasets in the global accuracy.
- Propose a reliability analysis of complex structures considering a Bi-fidelity Kriging surrogate model.

- Propose a Bi-Fidelity Deep Neural Network (BF-DNN) to reproduce the performance functions and evaluate the reliability index of complex structures.

- Propose the novel Bi-Fidelity Multi-task Learning Model based on a Deep Neural Network (BFMT-DNN) to reproduce the performance functions and measure the reliability of complex systems with multiple outputs.

1.2 MAIN CONTRIBUITIONS

The contributions of this thesis are listed as follows:

- The main contribution is developing a hyperparameter-optimized BFMT-DNN, considering the advantages of Bayesian Optimization, focusing on prediction accuracy, stability, and computational efficiency to obtain multiple outputs using the same surrogate model.

- In addition, the single-task BF-DNN surrogate model is presented to predict high-fidelity outputs, which is associated with the Subset Simulation, allowing rare events reliability assessments in high dimensional non-linear problems.

- Likewise, a bi-fidelity Kriging surrogate model associated with Subset simulation is defined to reproduce the performance functions and assess the reliability of stiffened panels under axial load.

1.3 THESIS OVERVIEW

The thesis structure, divided into three articles, can be summarized as follows:

- Chapter 2 proposes a reliability analysis of a stiffened panel considering a Bi-fidelity Kriging as a surrogate model of the performance function.

- Chapter 3 presents a Bi-Fidelity Deep Neural Network (BF-DNN) to reproduce the performance functions and evaluate the reliability index of a stiffened panel.

- Chapter 4 presents the offshore wind turbine reliability analysis proposed in the south of Brazil considering a Multi-task Learning Model based on a Bi-fidelity Deep Neural Network.

- Chapter 5 summarizes the Conclusions and Future Research.

2 BI-FIDELITY KRIGING MODEL FOR RELIABILITY ANALYSIS OF THE ULTIMATE STRENGTH OF STIFFENED PANELS

Authors: João Paulo Silva Lima, Francisco Evangelista Jr., Carlos Guedes Soares

Abstract: A method based on a Bi-fidelity Kriging model is proposed for structural reliability analysis. It is based on adding low-fidelity data samples to the model to predict high-fidelity values, thus saving computational effort. Distance Correlation develops the correlation between the Low and High-fidelity functions, initially proposed to assess the correlation between two variables. The bi-fidelity Kriging response surface model's efficiency as a surrogate model is assessed for structural reliability problems that demand high computational costs, such as non-linear finite element analysis structural models. The efficiency assessment is performed by comparing the accuracy of the failure probability predictions based on the Subset Simulation and First-order reliability method using the Bi-fidelity Kriging model as a surrogate for the performance function. The idea is illustrated by considering a representative component of marine structures analyzed by finite element analysis to create bi-fidelity scenarios to assess structural reliability with many variables. The results show that the proposed multi-fidelity method can provide an accurate failure probability estimation with less computational cost.

Keywords: Structural reliability analysis; Multi-fidelity; Surrogate models; Kriging; Non-linear finite element analysis.

2.1 INTRODUCTION

Structural reliability analysis is a central tool for assessing and designing engineering systems. It consists of estimating the probability of failure of a structural system to perform its required function under prescribed conditions, considering the uncertainties involved. The failure event usually involves many systems, and the failure probability is the probability of an undesired performance occurring in those systems.

In engineering practice, it is challenging to perform multidimensional numerical integration when the performance or limit state function involves many random variables. Several reliability analysis techniques have been studied to estimate the probability of failure with numerical methods and tools that can solve structural reliability problems representing the uncertainties involved in designing marine structures [38].

Since the establishment of structural reliability theory in the 1970s, its improvements have concentrated on making the performance functions (PF) more realistic, using proper strength assessment methods based on numerical methods [4]. In general, analytical formulations were substituted by finite element analysis (FEA), which are computationally demanding in non-linear analysis cases that occur when the collapse is considered. Furthermore, the PF became implicit, requiring a different approach to calculating the reliability index [8].

One approach adopted has been to create a surrogate PF based on a limited number of FE calculations reducing the computational burden. Initial formulations used polynomial fits [39, 40] and have been done to identify the best types of polynomials to adopt, how many sample points to use in the fitting and the choice of the initial set of points to be fitted by the surfaces. In marine structural analyses, the response surface method is applied, for example, by Kmiecik and Guedes Soares [41], which used polynomial fits in the reliability of compressed plates. Mohammad and Shiri [42] utilized the polynomial surfaces to predict the limit state function's mooring line tensions and anchor capacities. Dong et al. [43] approximated by polynomial surface the effect of position variations on fatigue notch factors of welded cruciform joints reliability analysis.

For smaller failure probabilities and/or complex failure boundaries involved in structural reliability problems can be computed efficiently using approximate methods such as the first-order or second-order reliability methods (FORM/SORM) (e.g. [7-9]) or using advanced simulation methods based on Monte Carlo simulation (MCS) with variance reduction techniques as MCS method with importance sampling (MCS-IS) (e.g. [10-12]), the directional IS (e.g. [11]), and the Subset Simulation (SUS) (e.g. [13-17]).

Kriging has been used as an essential method for sampling the points for generating response surfaces, and various contributions have appeared to these approaches. Morató and Sriramula [23] used Kriging in a reliability study and calibration of safety factors for offshore wind turbine support structures. Vosooghi *et al.* [24] developed a Kriging-based reliability analysis of subsea pipelines' critical lateral buckling force. Gaspar *et al*. [8] studied Kriging's efficiency for stiffened panels' reliability analysis. Similarly, in Shi *et al.* [25], the prediction accuracy of the Kriging is assessed to estimate the failure probability of marine structures with its application to a stiffened panel.

Adaptive methods have been proposed in which the points sampled for the response surface are decided stepwise. Early efforts proposed by Echard *et al*. [44] the Adaptive

Kriging - Monte Carlo Simulation (AK-MCS) method focused on an iterative approach based on MCS and Kriging surrogate model to assess the reliability of structures more efficiently. The AK-MCS has derived methods based on variance-reduction techniques as a reference for the adaptative methods. IS [10], SUS [17], and directional IS [11] adopted MCS instead. Gaspar *et al.* [3] associated the adaptative method with the Kriging model to reproduce the PF of stiffened panels.

As an alternative to polynomial regression models, approaches based on learning models in general and artificial neural networks (ANN) have been proposed [2, 26, 27]. Other approaches adopted Bayesian Networks (e.g., [18, 19]), Support Vector Machine Models (e.g., [12, 16, 20]), Polynomial Chaos Expansion [21], and Radial Basis Functions [9, 22].

The presented surrogate methods are built on the premise that data for the construction of the surrogate is available from a single fidelity source, often evaluated through fine high-fidelity (HF) simulations. However, experimental and computational simulations often are costly for many complex marine engineering systems [45]. In contrast, the low-fidelity (LF) models reduce either measurement or numerical computational costs but reduce the response's accuracy. The multi-fidelity (MF) investigations have proposed incorporating datasets at different levels and resolutions considering a multi-fidelity (MF) approximation. The MF methods are applied to combine, to some extent, the accuracy of HF solvers with the alleviation of the computational cost of many low-fidelity (LF) samples that estimate the same output [46, 47]. The bi-fidelity (BF) strategy uses the correlation between an HF model and one LF model.

Compared to the single-fidelity Gaussian methods, the MF Gaussian method leads to information sharing across outputs and the asymmetric knowledge transfer from the LF outputs to the HF output to estimate the unknown hyperparameters during the surrogate modeling for forecasting HF outputs [48]. The core feature of the Gaussian Process surrogate model requires prior information [49]. Differently, MF-Kriging does not require prior information [28, 50, 51]. Many extensions and improvements of MF-Kriging methods are gaining popularity. Qian and Wu [52] provided a hierarchical model framework. Chung and Alonso [53] and Han *et al.* [54] proposed incorporating gradient information in co-Kriging. Liu *et al.* [29] presented an adaptive sampling technique using MF-Kriging.

Due to its remarkable performance, MF has gained popularity in structural engineering applications such as the uncertainty quantification field [34, 49, 55] and the multidisciplinary, robust, and multi-objective optimization field [29, 32, 56]. Recently, MF models have also

been incorporated into structural reliability engineering issues. Aruna and Ganguli [57] developed an MF model based on response surface and uncertainty quantification in the context of beam vibrations, combining an FE-LF and HF model. Yoo *et al.* [9] developed a radial basis function MF modeling-based framework for reliability-based design optimization of composite structures under linear buckling. Zhang *et al.* [33] developed an adaptive reliability analysis for MF models using the Gaussian process applied to a transmission tower problem. The literature shows that other researchers have yet to fully address the development of a FE-Bi-fidelity mesh framework to evaluate the reliability of the stiffened panels under axial load.

The main goal of this paper is to propose a bi-fidelity Kriging surrogate model to reproduce the performance functions and assess the reliability analysis of complex structures such as stiffened panels under axial load. The application tackles a relevant problem in marine structures, consisting of a ship structure associated with the hull girder's ultimate strength described by a non-linear FE structural model with 13 variables. The correlation measured by Distance Correlation (DIC), the influence of the HF and LF dataset in the global accuracy, and a comparison of the reliability index predictions based on considering the BF-Kriging surrogate models in association with SUS and FORM are presented. The main contribution of this paper is the development of a multi-fidelity framework for an essential structural component of marine structures, which provides similar accuracy to computationally intensive HF modeling, but with considerable computational time savings.

The remainder of this paper is organized as follows: Section 2 introduces the stiffened panel problem and its FE model. Section 3 presents the reliability analysis steps for the considered problem and the methodology adopted for this proposal. Section 4 related concepts to elaborate the proposed BF-Kriging model; Section 5 discusses results obtained from the study extensively; Section 6 draws the main conclusions of the work conducted.

2.2   STRUCTURAL MODELING FOR STIFFENED PANELS

The efficiency of MF surrogate models based on an FEA for structural reliability analysis is assessed based on an application problem relevant to ship structures, although also applicable to box girder bridges. The issue of the reliability of ship structures associated with the hull girder's ultimate strength has been formulated in [58] and applied to more complex cases in [59-62], including the use of MCS [63]. The collapse strength of plates [41, 64] and stiffened panels [38, 65, 66] is critical to the reliability of ship hulls, and thus they have been much studied.

7

The problem studied here consists of the reliability assessment of the ultimate strength of a typical deck-stiffened plate element under axial compression in the net thickness state approach for local corrosion, as shown in Fig. 2.1, as presented by Gaspar *et al.* [3]. The stiffened panel comprises steel plates, transverse girders, and T-type cross-section longitudinal stiffeners.
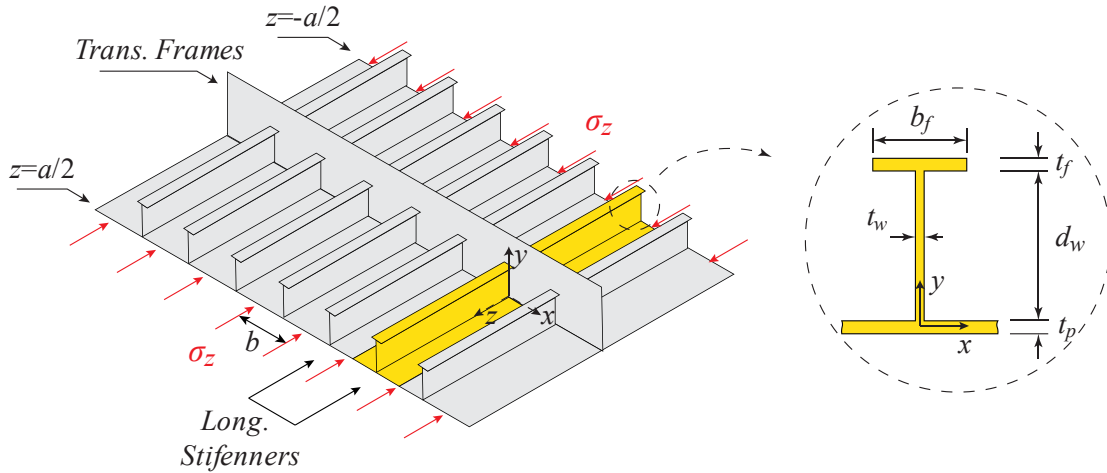
Figure 2.1 – Deck element under uniaxial compression and cross-sections of stiffeners in the stiffened panel.

For the Suezmax oil tanker considered as a case study in the present reliability analysis, the design value for the plate length is a=5450 mm, defined by the stiffener span or spacing between transverse frames; plate breadth b=900 mm for the spacing between stiffeners in the plate. The midship cross-section modulus at the deck for the corroded scantlings is Z_v=37.3 m^3. The plate and stiffener geometry variables are t_p=20.5 mm for the plate thickness, d_w=400 mm for the web height, b_f=100 mm for the flange breadth, t_w=9.5 mm for the web and t_f=14.0 mm for the flange thickness. The material of the plate is AH32 high-strength steel. This material has Young's modulus E=206000 MPa, yield stress σ_y=315 MPa, and Poisson's ratio v=0.3.

2.2.1   Performance function

The plates in the deck are subject to extreme axial compressive loads induced by the ship hull girder bending moments. Therefore, their ultimate compressive strength is an essential design requirement. The still water and vertical wave-induced bending moments are the primary loads that contribute to the longitudinal compressive load applied to a bottom-stiffened panel. The magnitude of these loads depends on several factors, such as the load condition of the ship, its operational profile, and uncontrollable environmental factors, such as the sea state

[67, 68]. Therefore, the safety margin associated with this hull beam failure mode can be described by a PF function of the form [64, 65] :

$$G(x) = \sigma_{zu} - \sigma_a \qquad \text{Equation Chapter 2 Section 1}(2.1)$$

with σ_{zu} the ultimate compressive strength of the plate elements, given as

$$\sigma_{zu} = \chi_u \left(\max \left(\sigma \left(\mathbf{x}^M, \mathbf{x}^G \right) \right) \right) \qquad (2.2)$$

where $\mathbf{x}^M$ is a vector of basic material variables, $\mathbf{x}^G$ is a vector of basic geometric variables, and the factor χ_u is an ultimate capacity calculation model uncertainty. The uniaxial compressive strength σ_a induced by the ship hull girder bending moments given by:

$$\sigma_a = \frac{M_{sw} + M_{wv}}{Z_v} \qquad (2.3)$$

where M_{sw} and M_{wv} are the random vertical still water and wave-induced bending moments, respectively. For the proposed model, $M_{sw,max}$=1483.7 MNm, and $M_{wv,max}$=4603.1 MNm were computed using the IACS-CSR design formulation [69].

2.2.2 Non-linear finite element method for structural analysis

The Finite Element (FE) analysis of the present study consists of the following steps: (*i*) Definition of the model characteristics and choice of the finite element, (*ii*) definition of imperfections, boundary, and load conditions, (*iii*) modeling of the structure's geometry, (*iv*) simulation and (*v*) post-processing and analysis of results. In the present study, the average stress-average strain curves that describe the structural behaviour of the plate under uniaxial compression is calculated using the FE software ABAQUS. The non-linear structural behaviour of the thin plate elements under axial compression is accounted for using the finite element S4R available in the finite element code. S4R is a 4-node general-purpose, quadrilateral, stress/displacement shell element with reduced integration and a large-strain formulation.

The non-linear material behaviour is modeled using a linear-elastic and ideally-plastic material law, neglecting the strain hardening effect. This material law was implemented using the Bilinear Isotropic Hardening (BISO).

Some initial imperfections are applied to the FE model to better represent the inevitable deformations of the ship steel structure's complex fabrication and welding process. Figure 2.2(a) depicts the considered pattern that analytical functions of the form approximate the shapes of these initial imperfections are [70]:

9

- Initial imperfection of local plate panel:

$$w_{0p}(y,z) = w_{op} \sin\frac{m\pi z}{a} \sin\frac{\pi y}{b} \qquad (2.4)$$

- Initial imperfection of stiffener column-type deflection:

$$w_{0c}(z) = w_{oc} \sin\frac{\pi z}{a} \qquad (2.5)$$

- Initial imperfection of the stiffener sideways deflection:

$$w_{0s}(x,z) = w_{os} \frac{x}{d_w} \sin\left(\frac{\pi z}{a}\right) \qquad (2.6)$$

where w_{op}, w_{oc}, and w_{os} are the following design values for the amplitudes given by Paik and Kim [71], with $w_{op}=b/200$ and $w_{oc}=w_{os}=a/1000$. The number of buckling half-waves in the longitudinal direction (m) usually is the integer of the ratio of the longer and shorter side of the plate $m=a/b$. The number of half-waves is adopted as the next largest integer, in this case, $m=7$ [2, 8, 25]. The imperfection amplitudes should be considered as a random variable to account for the contribution of the weld-induced initial imperfections to the uncertainty of the buckling collapse strength predictions.

(a) Initial imperfections (b) Boundary Conditions sections
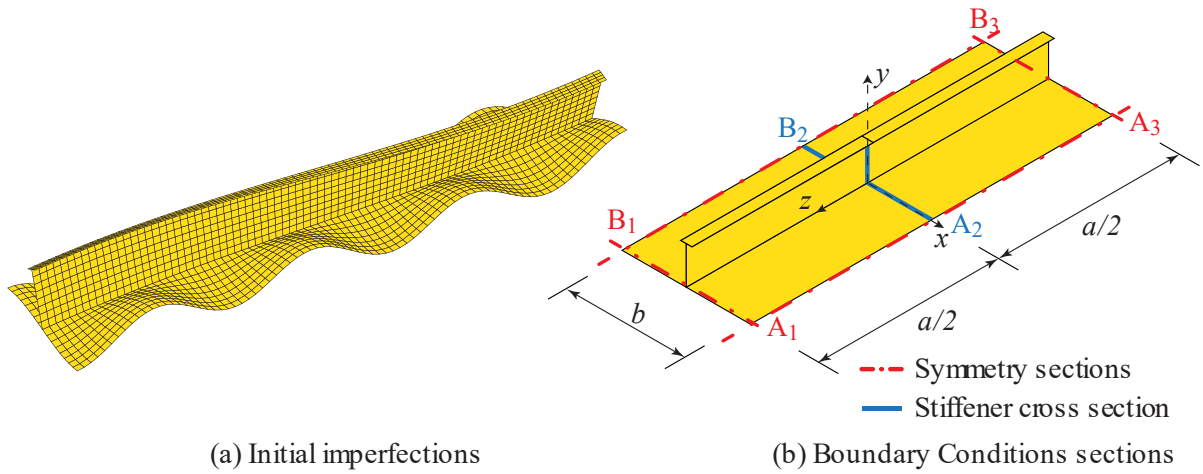
Figure 2.2 – Description of structural model: (a) Initial geometrical imperfections of the plate (scale factor 50×) (b), and Boundary conditions symmetry sections in the FEA.

According to Xu *et al.* [72], it is necessary for an adequate extension of the model used in the FEA to characterize the structure that surrounds the panel of interest and to reduce the uncertainties introduced by the boundary conditions on the final results. The FE structural model adopted for the reinforced plate elements is a half-plus half-span (1/2+1/2) model implemented by Paik *et al.* [71], and such an approach has been employed in Gaspar *et al.* [3]. Since the panel is symmetrical, only one of the reinforcements of the reinforced panel with

attached plating is analyzed, as shown in Fig. 2.2. The boundary conditions of the stiffened panel are represented by $\left[u_x, u_y, u_z, r_x, r_y, r_z \right]$, and are applied as per Fig. 2.2(b). The red dashed lines are the symmetry lines of the model, and the blue lines are the stiffener cross-section at the midspan. For the present analysis, consider the following boundary conditions, where "0" indicates translational or rotational constraint and "1" means no restriction:

- A_1-A_3 and B_1-B_3 border (symmetric conditions): [0,1,1,1,1,0];

- A_1-B_1 border (symmetric condition): [1,1,0,0,0,1];

- A_3-B_3 border (symmetric condition): [1,1, d_z,0,0,1];

- A_2-B_2 for plate nodes: [1,0,1,1,1,1]

- A_2-B_2 for stiffener web nodes: [0,1,1,1,1,1]

where d_z indicates the displacement imposed at the transverse edge. The displacements of nodes on borders A_1-A_3 and B_1-B_3 along the x-direction are linked. This process safeguards uniform displacement even under a point force applied to one of the coupled nodes and avoids the local imperfections of the profile.

2.2.3   Stochastic models for basic random variables

The ten basic random variables considered in the problem of Section 4 are given, and the two random variables considered in the load components of the performance function in Eq. 2.1 are shown in Table 2.1, both with their corresponding probability density distribution (PDF) and statistical moments mean (μ) and standard deviation (*std*) of the associated distribution functions.

The ultimate capacity calculation uncertainty factor χ_u is defined by Hørte *et al.* [73] as the e uncertainty measures dependent upon the methods used to estimate the ultimate structural capacity. The authors describe a normal distribution with μ=1.05, and a std=0.1 represents the model uncertainty χ_u in the prediction of ultimate capacity.

The stochastic models adopted for the basic geometric and material properties variables and the amplitude of the welding-induced initial imperfections proposed by Gaspar *et al.* [3] are commonly used in structural reliability analysis (e.g. [2, 8, 73]). The design value considered for σ_y=315 MPa corresponds to the probability density function's 5% percentile characteristic value of the lognormal distribution with COV=6% and lower limit=0, resulting in a mean of μ=348 MPa [74]. It is suggested that the spacing between transverse frames, the

midship cross-section modulus, the stiffener web height, and the stiffener flange breadth can be treated as deterministic variables. The thicknesses of the attached plating and stiffener cross-section should be treated as random variables, as well as the material yield stress and Young's modulus. The Poisson's ratio can be treated as a deterministic variable.

Table 2.1 – Stochastic models of the strength of basic random variables.

Variable	Units	P.D.	μ	COV (%)	*std*
t_p	mm	Normal	20.50	2.0	0.41
t_w	mm	Normal	9.50	2.0	0.19
t_f	mm	Normal	14.00	2.0	0.28
$\sigma_{y,p}$	MPa	Lognormal	348.00	6.0	20.88
$\sigma_{y,s}$	MPa	Lognormal	348.00	6.0	20.88
E_p	MPa	Lognormal	206000	6.0	12360
E_s	MPa	Lognormal	206000	6.0	12360
w_{op}	mm	Lognormal	2.30	50.0	1.15
w_{os}	mm	Lognormal	2.80	50.0	1.40
w_{oc}	mm	Lognormal	2.80	50.0	1.40
M_{sw}	MNm	Normal	1483.70	29.0	430.30
M_{wv}	MNm	Gumbel	4603.15	9.0	414.30
χ_u	-	Normal	1.05	9.5	0.10

For the welding-induced initial imperfections amplitudes stochastic modeling, some assumptions are usually made regarding the statistical parameters due to the lack of data available in the literature. In this study, the probabilistic model for the amplitudes of the initial deflections is based on Amlashi *et al.* [75]. The parameters of the lognormal distributions are derived considering the design amplitudes as 95% percentile characteristic values, and the coefficient of variation (COV) is assumed to equal 0.50.

The ship hull girder bending moments are the load basic random variables of the problem. The stochastic models adopted for the hull girder vertical bending moments along a ship voyage are defined based on Hørte *et al.* [73] and as specified in IACS [69]. A normal distribution with $\mu=0.7M_{sw,max}$ and $std=0.2M_{sw,max}$ describes the still water bending moment (M_{sw}).

Considering a reference period $T_r=1.0$ year for the extreme values of vertical wave-induced bending moment representative of a North Atlantic crossing, the stochastic model proposed by Guedes Soares *et al.* [58, 67] defines the wave-induced bending moment. The peak values of vertical wave-induced bending moment at a random moment are based on a two-parameter Weibull distribution with shape parameter $k=1$ and scale ξ parameter

satisfying: $P\left[M_{wv} > M_{wv,max}\right] = 10^{-8}$, where $M_{wv,max}$ is the maximum vertical wave-induced bending moment considered for design purposes [69]. The Gumbel distribution then describes the peak values over the T_r:

$$F_{M_{wve}}(M_{wv}) = \exp\left[-\exp\left(-\frac{M_{wv} - \varphi}{\zeta}\right)\right] \tag{2.7}$$

The Gumbel model parameters can be derived based on the initial Weibull distribution parameters φ and ζ, by Guedes Soares and Teixeira [76] as $\varphi = \xi(\ln n_c)^{1/k}$ and $\zeta = \xi/k \, (\ln n_c)^{1-k/k}$, where n_c corresponds to the mean number of load cycles expected over the operation period T_r [76]. Typically, the number of wave cycles corresponding to the return period is calculated considering an average wave period of $T_p=8$ s, which is the one that is applicable for the areas of the North Atlantic. More details can be found in Hørte *et al.* [73] and Gaspar *et al.* [65].

2.3 RELIABILITY ANALYSIS USING A BF SURROGATE MODEL

For the reliability analysis of the stiffened panels under axial load considering a BF surrogate model, two stages are considered as presented in the Fig. 2.3 flowchart.

In Stage 01, the Kriging surrogate model is associated with a Finite Element mesh analysis to determine the Bi-fidelity Kriging (BF-Kriging) surrogate model. The stages of the proposed algorithm may be summarized as follows:

Stage 01 - Step 1: In the first analysis, the impact of the correlation between LF and HF functions considering the Distance Correlation and the magnitude of the cost ratio are then investigated.

Stage 01 - Step 2: Two datasets is adopted to train the surrogate model. A total of n_{HF} HF samples are associated with a total of n_{LF} LF observations and is used for fit metamodeling of BF data. Therefore, an HF dataset $D_{tr,HF} = \left\{\left(\mathbf{x}_{tr,HF}^{(i)}, \mathbf{y}_{tr,HF}^{(i)}\right) : 1 \le i \le n_{HF}\right\}$ with $\mathbf{x}_{tr,HF}$ inputs, and $\mathbf{y}_{tr,HF}$ outputs, and an LF dataset $D_{tr,LF} = \left\{\left(\mathbf{x}_{tr,LF}^{(i)}, \mathbf{y}_{tr,LF}^{(i)}\right) : 1 \le i \le n_{LF}\right\}$ with $\mathbf{x}_{tr,LF}$ inputs, and $\mathbf{y}_{tr,LF}$ outputs are used to estimate the Kriging hyperparameters during the fitting stage of the BF-Kriging model; A total of n_{test} samples with $\mathbf{x}_{test}$ inputs, and $\mathbf{y}_{test,HF}$ are used in the dataset $D_{test} = \left\{\left(\mathbf{x}_{test}^{(i)}, \mathbf{y}_{test,HF}^{(i)}\right) : 1 \le i \le n_{test}\right\}$ to measure the BF-Kriging accuracy to predict HF outputs. The Latin hypercube technique is used for sampling the datasets.
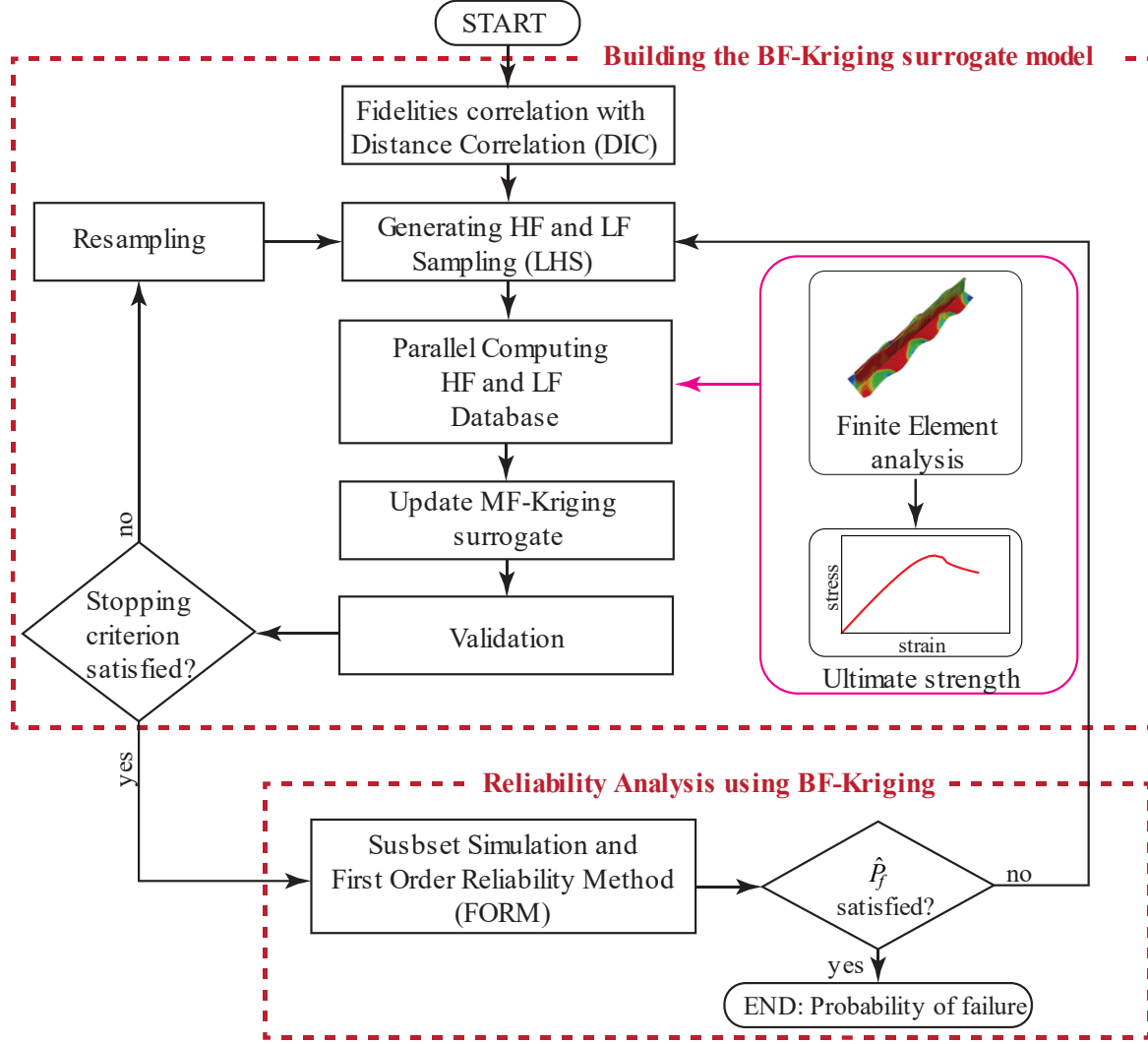
Figure 2.3 – Flowchart of the BF-Kriging using finite element mesh.

Stage 01 - Step 3: In parallel computing, the initial output $\mathbf{y}_{tr}^{(0)} = \left[\mathbf{y}_{tr,\text{HF}}; \mathbf{y}_{tr,\text{LF}} \right]$ and $\mathbf{y}_{test}$, which consists of the stiffened panel's ultimate strength, is obtained by non-linear FEA.

Stage 01 - Step 4: A hyperparameter Kriging optimization is realized considering Genetic Algorithm (GA).

Stage 01 - Step 5: Repeat Steps 2 and 3 and the corresponding hyperparameters optimization until the convergence criteria are satisfied. To quantify the accuracy of the proposed BF-Kriging surrogate model, the Mean relative error (MRE), considering the sum of absolute errors divided by the sample size, is given as:

$$\varepsilon_{\hat{\sigma}} = \frac{1}{n_{test}} \sum_{i=1}^{n_{test}} \left(\frac{\left| y_{test,i}^* - y_{test,i} \right|}{y_{test,i}} \right). \tag{2.8}$$

where $y_{test,i}^{*}$ denotes the output of the BF-Kriging surrogate model and $y_{test,i}$ the target value in the D_{test} obtained by FEA.

Stage 02 – The failure of probability P_f is estimated using the BF-Kriging surrogate model in association with the Subset Simulation (Sus) and FORM. The Sus is used to estimate the $P_{f\text{-}ref}$ using the FEA, and the FORM method is used to estimate the sensitivity analysis of the variables in the reliability analysis.

Mathematically, the failure of probability P_f can be estimated by solving the multifold probability integral defined as:

$$P_f(\mathbf{x}) = P(G(\mathbf{x}) < 0) = \int_{G(\mathbf{x})<0} f(\mathbf{x}) dx \tag{2.9}$$

where the random vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ represents the sources of uncertainty in the analysis and $G(\mathbf{x})$ is a PF function. In reliability analysis, equating $G(\mathbf{x})=0$ provides the limit state function, which separates the safe region $G(\mathbf{x})>0$ from the failure region $G(\mathbf{x})<0$. Moreover, $f(\mathbf{x})$ is the joint probability density function of $\mathbf{x}$. To find the local design point $\mathbf{u}^*(i)$ for the structural reliability problem, it is necessary to solve the following constrained optimization problem based on the FORM algorithm:

$$\mathbf{u}^* = \arg\min_{\mathbf{u}} \{\|\mathbf{u}\| \mid G(\mathbf{u}) = 0\} \tag{2.10}$$

where the design point, $\mathbf{u}^*$ is the point on $G(\mathbf{u})=0$ with the shortest distance $\|\mathbf{u}\|$ to the origin. The $\mathbf{u}^*$ value defines the so-called Hasofer-Lind reliability index $\beta = \|\mathbf{u}^*\|$ and the FORM failure probability prediction:

$$P_f \approx \Phi(-\beta) \tag{2.11}$$

where Φ is the standard normal probability distribution function. Sensitivity factors for each basic random variable can also be computed based on partial derivatives of the BF-Kriging surrogate model at the design point $\mathbf{u}^*$ [3, 25]:

$$\alpha_i = \frac{(\partial(\mathbf{u}) / \partial u_i)_{\mathbf{u}=\mathbf{u}^*}}{\|\nabla(\mathbf{u})\|_{\mathbf{u}=\mathbf{u}^*}}, \quad i = 1, \cdots, n. \tag{2.12}$$

The magnitude of the α-values quantifies the importance of the basic random variables. The sign of the sensitivity factor α_i ($i=1,\ldots n$) defines the contribution of variable u_i ($i=1,\ldots n$) to the reliability index β, in terms of increase (positive α) or decrease (negative α).

2.3.1 Subset simulation for reliability estimation

A brief introduction of the SS technique, largely following Ref. [13], is expressed. The Subset Simulation (SUS) is an adaptive Monte Carlo method proposed estimating small failure probabilities in high-dimensional problems, and details can be found in several other studies [14-16]. The SUS is based on nested sets $F_1 \supset F_2 \supset \cdots F_M$, where $F_M = F$ is the failure event, and expresses the failure probability P_f as a product of conditional probabilities:

$$\hat{P}_f = \mathrm{P}(F) = P_{F_1} \prod_{j=2}^{M} \mathrm{P}(F_j \mid F_{j-1}) \tag{2.13}$$

where P_{F_1} is the first unconditional failure probability computed as the fraction of samples exceeding the current threshold level L_1, and $\mathrm{P}(F_j \mid F_{j-1})$ are the subsequent conditional failure probabilities conditional on exceeding the prior intermediate thresholds in level L_{j-1}, and are computed as the fraction of samples exceeding the threshold level L_j.

In expressing $\hat{P}_f$ as a product of larger failure probabilities, SUS creates intermediate failure thresholds L before the required zero thresholds. As for SUS, the algorithm starts with a direct MCS to estimate P_{F_1}, while a Markov chain Monte Carlo is used to estimate $P_{F_j \mid F_{j-1}}$.

In SUS, the values $F(\mathbf{x}^{(k)})$, $k = 1, ..., N_n$, of the N_n samples are in the subset $S^{(1)}$. Crude MCS generates the samples independently and identically distributed, assuming the Nataf transformation. In contrast, samples in the subsequent subsets are generated by Markov Chain Monte Carlo (MCMC) simulation and correlated, considering the dependence of the seeds for each Markov chain. The samples falling between two subsequent intermediate failure thresholds L_{j-1} and L_j constitute a subset. The intermediate failure thresholds must be specified to estimate these intermediate failure probabilities. The range of the samples $N_s = p_0 N_n$ are retained in each step and serves as the seed for the Markov chains. The value $p_0 = 0.1$ is often used in the literature, which makes F_1 a relatively frequent event.

The authors Au and Beck [13] proposed a Metropolis–Hastings algorithm, and this method is popular for simulating the conditional samples in $P_{F_j \mid F_{j-1}}$. However, occasionally it can lead to degenerate sampling when dealing with models having geometrically complex performance functions. Papaioannou *et al.* [15] proposed a delayed rejection MCMC method to alleviate these issues and is used in this study.

In addition, an approximate COV estimate of the intermediate failure probability in the subset is given by [16]:

$$\hat{\delta}_j = \begin{cases} \sqrt{\left(1 - P_{F_1}\right) \Big/ \left(P_{F_1} N_n\right)} & \forall \quad j = 1 \\ \sqrt{\left(1 - P_{F_1 | F_{j-1}}\right) \Big/ \left(P_{F_1 | F_{j-1}} N_n\right)} & \forall \quad 1 < j \le N_s \end{cases} \tag{2.14}$$

where N_n is the number of samples in a subset. The overall COV estimate over the required failure probability is:

$$COV_{\hat{P}_f} = \sqrt{\sum_{i=1}^{N_S} \hat{\delta}_i^2} \tag{2.15}$$

2.4   BI-FIDELITY SURROGATE MODEL FOR RELIABILITY ANALYSIS

A surrogate model is an approximate model for studying complex input-output relationships exhibited by another more complex model. Therefore, as mathematical models are an abstraction from the real world, surrogated models are yet another abstraction of that mathematical model. To approximate a function f is initially defined as a set of sample data computed by a sampling plan at a set of points in the domain of interest determined [28, 77]. In this way, the surrogate is trained to predict (interpolate) the outputs resulting from other input values not included in the training sample.

Several MF surrogate models based on different single-fidelity surrogate models have been developed to take advantage of both HF and LF models. The variable fidelity analysis is an essential tool that makes it possible to break the barriers of computational cost using a combination of LF and HF data.

2.4.1   The Fidelities correlation considering Distance Correlation (DIC)

Although LF samples are noisy and skewed, they usually correlate strongly with HF samples [78]. As such, leveraging this correlation to avoid total reliance on HF data is possible. Preview researchers such as Toal [79] and Shi et al. [80] adopted the Pearson correlation coefficient, commonly used as an indicator to measure the correlation between two random variables, to measure the correlation of the HF and LF functions, which can be calculated as :

$$r = \frac{\text{cov}\left(y_{\text{HF}}, y_{\text{LF}}\right)}{\sqrt{\text{var}\left(y_{\text{HF}}\right) \text{var}\left(y_{\text{LF}}\right)}}. \tag{2.16}$$

Respectively, y_{HF} and y_{LF} are a set of n observations of the HF and LF data for identical inputs with finite and positive variances. Since the calculation of r is straightforward, it can just measure the linear correlation of the HF and LF functions.

Distance correlation (DIC) is introduced by Székely *et al.* [81-83] to measure the dependence between random vectors providing a scalar measure of multivariate independence that characterizes independence of random vectors, not necessarily in the same dimension. It is more general than the classical Pearson product-moment correlation because the DIC has a significant benefit for characterizing a non-linear relationship of random vectors. Moreover, DIC satisfies $0 \le \text{DIC}\left(y_{\text{HF}}, y_{\text{LF}}\right) \le 1$ and equals zero if independence holds.

In the correlation fidelities functions context, Distance covariance (*d*cov) and DIC are dependency measures between two random vectors *X* and *Y,* discrete observation points where the fidelities y_{HF} and y_{LF} are observed. Moreover, *d*cov and DIC are defined for vectors of arbitrary dimensions, $X \in \mathbb{R}^p$, and $Y \in \mathbb{R}^q$.

The squared $d\text{cov}^2(X,Y)$ is defined as a weighted distance between the joint characteristic function $\phi_{X,Y}\left(t,s\right)$ and the product of the marginals $\phi_X\left(t\right)$ and $\phi_Y\left(t\right)$, considering the analyzed vectors *t* and *s*. The *d*cov(*X,Y*) is then the nonnegative number that verifies:

$$\text{dcov}^2\left(X,Y\right) = \int\limits_{\mathbb{R}^{p+q}} \left\| \phi_{X,Y}\left(t,s\right) - \phi_X\left(t\right)\phi_Y\left(s\right) \right\|^2 w\left(t,s\right)dtds \tag{2.17}$$

where $\left\| \cdot \right\|$ stands for the Euclidean norm in $\mathbb{R}^d$, *p* and *q* are the dimensionalities of *X* and *Y,* and the analyzed vectors *t* and *s*. The choice of the weight function is detailed in Lemma 1 of Székely *et al.* [82]. According to the authors, it is natural to choose the weight function given by:

$$w\left(t,s\right) = \left(c_p c_q \left\| t \right\|_p^{1+p} \left\| s \right\|_q^{1+q}\right)^{-1} \tag{2.18}$$

where c_d is half the surface area of the unit sphere in $\mathbb{R}^d$, determined by $c_l = \pi^{(l+1)/2} \big/ \Gamma\left(\left(l+1\right)/2\right)$, for $l \in \mathbb{N}$, and $\Gamma\left(\cdot\right)$ is the gamma function.

Analogously to classical Pearson correlation, the DIC(*X,Y*) is defined from the distance covariance as:

$$DIC^2\left(X,Y\right) = \begin{cases} \dfrac{d\text{cov}^2\left(X,Y\right)}{\sqrt{d\text{cov}^2\left(X,X\right)d\text{cov}^2\left(Y,Y\right)}} & \text{if } d\text{cov}^2\left(X,X\right)d\text{cov}^2\left(Y,Y\right) > 0 \\ 0 & \text{if } d\text{cov}^2\left(X,X\right)d\text{cov}^2\left(Y,Y\right) = 0. \end{cases} \tag{2.19}$$

Although the apparent complexity of the definitions in Eq. 2.17 and 2.18, *d*cov and DIC have an empirical simple parameter-free estimator (see Theorem 1, Székely *et al.* [82]). Given a

sample $\left\{\left(x_i, y_i\right)\right\}_{i=1}^{N}$ of N observations of the joint random vector (X,Y), is defined the double-centered distance matrices A and B as follow:

$$A_{i,j} = a_{i,j} - \frac{1}{N}\sum_{l=1}^{N} a_{il} - \frac{1}{N}\sum_{k=1}^{N} a_{kj} + \frac{1}{N^2}\sum_{k,l=1}^{N} a_{kl}, \text{ and} \tag{2.20}$$

$$B_{i,j} = b_{i,j} - \frac{1}{N}\sum_{l=1}^{N} b_{il} - \frac{1}{N}\sum_{k=1}^{N} b_{kj} + \frac{1}{N^2}\sum_{k,l=1}^{N} b_{kl}, \tag{2.21}$$

where $a_{ij} = \left\| x_i - x_j \right\|_p$ and $b_{ij} = \left\| y_i - y_j \right\|_q$. Then, the sample $d$cov is the square root of

$$d\operatorname{cov}_N^2 = \frac{1}{N^2}\sum_{i,j=1}^{N} A_{i,j} B_{i,j}. \tag{2.22}$$

Similarly, the sample distance correlation is the standardized sample covariance:

$$DIC_N^2(X,Y) = \begin{cases} \dfrac{d\operatorname{cov}_N^2(x,y)}{\sqrt{d\operatorname{cov}_N^2(x,x)\,d\operatorname{cov}_N^2(y,y)}} & \text{if } d\operatorname{cov}_N^2(x,x)\,d\operatorname{cov}_N^2(y,y) > 0 \\ 0 & \text{if } d\operatorname{cov}_N^2(x,x)\,d\operatorname{cov}_N^2(y,y) = 0. \end{cases} \tag{2.23}$$

2.4.2 A BF-Kriging surrogate model

A more accurate HF dataset with $\mathbf{x}_{\mathrm{HF}}$ input points and $\mathbf{y}_{\mathrm{HF}}$ output values associated with a less accurate LF dataset with $\mathbf{x}_{\mathrm{LF}}$ inputs and $\mathbf{y}_{\mathrm{LF}}$ outputs is necessary to fit a bi-fidelity model. According to Forrester and Sóbester [84], to create the BF model considering the Kriging surrogate, the formulation of a correction process is simplified if the HF function sample locations coincide with a subset of the LF sample locations $\left(\mathbf{x}_{\mathrm{HF}} \subset \mathbf{x}_{\mathrm{LF}} \right)$. Using both sets of data, LF and HF, the concatenated sample gives the combined BF set of sample points:

$$\mathbf{x}_{\mathrm{BF}} = \begin{pmatrix} \mathbf{x}_{\mathrm{LF}} \\ \mathbf{x}_{\mathrm{HF}} \end{pmatrix} = \left(\mathbf{x}_{\mathrm{LF}}^{(1)}, \ldots, \mathbf{x}_{\mathrm{LF}}^{(n_{\mathrm{LF}})}, \mathbf{x}_{\mathrm{HF}}^{(1)}, \ldots, \mathbf{x}_{\mathrm{HF}}^{(n_{\mathrm{HF}})} \right)^{\mathrm{T}}. \tag{2.24}$$

As with Kriging, the responses in $\mathbf{x}_{\mathrm{BF}}$ are treated as realizations of a stochastic process. The random field is given as

$$\mathbf{y}_{\mathrm{BF}} = \begin{pmatrix} \mathbf{y}_{\mathrm{LF}}\left(\mathbf{x}_{\mathrm{LF}}\right) \\ \mathbf{y}_{\mathrm{HF}}\left(\mathbf{x}_{\mathrm{HF}}\right) \end{pmatrix} = \left(\mathbf{y}_{\mathrm{LF}}\left(\mathbf{x}_{\mathrm{LF}}^{(1)}\right), \ldots, \mathbf{y}_{\mathrm{LF}}\left(\mathbf{x}_{\mathrm{LF}}^{(n_{\mathrm{LF}})}\right), \mathbf{y}_{\mathrm{AF}}\left(\mathbf{x}_{\mathrm{HF}}^{(1)}\right), \ldots, \mathbf{y}_{\mathrm{HF}}\left(\mathbf{x}_{\mathrm{HF}}^{(n_{\mathrm{HF}})}\right) \right)^{\mathrm{T}}. \tag{2.25}$$

The auto-regressive model of Kennedy and O'Hagan (2000) is adopted, which assumes that $\operatorname{cov}\left\{ \mathbf{y}_{\mathrm{HF}}\left(\mathbf{x}^{(i)}\right), \mathbf{y}_{\mathrm{LF}}(\mathbf{x}) \middle| \mathbf{y}_{\mathrm{LF}}\left(\mathbf{x}^{(i)}\right) \right\} = 0$, $\forall \mathbf{x} \neq \mathbf{x}^{(i)}$. This means that no more can be learned about $\mathbf{y}_{\mathrm{HF}}(\mathbf{x}^{(i)})$ using the LF values as a reference if the value of the HF function at $\mathbf{x}^{(i)}$ is known.

Consider that the HF model is approximated by scaling the LF model by multiplying it with a constant scaling factor ρ and summing it with a Gaussian process $Z_d(\cdot)$ that is the difference between $\rho Z_{LF}(\cdot)$ and $Z_{HF}(\cdot)$.

$$Z_{HF}(\mathbf{x}) = \rho Z_{LF}(\mathbf{x}) + Z_d(\mathbf{x}). \tag{2.26}$$

Here $Z_{HF}(\cdot)$ and $Z_{LF}(\cdot)$ are Gaussian processes that represent the local features of the HF and LF model, respectively. A Kriging prediction defines the Z_{LF} Gaussian processes. The Kriging surrogate model has a mean base term $\hat{\mu}$ (the circumflex denotes a maximum likelihood estimate, MLE) plus a stationary Gaussian process, $Z(x)$, with zero mean and covariance:

$$\text{cov}\left[y\left(x^{(i)}\right), y\left(x^{(j)}\right) \right] = \sigma^2 \psi\left(x^{(i)}, x^{(j)}\right) \tag{2.27}$$

where σ^2 is the variance, and $\psi(\cdot)$ are correlations between a random variable at the point to be predicted and at the sample data points:

$$\psi\left(x^{(i)}, x^{(j)}\right) = \exp\left(-\sum_{k=1}^{d} \hat{\theta}_k \left\| x_k^{(i)} - x_k^{(j)} \right\|^{\hat{p}_k} \right) \tag{2.28}$$

where d is the number of design variables and $\hat{\theta}$, $\hat{p}_k$ are spatially related parameters. In several situations, it is assumed that $\hat{p}_k = 2$ rather than using an MLE. The covariance matrix based on the Kriging model is then:

$$\text{cov}\left[\mathbf{y}(\mathbf{x}), \mathbf{y}(\mathbf{x}) \right] = \sigma^2 \tag{2.29}$$

Similarly, several quantities need to be defined to obtain the BF-Kriging estimator. Whereas Kriging is had a covariance matrix (Eq. 2.29), BF-Kriging has a more complex covariance matrix involving individual terms associated with both high-fidelity and low-fidelity models:

$$
\begin{aligned}
cov\{\mathbf{y}_{LF}(\mathbf{x}_{LF}), \mathbf{y}_{LF}(\mathbf{x}_{LF})\} &= \text{cov}\{Z_{LF}(\mathbf{x}_{LF}), Z_{LF}(\mathbf{x}_{LF})\} \\
&= \sigma_{LF}^2 {}_{LF}(\mathbf{z}_{LF}, \mathbf{z}_{LF}) \\
cov\{\mathbf{y}_{HF}(\mathbf{x}_{HF}), \mathbf{y}_{LF}(\mathbf{x}_{LF})\} &= \text{cov}\{\rho Z_{LF}(\mathbf{x}_{LF}) + Z_d(\mathbf{x}_{LF}), Z_{LF}(\mathbf{x}_{HF})\} \\
&= \rho \sigma_{LF}^2 {}_{LF}(\mathbf{z}_{LF}, \mathbf{z}_{HF}) \\
cov\{\mathbf{y}_{HF}(\mathbf{x}_{HF}), \mathbf{y}_{HF}(\mathbf{x}_{HF})\} &= \text{cov}\{\rho Z_{LF}(\mathbf{x}_{HF}) + Z_d(\mathbf{x}_{HF}), \rho Z_{LF}(\mathbf{x}_{HF}) + Z_d(\mathbf{x}_{HF})\} \\
&= \rho^2 \text{cov}\{Z_{LF}(\mathbf{x}_{HF}), Z_{LF}(\mathbf{x}_{HF})\} + \text{cov}\{Z_d(\mathbf{x}_{HF}), Z_d(\mathbf{x}_{HF})\} \\
&= \rho^2 \sigma_{LF}^2 {}_{LF}(\mathbf{z}_{HF}, \mathbf{z}_{HF}) + \sigma_d^2 {}_d(\mathbf{z}_{HF}, \mathbf{z}_{HF})
\end{aligned}
\tag{2.30}
$$

For example, the notation $\Psi_{\mathrm{LF}}\left(\mathbf{z}_{\mathrm{LF}},\ \mathbf{z}_{\mathrm{HF}}\right)$ represents the matrix constituted by the correlation coefficient of $\mathbf{x}_{\mathrm{HF}}$ and $\mathbf{x}_{\mathrm{LF}}$. in the low-fidelity Kriging model. The complete covariance matrix is given as follows:

$$\mathbf{C}=\begin{pmatrix} \sigma_{\mathrm{LF}}^{2}\,\Psi_{\mathrm{LF}}\left(\mathbf{z}_{\mathrm{LF}},\ \mathbf{z}_{\mathrm{LF}}\right) & \rho\sigma_{\mathrm{LF}}^{2}\,\Psi_{\mathrm{LF}}\left(\mathbf{z}_{\mathrm{LF}},\mathbf{z}_{\mathrm{HF}}\right) \\ \rho\sigma_{\mathrm{LF}}^{2}\,\Psi_{\mathrm{LF}}\left(\mathbf{z}_{\mathrm{HF}},\ \mathbf{z}_{\mathrm{LF}}\right) & \rho^{2}\sigma_{\mathrm{LF}}^{2}\,\Psi_{\mathrm{LF}}\left(\mathbf{z}_{\mathrm{HF}},\mathbf{z}_{\mathrm{HF}}\right)+\sigma_{\mathrm{d}}^{2}\,\Psi_{\mathrm{d}}\left(\mathbf{z}_{\mathrm{HF}},\ \mathbf{z}_{\mathrm{HF}}\right) \end{pmatrix} \tag{2.31}$$

The correlations used in the correlation matrices $\boldsymbol{\Psi}$ are presented in Eq. 2.28. Since there are two ψ_{LF} and ψ_{d} correlations, it is necessary to estimate the hyper-parameters $\boldsymbol{\theta}_{\mathrm{LF}}$, $\boldsymbol{\theta}_{\mathrm{d}}$, $\mathbf{p}_{\mathrm{LF}}$, $\mathbf{p}_{\mathrm{d}}$, and a scale factor ρ. Therefore, by maximizing the ln-likelihood of the LF data field, obtain MLEs for the parameters μ_{LF}, σ_{LF}^{2}, $\boldsymbol{\theta}_{\mathrm{LF}}$, and $\mathbf{p}_{\mathrm{LF}}$:

$$\mathrm{MLE}_{\mathrm{LF}}=-\frac{n_{\mathrm{LF}}}{2}\ln\left(\sigma_{\mathrm{LF}}^{2}\right)-\frac{1}{2}\ln\left|\det\left(\Psi_{\mathrm{LF}}\left(\mathbf{z}_{\mathrm{LF}},\ \mathbf{z}_{\mathrm{LF}}\right)\right)\right|-\frac{\left(\mathbf{y}_{\mathrm{LF}}-\mathbf{1}\mu_{\mathrm{LF}}\right)^{\mathrm{T}}\Psi_{\mathrm{LF}}\left(\mathbf{z}_{\mathrm{LF}},\ \mathbf{z}_{\mathrm{LF}}\right)^{-1}\left(\mathbf{y}_{\mathrm{LF}}-\mathbf{1}\mu_{\mathrm{LF}}\right)}{2\sigma_{\mathrm{LF}}^{2}} \tag{2.32}$$

By setting the derivatives of Eq. 2.32 with respect to μ_{LF} concerning and σ_{LF}^{2} to 0 and solving, find MLEs of

$$\hat{\mu}_{\mathrm{LF}}=\mathbf{1}^{\mathrm{T}}\,\Psi_{\mathrm{LF}}\left(\mathbf{z}_{\mathrm{LF}},\ \mathbf{z}_{\mathrm{LF}}\right)^{-1}\mathbf{y}_{\mathrm{LF}}\Big/\mathbf{1}^{\mathrm{T}}\Psi_{\mathrm{LF}}\left(\mathbf{z}_{\mathrm{LF}},\ \mathbf{z}_{\mathrm{LF}}\right)^{-1}\mathbf{1}\quad\text{and} \tag{2.33}$$

$$\hat{\sigma}_{\mathrm{LF}}^{2}=\left(\mathbf{y}_{\mathrm{LF}}-\mathbf{1}\hat{\mu}_{\mathrm{LF}}\right)^{\mathrm{T}}\Psi_{\mathrm{LF}}\left(\mathbf{z}_{\mathrm{LF}},\ \mathbf{z}_{\mathrm{LF}}\right)^{-1}\left(\mathbf{y}_{\mathrm{LF}}-\mathbf{1}\hat{\mu}_{\mathrm{LF}}\right)\big/n_{\mathrm{LF}} \tag{2.34}$$

Substituting Eqs. 2.33 and 2.34 back into ln-likelihood, yields the concentrated ln-likelihood function:

$$\mathrm{MLE}_{\mathrm{LF}}=-\frac{n_{\mathrm{LF}}}{2}\ln\left(\hat{\sigma}_{\mathrm{LF}}^{2}\right)-\frac{1}{2}\ln\left|\det\left(\Psi_{\mathrm{LF}}\left(\mathbf{z}_{\mathrm{LF}},\ \mathbf{z}_{\mathrm{LF}}\right)\right)\right|. \tag{2.35}$$

Since $\hat{\boldsymbol{\theta}}_{\mathrm{LF}}$ and $\hat{\mathbf{p}}_{\mathrm{LF}}$ cannot be obtained theoretically, they can only be obtained by numerical maximization of Eq. 2.35. Single-objective Genetic Algorithm (GA) [28] is applied in this study. This Gaussian process construction to approximate the $Z_{\mathrm{LF}}(\cdot)$ with the LF sample data in the BF-Kriging model is similar to a single-fidelity Kriging surrogate model. To find $\mu_{\mathrm{d}}, \sigma_{\mathrm{d}}^{2}, \hat{\boldsymbol{\theta}}_{\mathrm{d}}, \hat{\mathbf{p}}_{\mathrm{d}}$, and $\rho,$ define:

$$\mathbf{d}=\mathbf{y}_{\mathrm{HF}}-\rho\mathbf{y}_{\mathrm{LF}}\left(\mathbf{x}_{\mathrm{HF}}\right) \tag{2.36}$$

where $\mathbf{y}_{\mathrm{LF}}(\mathbf{x}_{\mathrm{HF}})$ are the responses of the LF model at the HF locations. If $\mathbf{y}_{\mathrm{LF}}$ is not available at $\mathbf{x}_{\mathrm{HF}}$, it is possible to estimate ρ at little additional cost by using Kriging estimates $\hat{\mathbf{y}}_{\mathrm{LF}}\left(\mathbf{x}_{\mathrm{HF}}\right)$, using the determined hyper-parameters $\hat{\boldsymbol{\theta}}_{\mathrm{LF}}$ and $\hat{\mathbf{p}}_{\mathrm{LF}}$. The ln-likelihood of $\mathbf{d}$, given $\mathbf{y}_{\mathrm{LF}}$, is:

$$\text{MLE}_d = -\frac{n_{\text{HF}}}{2}\ln\left(\hat{\sigma}_d^2\right) - \frac{1}{2}\ln\left|\det\left(\;_d\left(\mathbf{z}_{\text{LF}},\;\mathbf{z}_{\text{LF}}\right)\right)\right| - \frac{\left(\mathbf{d}-\mathbf{1}\mu_d\right)^{\text{T}}\;_d\left(\mathbf{z}_{\text{HF}},\;\mathbf{z}_{\text{HF}}\right)^{-1}\left(\;-\mathbf{f}\,\mu_d\right)}{2\sigma_d^2} \tag{2.37}$$

yielding MLE$_d$ of

$$\hat{\mu}_d = \mathbf{1}^{\text{T}}\;_d\left(\mathbf{z}_{\text{HF}},\;\mathbf{z}_{\text{HF}}\right)^{-1}\;\Big/\;\mathbf{f}^T\;_d\left(\;_{\text{HF}},\;\mathbf{z}_{\text{HF}}\right)^{-1}\mathbf{z},\;\text{and} \tag{2.38}$$

$$\hat{\sigma}_d^2 = \left(\mathbf{d}-\mathbf{1}\hat{\mu}_d\right)^{\text{T}}\;_d\left(\mathbf{z}_{\text{HF}},\;\mathbf{z}_{\text{HF}}\right)^{-1}\left(\;-\mathbf{f}\,\hat{\mu}_d\right)\big/ n_{\text{HF}} \tag{2.39}$$

with $\hat{\;}_d$, $\hat{\mathbf{p}}_d$ and $\hat{\rho}$ found by maximizing:

$$\text{MLE}_d = -\frac{n_{\text{HF}}}{2}\ln\left(\hat{\sigma}_d^2\right) - \frac{1}{2}\ln\left|\det\left(\;_d\left(\mathbf{z}_{\text{HF}},\;\mathbf{z}_{\text{HF}}\right)\right)\right| \tag{2.40}$$

Likewise, the estimation of $\hat{\;}_d$, $\hat{\mathbf{p}}_d$ and ρ is also obtained by single-objective GA maximizing the MLE$_d$.

Similar to the approach in Kriging, the prediction of a new HF point should be consistent with the observed data and the correlation parameters. The new prediction is augmented with the existing data, and the likelihood of this augmented data is maximized to obtain the MLE for the prediction $\hat{y}_{\text{HF}}\left(\mathbf{x}^{(n+1)}\right)$. The new dataset is defined as $\mathbf{X} = \left[\mathbf{x}_{\text{LF}}^T,\mathbf{x}_{\text{HF}}^T,\mathbf{x}^{(n+1)T}\right]^T$ with values of $\mathbf{y} = \left[\mathbf{y}_{\text{LF}}^T,\mathbf{y}_{\text{HF}}^T,\hat{\mathbf{y}}_{\text{HF}}^T\left(\mathbf{x}^{(n+1)}\right)\right]^T$, and the correlation vector that defines the correlation between the observed data and the new prediction is:

$$\mathbf{c} = \begin{bmatrix} \hat{\rho}\hat{\sigma}_{\text{LF}}^2\psi_{\text{LF}}\left(\mathbf{x}_{\text{LF}},\mathbf{x}^{(n+1)}\right) \\ \hat{\rho}\hat{\sigma}_{\text{LF}}^2\psi_{\text{LF}}\left(\mathbf{x}_{\text{HF}},\mathbf{x}^{(n+1)}\right)+\hat{\sigma}_d^2\psi_d\left(\mathbf{x}_{\text{HF}},\mathbf{x}^{(n+1)}\right) \end{bmatrix}. \tag{2.41}$$

The augmented covariance matrix is:

$$\tilde{\mathbf{C}} = \begin{bmatrix} \mathbf{C} & \mathbf{c} \\ \mathbf{c}^T & \rho^2\sigma_{\text{LF}}^2+\sigma_d^2 \end{bmatrix} \tag{2.42}$$

Similar to the single-fidelity Kriging, to maximize $MLE\;\hat{y}_{\text{HF}}\left(\mathbf{x}^{(n_{\text{HF}}+1)}\right)$, it is necessary to maximize:

$$\max\; f = -\frac{1}{2}\left(\tilde{\mathbf{y}}-\mathbf{1}\mu\right)^T\tilde{\mathbf{C}}^{-1}\left(\tilde{\mathbf{y}}-\mathbf{1}\mu\right), \tag{2.43}$$

which may be expressed as:

$$\max\; f = -\frac{1}{2}\begin{pmatrix} \mathbf{y}-\mathbf{1}\hat{\mu} \\ \hat{y}_{\text{HF}}\left(\mathbf{x}^{(n_{\text{HF}}+1)}\right)-\hat{\mu} \end{pmatrix}^T\begin{bmatrix} \mathbf{C} & \mathbf{c} \\ \mathbf{c}^T & \rho^2\hat{\sigma}_{\text{LF}}^2+\hat{\sigma}_d^2 \end{bmatrix}^{-1}\begin{pmatrix} \mathbf{y}-\mathbf{1}\hat{\mu} \\ \hat{y}_{\text{HF}}\left(\mathbf{x}^{(n_{\text{HF}}+1)}\right)-\hat{\mu} \end{pmatrix} \tag{2.44}$$

Ignoring terms without $\hat{y}_{\mathrm{HF}}\left(\mathbf{x}^{(n_{\mathrm{HF}}+1)}\right)$, and set it to 0,

$$\left(\frac{-1}{\rho^2\hat{\sigma}_{\mathrm{LF}}^2 + \hat{\sigma}_d^2 - \mathbf{c}^T\mathbf{C}^{-1}\mathbf{c}}\right)\left(\hat{y}_{\mathrm{HF}}\left(\mathbf{x}^{(n_{\mathrm{HF}}+1)}\right) - \hat{\mu}\right) + \frac{\mathbf{c}^T\mathbf{C}^{-1}(\mathbf{y} - \mathbf{1}\hat{\mu})}{\rho^2\hat{\sigma}_{\mathrm{LF}}^2 + \hat{\sigma}_d^2 - \mathbf{c}^T\mathbf{C}^{-1}\mathbf{c}} = 0 \tag{2.45}$$

Solving for $\hat{y}_{\mathrm{HF}}\left(\mathbf{x}^{(n_{\mathrm{HF}}+1)}\right)$:

$$\hat{y}_{\mathrm{HF}}\left(x^{(n_{\mathrm{HF}}+1)}\right) = \frac{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{y}}{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}} + \mathbf{c}^T\mathbf{C}^{-1}(\mathbf{y} - \mathbf{1}\hat{\mu}). \tag{2.46}$$

By using the standard stochastic-process approach can also derived the mean-squared error $\hat{s}^2(\mathbf{x})$ of this predictor:

$$\hat{s}^2(\mathbf{x}) \approx \rho^2\hat{\sigma}_{\mathrm{LF}}^2 + \hat{\sigma}_d^2 - \mathbf{c}^T\mathbf{C}^{-1}\mathbf{c} + \frac{\left(1 - \mathbf{1}^T\mathbf{C}^{-1}\mathbf{c}\right)^2}{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}}.$$

2.5 RESULTS AND DISCUSSION

In the present section, a representative component of marine structures is presented in Section 2 to validate the proposed BF-Kriging surrogate model. Specifically, the first analysis investigated the correlation between the HF and three different LF models. Then, the effectiveness of the number of $D_{tr,\mathrm{HF}}$, and $D_{tr,\mathrm{LF}}$ training points in the BF-Kriging model is explored. After that, the success of the proposed method is examined to assess the global σ_{zu} approximation considering $D_{tr,\mathrm{HF}}$, and $D_{tr,\mathrm{LF}}$ points centered at the mean value μ^* and the reliability of a stiffened panel under axial compression centered at the FORM design point x^*.

2.5.1 The mesh fidelity selection and validation of the FEM

The influence of mesh must be studied and refined enough to provide accurate and valid results. However, an equilibrium between the precision required and computational effort is needed to save computing time. The discretization level of FE models with HF and LF is determined through a mesh verification test considering the dimensions to be discretized by the FE Method. Consider N_x the number of divisions along the plate x-axis, N_y the number of divisions in the stiffener web (y-axis), N_z the number of divisions along the plate and stiffener z-axis, and N_{ts} the number of divisions along the flange stiffener x-axis. In the present study, $N_x = N_y$ is adopted because they are similar in size, and for all analyses, $N_{ts}=2$ is adopted.

The plate with the mean value of the material and geometric variables (μ^*) of the proposed model under axial compression presented in Table 2.1 is used for a mesh

convergence study. After several combinations of mesh size, one combination with $N_x=N_y=8$ and $N_z=80$ is defined as the HF model, depicted in Fig. 2.4(a). Moreover, three sets of division numbers are selected as LF models to evaluate the fidelities correlation presented in Table 2.2.

(a) HF Mesh

(b) LF Mesh
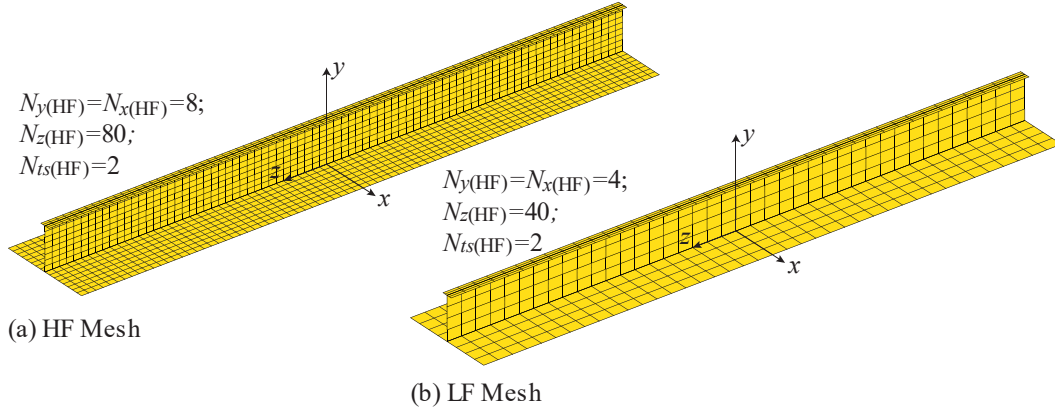
Figure 2.4 – Mesh (a) divisions, (b) HFM mesh size, and (c) LFM mesh size.

Table 2.2 – Fidelities correlation and cost functions.

Model Fidelity	$N_x=N_y$	N_z	Comp. Cost (s)	Ratio to HF	DIC
Low 01	2	20	16.200	0.2364	0.975
Low 02	3	30	17.300	0.252	0.983
Low 03	4	40	18.200	0.265	0.991
High	8	80	68.500	1.000	-

The computational cost in the stiffened panels analysis is the sum of the time for geometry and mesh creation, mesh modification, and non-linear analysis. The construction of the geometry and mesh presents similar computational costs for all the presented models. However, the mesh is modified, applying the initial imperfections at every point. This task that the more refined the mesh is, the greater the computational time devoted to it. Unlike the HF model, the fewer nodes for the Low 01, Low 02, and Low 03 models did not present a significant difference in the computational time of the mesh modification. In addition, the computational time of the non-linear analysis is also increased with the refinement of the mesh.

To evaluate the relationship of the mesh fidelities model, a test dataset D_{test} with $n_{test}=200$ is used to determine the DIC correlation parameter. The computational cost presented in Table 2.2 is the average time consumed by the D_{test} to obtain the σ_{zu} output. Seemingly the computational costs for the LF models are close, showing that the mesh size does not significantly influence the computational time of LF models. Still considering the information in Table 2.2, the LF models presented larger values of DIC, indicating a stronger relationship between each LF about the HF model. Based on these presented results, the LF

Model 03, depicted in Fig. 2.4(b), with $N_x=N_y=4$ and $N_z=40$, is adopted as a reference LF model in the present study. Fig. 2.5(a) shows the D_{test} points, with the HF and LF model correlation.

(a)

(b)
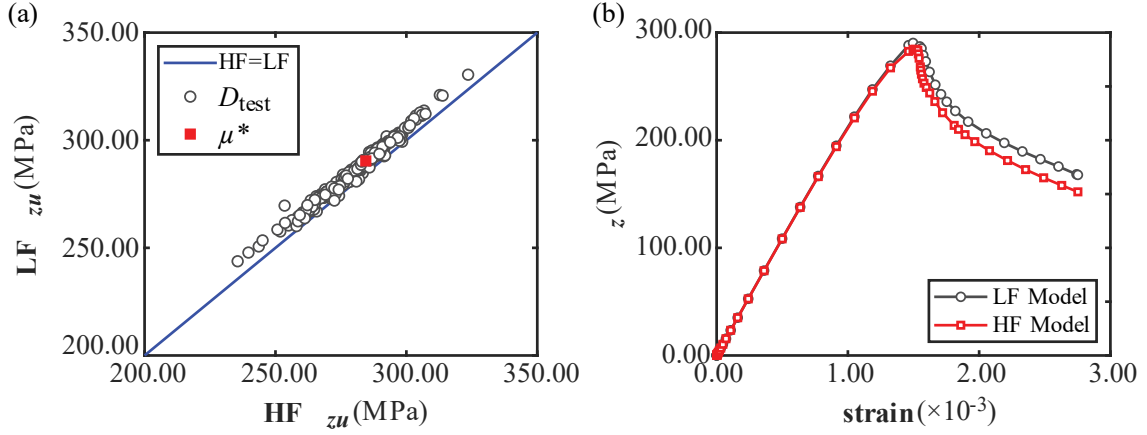
Figure 2.5 – FE results: (a) HF and (b) LF deflection shape and stress distribution, (b) stress-strain.

Figure 2.5 shows the FE results for the HF and LF analysis assuming the mean value of the basic random variables (red filled square), comparing with the D_{test} dataset (Fig. 2.5(a)), and the stress-strain considering the HF and LF model (Fig. 2.5(b)). The σ_{zu}=284.260 MPa considering the HF model, and σ_{zu}=290.39 MPa considering the LF model. The HF result obtained by the proposed FEA model agrees well with the ones obtained by Gaspar *et al.* [3].

2.5.2   Assessment of BF-Kriging models efficiency in global σ_{zu} predictions

The predictive accuracy of the failure probability predictions depends on the σ_{zu} prediction of the stiffened panel elements. This accurate prediction depends on the number of surrogate training points and the hyperparameters of the BF-Kriging model. Therefore, to obtain an accurate response surface BF approximation for this example, it is necessary to evaluate the $D_{tr,HF}$, and $D_{tr,LF}$ dataset sizes n_{HF} and n_{LF}. The n_{LF} value is determined by multiplying n_{HF} by a sample size ratio between HF and LF models (λ_{BF}) to obtain the surrogate improvement level with potential LF sample points. The potential and impact of the LF sample points were evaluated with the HF samples fixed to n_{HF}=500 and 1000 training points, while the infill strategy of LF samples is increased as the ratio λ_{BF} ranges from 3 to 8.

The total cost is the number of calls to the performance function N_{call} that, based on the cost functions, equals the equivalent number of HF samples. The total computational cost for each studied case is depicted in Fig. 2.6(a). By plotting PF N_{call} in terms of the λ_{BF} evolution, it can be seen that for each value of n_{HF}, the computational cost and N_{call} increase in

the associated linear rate. For each value of n_{HF} and comparing the extreme cases of $\lambda_{BF}=3$ and 8, N_{call} has a similar increase of 73.9%.
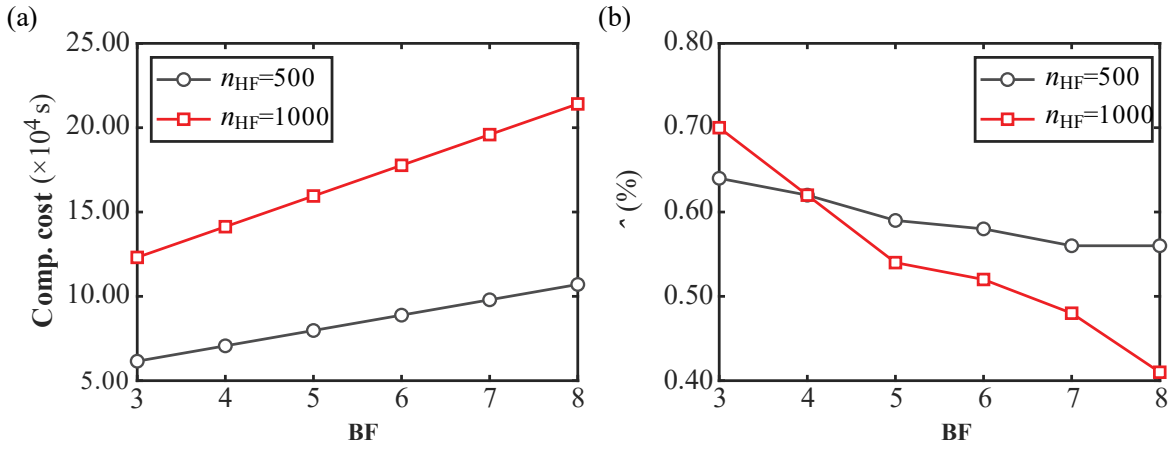
Figure 2.6 – The tendency of the global: (a) Computational cost, and (b) relative error of the BF-Kriging surrogate model.

For each fitted BF-Kriging surrogate model under different sample ratios λ_{BF}, the MRE accuracy metrics are obtained predicting the same D_{test} with $n_{test}=200$ used in the model's correlation analysis. The $\text{MRE}_{\hat{\sigma}}$ accuracy assessment is presented in Fig. 2.6(b). In both analyses for the values of n_{HF}, when $\lambda_{BF}\leq5$, as the number of LF samples increases, the prediction accuracy of the BF surrogate models improves rapidly. However, when $\lambda_{BF}>5$, as the number of LF samples increases, the models with $n_{HF}=500$ slightly improve their fit after that little to no variation in their metrics. The BF models with $n_{HF}=500$ provide the best prediction accuracy when $\lambda_{BF}=8$. This scenario affords similar accuracy metrics compared to the cases with $n_{HF}=1000$ and low λ_{BF} values. Otherwise, in the model with $n_{HF}=1000$, the increment of λ_{BF} guarantees a continuous improvement in the model, accordingly, sufficient accuracy to the validation data.

Considering the presented accuracy results, the case with $n_{HF}=1000$ and $\lambda_{BF}=8$ provides accurate metrics estimates and can be viewed as an appropriate BF-Kriging surrogate model to predict the ultimate compressive stress σ_{zu} of stiffened panels. Figure 2.7 depicts a scatter plot of σ_{zu} predicted by the FEA model and the BF-Kriging model at the D_{test} dataset showing the accuracy of the BF-Kriging model approximation.

Figure 2.7(b) illustrates the relative errors for the D_{test} dataset. The selected BF-Kriging surrogate model has a mean absolute error $\varepsilon_{\hat{\sigma}}=0.41\%$, and a coefficient of determination $R^2=0.991$ that implies that this BF-Kriging surrogate model ensures sufficient accuracy in the validation data. Figure 2.7(b) shows that for 95% of the verification points,

the squared error satisfies $\varepsilon_{\hat{\sigma}} \leq 1.0\%$. Beyond that, σ_{zu}=284.402 MPa for the mean value μ^* of the basic random variables (Table 2.1) obtained by the BF-Kriging model with an $\varepsilon_{\hat{\sigma}} = 0.05\%$ when compared with the σ_{zu} obtained by FEA and described in Section 5.1.
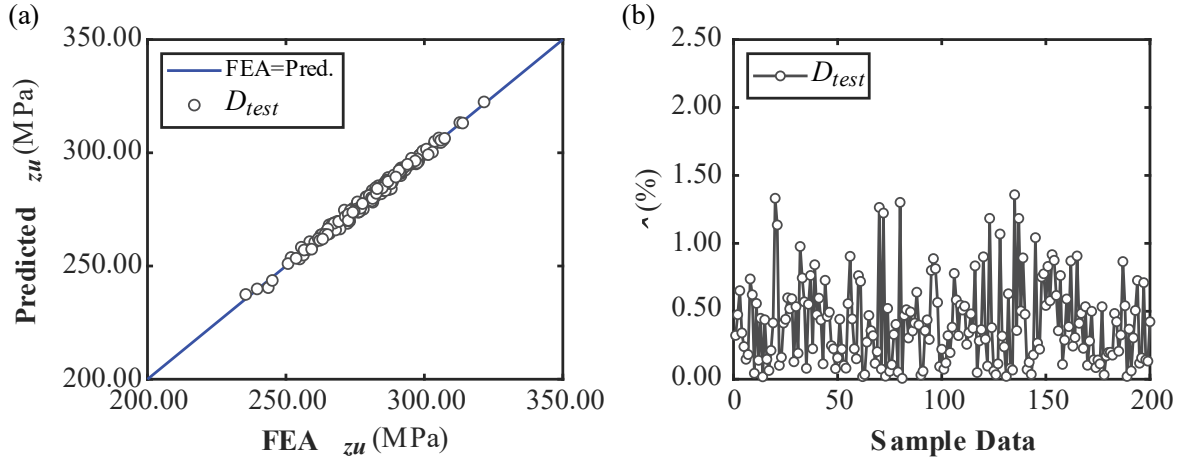
Figure 2.7 – BF-Kriging results for the σ_{zu} of the stiffened plate: (a) predictions compared with the FEA results, and (b) mean relative errors.

2.5.2.1 Sensitivity analysis

It must be noted that the hyperparameter θ_l in the BF-Kriging model can reflect the influence of the lth design variable on the output in a way. Generally speaking, a larger θ_l indicates that the output varies more dramatically in the lth coordinate direction [28, 77]. Based on the HF and LF sets, a sensitivity analysis is first conducted by evaluating the rank-correlation coefficients for the response variable concerning the random variables in θ. The resulting tornado plots are shown in Fig. 2.8 for both LF and HF models.
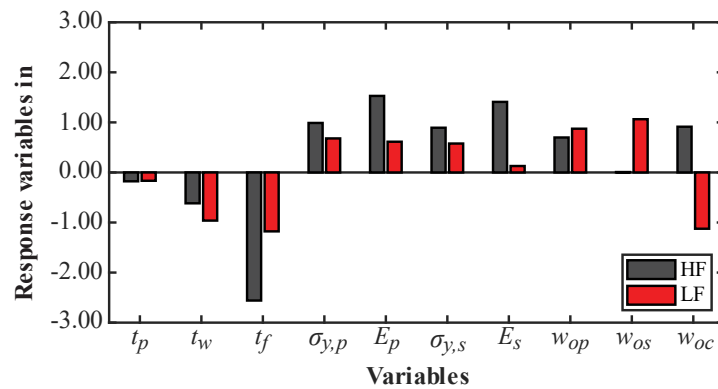
Figure 2.8 – Rank-correlation coefficients for the input variables.

For the global σ_{zu} prediction, the variables do not have broadly similar importance on the problem and differ significantly in nature and significance. The HF and LF models generally agree in their sensitivity indications. For the HF and LF models, the random

variables characteristic of the material $\sigma_{y,p}$, $\sigma_{y,s}$, E_p, and E_s are sensitive to the σ_{zu}. In contrast, the stiffened panel geometry variables t_p, t_w, and t_f show low sensitivity to the response variable in both LF and HF models.

There is an apparent divergence between the HF and LF model sensitivities for initial imperfection amplitudes w_{os}, and w_{oc} showing inconsistency in sensitivity between the models. Some divergence in the sensitivity can occur between multi-fidelity models due to the different nature of their states (e.g., 2D vs. 3D models, models with different domains, coarse vs. refined meshes). In this paper's specific case, this behaviour is directly linked to the initial deformation that varies with the increase of the FE mesh. An infill strategy can be used in cases where there is no agreement on the sensitivity of the variables between the two models sampling HF data used for the HF model or to ensure that the LF model is globally accurate. However, in the present study, despite the divergence of sensitivities predominantly in the initial imperfections, the HF and LF models resulted in acceptable global accuracy, showing that both are globally accurate.

2.5.3   Reliability assessment considering BF-Kriging model

The assessment of efficiency, accuracy, and fidelity of the BF-Kriging surrogate model in the shipbuilding case presented in this paper is also performed through a systematic comparison of their accuracy in the approximation of the failure of probability and reliability index of the application problem based on SUS and FORM. The FEA is used to obtain the implicit performance function (FE-PF). For the FEA+SUS analysis, the initial N_n is set to 3×10^3 and $p_0 = 0.1$, also of the other intermediate failure events size. For the FEA+FORM, the finite difference method determines the partial derivates in Eq. 2.12 for probability prediction. The results of the reliability analyses reference values $P_{f\text{-}ref}$ and β_{ref} are summarized in Table 2.3. The N_{call} value is the equivalent number of HF FEA simulations to estimate the failure probability, and the total runs are the number of independent simulations to estimate the statistics of $\hat{P}_f$. The relative error for the β is $\varepsilon_{\hat{\beta}} = \left| \hat{\beta} - \beta_{ref} \right| / \beta_{ref}$ computed considering the $\hat{\beta}$ predicted by BF-Kriging with the reference β_{ref} obtained by FE-PF. The other results presented in Table 2.3 are discussed throughout this section.

Table 2.3 – FORM design points and sensitivity factors for stiffened panel.

Method	N_{call} (runs)	$\hat{P}_f$	$\hat{\beta}$	$MRE_{\hat{\beta}}$ (%)
FEA+SUS (ref.)	15000	1.48×10^{-4}	3.618	-
FEA +FORM	1350	9.33×10^{-5}	3.736	3.26
Kriging+SUS	3000 (50)	1.31×10^{-4}	3.640	0.89
BF-Kriging+FORM	3120 (50)	1.04×10^{-4}	3.708	2.49
BF-Kriging+SUS (proposed)	3120 (50)	1.32×10^{-4}	3.650	0.88

After each BF-Kriging tuned hyperparameter and successfully fitted considering different n_{HF} and λ_{BF}, the BF-Kriging model is associated with SUS to predict $\hat{P}_f$. As presented in Step 2 of Fig. 2.3, the values of n_{HF} and λ_{BF} are also increased if $\varepsilon_{\hat{\beta}} \geq 1.0\%$. The initial N_n is set to 3×10^3 and $p_0 = 0.1$, also of the other intermediate failure events size. In this analysis, the outputs are obtained by the BF-Kriging surrogate model. The predicted $\hat{\beta}$ using BF-Kriging+SUS are the continuous lines presented in Fig. 2.9(a) w.r.t. the sample ratio λ_{BF}. Each point is presented with the 95% confidence interval (CI) amplitude of $\hat{\beta}$. The relative error $\varepsilon_{\hat{\beta}}$ is the continuous lines presented in Fig. 2.9(b), comping the $\hat{\beta}$ predicted by BF-Kriging+SUS with the reference β_{ref} obtained by FEA+SUS.
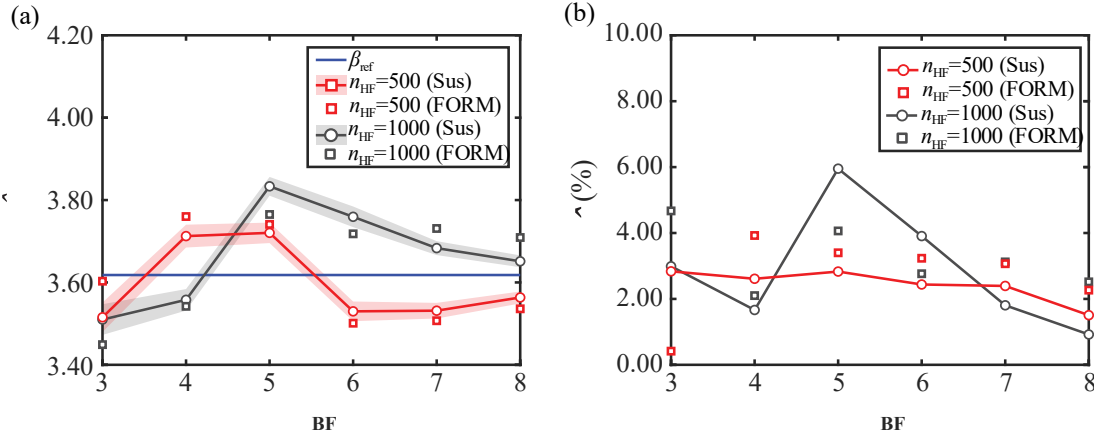
Figure 2.9 – Reliability index: (a) predicted β and, (b) β relative error.

As displayed in Fig. 2.9, CI becomes narrower as the number of λ_{BF} increases due to the effect of reducing the standard error of the probability of failure estimations by BF-Kriging+SUS. Considering the different values of n_{HF}, these BF-Kriging models give bad results for a small number of samples and show fluctuation as the number of LF samples increases in $\hat{\beta}$ prediction in the cases with $\lambda_{BF} < 7$. The models provide less accurate local predictions close to the limit state region for these cases. The results curve of the $\hat{\beta}$ predicted

by BF-Kriging+SUS shows little to no fluctuation in the cases with $\lambda_{BF} \geq 7$, providing relative values compared with the β results obtained by FEA. However, it is essential to emphasize that although the BF-Kriging model with $n_{HF}=500$ and $\lambda_{BF} \geq 7$ presents a good local prediction, ensuring good accuracies in the $\hat{\beta}$ predicted by SUS, it does not ensure satisfactory global accuracy results, as shown in Fig. 2.6(b). For the best circumstances with $n_{HF}=500$, the most accurate reliability index is obtained with $\lambda_{BF}=8$, which shows $\varepsilon_{\hat{\beta}} = 1.73\%$ when compared with β_{ref}. When $n_{HF}=1000$ is considered, the most accurate reliability estimate is obtained with $\lambda_{BF}=8$, which shows a small variability of $\varepsilon_{\hat{\beta}} = 0.88\%$ about β_{ref}.

From Fig. 2.9, with the rising of n_{LF} and λ_{BF}, a better approximation is obtained for the reliability metrics of $\hat{P}_f$ and $\hat{\beta}$ in the BF model compared to that obtained through the FEA. Finally, through the results presented, the chosen convergence criterion accepts the BF model with $n_{HF}=1000$ and $\lambda_{BF}=8$. This model resulted in a $\hat{\beta}=3.65$ and $\hat{P}_f=1.32\times10^{-4}$.

Furthermore, BF-Kriging+FORM is also performed to predict $\hat{P}_f$ and $\hat{\beta}$. The performance assessment is presented as squares scattered in Fig. 2.9. As depicted in Fig. 2.9b, for the most cases evaluated, the MRE_β are higher for most of the evaluated cases, showing that for this study, BF-Kriging+SUS presented a better prediction of $\hat{\beta}$ compared to the reference value. However, the BF-Kriging+FORM performance assessment is appropriate to evaluate the corresponding design point coordinates of the random variables (i.e., the u^* in $G(\mathbf{x}^*)=0$ even the sensitivity factors (α), based on Eq. 2.12. Also, the value for each variable is conceivably compared about the obtained by FE-FORM, given in Table 2.4.

Analyzing the BF-Kriging results, compared with the reference design point obtained by FE-FORM, the design point is similar with a small error. Unlike the Sensitivity analysis in Fig. 2.8, which evaluates the influence of variables in the global prediction of σ_{zu}, the sensitivity factors (α) introduced in Table 2.4 show the effect of each variable in the $\hat{P}_f$ determination. The ultimate capacity model factor χ_u *is the most* decisive variable in the $\hat{P}_f$ prediction. The basic random variables with a positive contribution to the reliability index are the geometry variables (t_p, t_w, and t_f), the yield stress $\sigma_{y,p}$ and Young's modulus E_p of the plate, and the yield stress $\sigma_{y,s}$ of the stiffeners. The young modulus E_s of the stiffeners, the weld-induced initial imperfections (w_{op}, w_{os}, and w_{oc}), and the load variables (M_{sw} and M_{wv}) contribute negatively to the $\hat{P}_f$ prediction, as shown in Table 2.4.

Table 2.4 – FORM design points and sensitivity factors for stiffened panel.

Variable	FE-PF			BF-Kriging		
	u^*	x^*	α	u^*	x^*	α
t_p	-0.156	20.436	0.041	-0.162	20.433	0.043
t_w	-0.066	9.487	0.0125	-0.045	9.491	0.0122
t_f	-0.025	13.992	0.0045	0.002	14.000	0.004
$\sigma_{y,p}$	-0.024	346.862	0.0075	0.119	349.878	0.008
E_p	-1.115	1.923e5	0.286	-0.833	1.956e5	0.293
$\sigma_{y,s}$	-1.061	325.965	0.190	-1.292	321.480	0.187
E_s	0.583	2.129e5	-0.162	0.635	2.136e5	-0.160
w_{op}	0.793	2.992	-0.259	0.750	2.932	-0.254
w_{os}	0.982	3.983	-0.304	0.912	3.853	-0.302
w_{oc}	0.074	2.593	-0.029	0.003	2.508	-0.024
M_{sw}	1.077	52.234	-0.281	1.130	52.843	-0.285
M_{wv}	1.825	147.534	-0.502	1.882	148.678	-0.493
χ_u	-2.263	0.823	0.601	-2.544	0.795	0.607

The results of the proposed BF-Kriging+SUS and other methods are shown in Table 2.3. The reduction of computing time is remarkable considering the proposed BF-Kriging+SUS and compared with the FEA-SUS. The proposed method reduces the N_{call} by about 80%, getting a similar value of $\hat{P}_f$. Comparing BF-Kriging+SUS with the FEA-FORM, the proposed method needs more N_{call} to fit the PF. However, considering the PF is highly non-linear, the proposed method ensures a more reliable $\hat{P}_f$ prediction value. The BF-Kriging+SUS is also compared with the single-fidelity Kriging surrogate model [77], previously adopted to predict σ_{zu} of stiffened panels (e.g. [8, 25]). The proposed method has similar $\hat{P}_f$ prediction value and costs similar N_{call} necessary to fit the surrogate accurately.

These results mean that for practical applications, a dataset using around 1000 HF and 8000 LF FEAs can be an initial estimative to accurately fit a BF-Kriging model to make reliability predictions for stiffened panels with similar geometry, materials, and boundary conditions. Table 2.3 demonstrates that the BF-Kriging+SUS reliability approach has advantages over the SUS with FEA-based PF at identical computational effort compared with classic surrogate models. In addition, comparing accuracy, the proposed BF-Kriging+SUS reliability approach has advantages over the FORM with FEA-based PF, considering highly non-linear PF. Furthermore, in highly non-linear applications, the non-linear FEA numerical algorithm may produce noisy derivatives due to oscillations of the LSF for slight variations of

the input parameters, which can affect the FORM convergence. In these cases, the BF-Kriging+SUS also has the advantage of obtaining a reliable approximated PF.

2.6   CONCLUSIONS

In this paper, a Bi-Fidelity Kriging surrogate model is proposed to become explicit a 13-*d* highly non-linear stiffened panel performance function to facilitate the reliability analysis of the presented structure.

The distance correlation, frequently used for non-linear variables correlation, is innovatively imported for constructing spatial correlation functions to estimate the relationship between the high and low-fidelity models. Estimating the correlation between the fidelities reduces the complexity of the high and low-fidelity models' choice. After the low-fidelity model selection, the reduction in the computational time provided by the proposed BF-Kriging has been up to 75% of the computational cost in the FE analysis.

Moreover, the present paper shows the influence of the number of high and low-fidelity data points in the BF-Kriging accuracy for the case study. A global prediction system for the ultimate strength is developed, and the BF-Kriging model provides immediate solutions to the Finite Element. Moreover, the BF reliability analyses using Subset Simulation present accurate results, showing that the BF model allows an excellent local prediction under the given performance function.

Also, the importance of each design variable is evaluated in two different aspects. Firstly, it concerned the ultimate strength global output estimated through the corresponding BF-Kriging. After that, for reliability index prediction using Subset Simulation and FORM.

These results suggest that the BF framework can be applied to the reliability analysis of structures, using HF associated with low-fidelity samples to predict high-fidelity values. This framework provides an acceptable level of accuracy and considerable computation time savings compared to the conventional methods that use only high-fidelity samples.

3 HYPERPARAMETER-OPTIMIZED BI-FIDELITY DEEP NEURAL NETWORK MODEL ASSOCIATED WITH SUBSET SIMULATION FOR STRUCTURAL RELIABILITY ANALYSIS

Authors: João Paulo Silva Lima, Francisco Evangelista Jr., Carlos Guedes Soares

Abstract: The present study proposes a two-stage Bi-Fidelity Deep Neural Network surrogate model to quantify the uncertainty of structural analysis using low-fidelity data samples added to the model to predict high-fidelity responses. This multi-fidelity surrogate model efficiently reduces the high computational cost for highly non-linear and high dimensional structural reliability problems. The framework is demonstrated using three different representative examples. First, it demonstrates the multi-fidelity model's accuracy for approximating a high non-linear 20-dimensional standard benchmark function that is hard to approximate with other methods and compared with another multi-fidelity neural network framework. In the other examples, the multi-fidelity framework is associated with Subset Simulation to efficiently estimate rare events considering a benchmark case study including high-dimensional scenarios, and the third example considers finite element model case studies. The proposed framework is also compared with a multi-fidelity Co-Kriging method. The results show that the proposed multi-fidelity framework, with its optimized hyperparameters using Bayesian Optimization, is an excellent strategy for reducing the number of samples used to construct the performance function's surrogate model. Moreover, the proposed framework can provide an accurate failure probability estimation with a lower computational cost in high non-linear, high dimensional, and rare events.

Keywords: Multi-fidelity; Artificial Neural Networks; Structural reliability analysis; Non-linear Finite Element Analysis; Stiffened panel.

3.1 INTRODUCTION

Reliability analysis is of significant importance in structural engineering, as it aims to obtain the failure probability of a system of interest under uncertainties that influence some event concerning some performance criterion [1-3, 7]. Engineers and researchers have well-recognized the importance of reliability in the past few decades. Its improvements have concentrated on making the performance functions more realistic, using proper strength and loading assessment methods based on numerical methods [4]. In general, analytical

formulations of the performance function (PF) were substituted by finite element assessments. The solution of the failure probability integral with accuracy in problems involving non-linear limit state functions is computationally demanding [3]. Furthermore, the limit state functions became implicit for most practical engineering problems, requiring different approaches to render them explicit and to calculate the reliability index [5, 6].

Many structural reliability problems can be computed efficiently using approximate methods such as the first-order or second-order reliability methods (FORM/SORM) (e.g. [7-9]) but for very small failure probabilities and complex failure boundaries use can be made of advanced simulation methods based on Monte Carlo simulation (MCS) with variance reduction techniques as MCS method with importance sampling (MCS-IS) (e.g. [10-12]), the directional IS (e.g. [11]), and the Subset Simulation (SUS) (e.g. [13-17]).

Surrogate models have been widely applied in designs to replace physical experiments or expensive simulations, reproducing the PF function based on limited expensive calculations to obtain samples to reduce computational burden [8, 22, 85]. Initial formulations using polynomial fits [39, 40] have been done to identify the best types of polynomials to adopt, how many sample points to use in the fitting, and the choice of the initial set of points to be fitted by the surfaces. Methods based on quadratic response surface (e.g. [86]), quadratic polynomial–based moving least square (e.g. [87, 88]), and weight regression method (e.g. [89]) were applied to approximate the PF. Alternately to polynomial regression models, approaches based on the Kriging surrogate model have been proposed to approximate the PF [8, 30, 90, 91]. Other approaches adopted Radial Basis Functions [9, 22], Support Vector Machine (e.g. [92, 93]), Polynomial Chaos Expansion (e.g. [21, 94, 95]), and Gaussian Processes (e.g. [96-98]).

Artificial Neural Networks (ANN) are amongst the most popular Machine Learning (ML) methods used to generate the response surface in structural reliability problems. Chojaczyk *et al*. [2] reviewed and applied ANN models in the reliability analysis of steel structures used in the marine industry. Xu *et al.* [99] and Afshari *et al.* [100] summarize the main applications of ANN variations and ML models in structural reliability problems. Typically, an ANN regression model contains model parameters and hyperparameters. Selecting an optimal model that achieves the best performance on the data in a suitable amount of time becomes a problem of tuning or optimizing these hyperparameters [101-103].

The above surrogates are based on the fact that data for their construction are available from a single-fidelity source, often evaluated through fine high-fidelity (HF) simulations.

Indeed, the use of an HF simulation can often provide more reliable and accurate simulation results. Using low-fidelity (LF) models to estimate system performance involves significantly lower computational requirements using coarse discretization or simplified physics-based models. However, LF models may result in inaccurate metamodels or even distorted ones. Hence, Multi-fidelity (MF) surrogate approaches have attracted significant attention recently for data regression which maximizes the use of cheaper LF models to predict HF outputs [28]. The Bi-fidelity (BF) strategy uses the correlation between an HF model and one LF model.

The ability to handle large data sets and deal with arbitrary non-linearities in high dimensions have made ANNs a good candidate for MF regression. Aydin *et al.* [104] use a training strategy starting the ANN with LF computational models and switching to HF training data when the overall performance of the ANN stops increasing. Motamed *et al.* [105] and Liu and Wang [106] propose MF NN surrogate models for the uncertainty quantification of physical systems described by partial differential equations. Meng and Karniadakis [107] developed a 3-step MF Deep neural network (DNN). The first stage train LF data and is associated with two HF NNs to define linear and non-linear relationships between the LF and HF models. Guo *et al.* used the 3-step model to approximate input–output maps in problems governed by PDEs. Chen *et al.* [108] used the 3-step MF considering Convolutional Neural Networks (CNN), treating the multi-fidelity data as image data and processing them using CNN. Applications using the 3-step proposed by Meng and Karniadakis [107] in engineering problems can be found in [109-111].

The use of MF models has gained popularity in structural engineering applications such as multidisciplinary, robust, and multi-objective optimization fields [28, 29, 32, 56] and uncertainty quantification fields (e.g., [34, 55, 112]). The applications of MF models in reliability analysis have emerged recently. Zhang *et al.* [33] developed an adaptive reliability analysis for MF models using the Gaussian process and applied it to a transmission tower problem; however, this method is constrained in treating high-dimensional problems due to the general limitation of Gaussian process metamodels. Yi *et al.* [30] presented an adaptive MF kriging algorithm for structural reliability analysis using five examples with low-to-moderate dimensions. Skandalos et al. [51] used Co-Kriging as a multi-fidelity surrogate model for conducting seismic reliability analysis of base-isolated buildings. The Co-kriging method has two significant benefits. Firstly, the method can be cast in the Bayesian framework, and the regression results naturally include an uncertainty estimation, which is usually desirable. Moreover, it is suitable for many applications as a non-parametric

regression tool. Nevertheless, the Co-Kriging is unsuitable for high-dimensional drawbacks due to the curse of dimensionality.

Recently, MF frameworks have been proposed for highly non-linear reliability problems and rare event simulations. Dhulipala *et al.* [113] propose active learning based on adaptive Subset Simulation (SUS) for robustly predicting small failure probabilities using a Gaussian Process active learning function that decides when to call the HF model. In another study, Dhulipala *et al.* estimated the small failure probabilities of an advanced nuclear fuel using coupled active learning, MF modeling, and the SUS algorithm. Proppe and Kaupp [114] presented efficient MF estimators for rare events by combining multiplicative and additive information fusion with importance sampling.

The main goal of this paper is to associate MF modeling and the SUS algorithm to estimate the failure probabilities. The main contribution is developing a hyperparameters-optimized BF Deep Neural Network (BF-DNN) model for global prediction, allowing similar accuracy to the conventional HF modeling, and offering a considerable decrease in dataset size and computational time. Comparisons of the failures probability predictions accuracy based on SUS through the BF-DNN prediction models as surrogates for the FE-PF are presented. The proposed framework is demonstrated in two academic case studies and on reliability assessment of the ultimate compressive strength of plate elements typical of the deck structure of double-hull oil tanker ships described by a non-linear Finite Element (FE) Analysis. A comparison with another MF-DNN model and with the traditional strategy Co-Kriging is also presented. The main advantage of BF-DNN+SUS is expected to be computationally efficient and feasible for dealing with rare events with complex, highly non-linear implicit performance functions, which require time-consuming numerical procedures in high-dimensional practical engineering problems.

This paper is organized as follows: Section 2 introduces related concepts to elaborate on the proposed BF-DNN model and the methodology adopted to propose the reliability analysis; The application problem, considered as a numerical example, and the FE model for the present study is described in Section 3; In section 4, results obtained from the study are discussed extensively; Section 5 draws the main conclusions of the work conducted.

3.2 PROPOSED BI-FIDELITY DEEP NEURAL NETWORK WITH BAYESIAN OPTIMIZATION

This section outlines the computational framework of the Bi-fidelity Deep Neural Network (BF-DNN) to assess structures' uncertainty quantification and failure probability. This

network intends to employ a small amount of HF data and a large amount of LF data, both generated by FE analysis, to train the surrogate model.

3.2.1 A bi-fidelity surrogate model

The variable fidelity analysis is essential for tackling computational costs using a combination of LF and HF data. Although LF samples are noisy and somewhat inaccurate, they usually correlate strongly with HF samples [78]. As such, it may be possible to leverage this correlation to avoid total reliance on HF data correcting the HF model using LF information.

The combinations of fidelities can be categorized into three groups [115]: (*i*) Adaptation approaches that use adaptation to enhance LF models with information from HF models while the computation proceeds; (*ii*) Fusion strategies that evaluate LF models and HF models and then combine information from all outputs and (*iii*) Filtering approaches that use the HF model if the LF model is inaccurate, or when the candidate point meets some criterion.

Currently, there are several available correction methods, which can be briefly divided into three categories [116]: (*i*) Additive and multiplicative corrections that construct a surrogate model of the difference or the ratio between the HF and the LF models; (*ii*) Comprehensive corrections, where both corrections (additive and multiplicative) are used in the same bi-fidelity surrogate model; and (*iii*) Space mapping (input correction) where instead of correcting the output of the LF model, it is also possible to correct the input variables.

Suppose an *n*-dimensional random vector $\mathbf{x} \in \mathbb{R}^d$ is mapped through a model to obtain a desired output $y(\mathbf{x})$. Let $y_{\mathrm{LF}}(\mathbf{x})$ and $y_{\mathrm{HF}}(\mathbf{x})$ denote the approximated values of $y(\mathbf{x})$ by an LF and HF computational model, respectively. In the current study, a fusion MF with a comprehensive correction model adopted is expressed as follows:

$$y_{\mathrm{HF}}(\mathbf{x}) = \rho(\mathbf{x}) \cdot y_{\mathrm{LF}}(\mathbf{x}) + z(\mathbf{x}) \hspace{2cm} \text{Equation Chapter (Next) Section 1 (3.1)}$$

where $\rho(\mathbf{x})$ is a scaling factor that quantifies the correlation between [y_{HF}, y_{LF}], and $z(\mathbf{x})$ the corresponding noise. However, one of the main disadvantages of this scheme is that it can only handle linear correlation between two fidelity data. In contrast, many interesting cases follow non-linear relationships between LF and HF data [107]. A general correlation is considered the relation between the two models is expressed as:

$$y_{\mathrm{HF}}(\mathbf{x}) = F\big(\mathbf{x}, y_{\mathrm{LF}}(\mathbf{x})\big) \hspace{5cm} (3.2)$$

where $F(\cdot)$ is an unknown function that maps the LF data to the HF level that is based on the input data $\left\{ \left(\mathbf{x}_{\mathrm{HF}}^{(i)}, y_{\mathrm{LF}}\left(\mathbf{x}_{\mathrm{HF}}^{(i)} \right) \right) : 1 \leq i \leq n_{\mathrm{HF}} \right\}$, and the output data $\left\{ \mathbf{y}_{\mathrm{HF}}^{(i)} : 1 \leq i \leq n_{\mathrm{HF}} \right\}$.

3.2.1.1 The Fidelities correlation considering Distance Correlation (DIC)

Although LF samples are noisy and skewed, they usually correlate strongly with HF samples [78]. As such, leveraging this correlation to avoid total reliance on HF data is possible. The empirical Distance Correlation (DIC) was introduced by Székely *et al.* [81, 82] to measure associations between two random vectors, not necessarily in the same dimension. As an advantage over the classical Pearson product-moment correlation, the DIC has a significant benefit for characterizing a non-linear relationship of random vectors. Moreover, DIC satisfies $0 \le \mathrm{DIC}\left(y_{\mathrm{HF}}, y_{\mathrm{LF}}\right) \le 1$ and equals zero if independence holds. Given a sample $\left\{\left(y_{HF}^{(i)}, y_{LF}^{(i)}\right)\right\}_{i=1}^{N}$ of N observations of the joint random vector (y_{HF},y_{LF}) , is defined the double-centered distance matrices *A* and *B* as follows:

$$A_{i,j} = a_{i,j} - \frac{1}{N}\sum_{l=1}^{N} a_{il} - \frac{1}{N}\sum_{k=1}^{N} a_{kj} + \frac{1}{N^2}\sum_{k,l=1}^{N} a_{kl}, \text{ and} \tag{3.3}$$

$$B_{i,j} = b_{i,j} - \frac{1}{N}\sum_{l=1}^{N} b_{il} - \frac{1}{N}\sum_{k=1}^{N} b_{kj} + \frac{1}{N^2}\sum_{k,l=1}^{N} b_{kl}, \tag{3.4}$$

where $a_{ij} = \left\| y_{HF}^{(i)} - y_{HF}^{(j)} \right\|_p$ and $b_{ij} = \left\| y_{LF}^{(i)} - y_{LF}^{(j)} \right\|_q$. Then, the sample dcov is the square root of

$$\mathrm{dcov}^2 = \frac{1}{N^2}\sum_{i,j=1}^{N} A_{i,j} B_{i,j}. \tag{3.5}$$

Similarly, the sample distance correlation is the standardized sample covariance given by:

$$DIC^2\left(y_{\mathrm{LF}}, y_{\mathrm{HF}}\right) = \begin{cases} \dfrac{\mathrm{dcov}^2\left(y_{\mathrm{LF}}, y_{\mathrm{HF}}\right)}{\sqrt{\mathrm{dcov}^2\left(y_{\mathrm{LF}}, y_{\mathrm{LF}}\right)\mathrm{dcov}^2\left(y_{\mathrm{HF}}, y_{\mathrm{HF}}\right)}} & \text{if } \mathrm{dcov}^2\left(y_{\mathrm{LF}}, y_{\mathrm{LF}}\right)\mathrm{dcov}^2\left(y_{\mathrm{HF}}, y_{\mathrm{HF}}\right) > 0 \\ 0 & \text{if } \mathrm{dcov}^2\left(y_{\mathrm{LF}}, y_{\mathrm{LF}}\right)\mathrm{dcov}^2\left(y_{\mathrm{HF}}, y_{\mathrm{HF}}\right) = 0. \end{cases} \tag{3.6}$$

3.2.2 Artificial neural networks for bi-fidelity regression

To learn the relationship $F\left(\mathbf{x}, y_{\mathrm{LF}}(\mathbf{x})\right)$ in Eq. 3.2, an Artificial Neural Network (ANN) surrogate model is constructed. According to the literature, different ML methods have been applied based on their computational efficiency and accuracy advantages to model an unknown function or distribution $y(x)$. Fundamentally, the ANN architecture is specified by the number of layers, neurons, activation functions, and the training process. Before the learning process, it is essential to define a set of hyperparameter values that achieve the best performance on the data in a reasonable amount of time [100]. Hyperparameters differ from the internal model parameters, such as the neural network's weights and biases, which can be learned from the data during the model training.

This study uses the multi-layer feed-forward perceptron (MLP) to train the data [2, 100]. Three different datasets are adopted. Namely, a training dataset, $D_{tr,\text{HF}} = \left\{ \left(\mathbf{x}_{tr,\text{HF}}^{(i)}, \mathbf{y}_{tr,\text{HF}}^{(i)} \right) : 1 \leq i \leq n_{\text{HF}} \right\}$ and $D_{tr,\text{LF}} = \left\{ \left(\mathbf{x}_{tr,\text{LF}}^{(i)}, \mathbf{y}_{tr,\text{LF}}^{(i)} \right) : 1 \leq i \leq n_{\text{LF}} \right\}$, are used to estimate the gradients during the training of the neural networks in the two different stages of the BF-DNN model. A holdout validation dataset $D_{val} = \left\{ \left(\mathbf{x}_{val}^{(i)}, \mathbf{y}_{val,\text{HF}}^{(i)}, \mathbf{y}_{val,\text{LF}}^{(i)} \right) : 1 \leq i \leq n_{val} \right\}$ is applied to optimize the model's hyperparameters, and a test dataset $D_{test} = \left\{ \left(\mathbf{x}_{test}^{(i)}, \mathbf{y}_{test,\text{HF}}^{(i)}, \mathbf{y}_{test,\text{LF}}^{(i)} \right) : 1 \leq i \leq n_{test} \right\}$, never seen before during training and validation, estimates the performance of the final tunned model.

A Multilevel DNN is applied to capture the non-linear data correlation in a BF context and learn the multidimensional function $F = \left[F_{\text{HF}}(\mathbf{x}), F_{\text{LF}}(\mathbf{x}) \right]^{\text{T}}$ using different NNs to model these functions. The Bi-Fidelity Deep Neural Network (BF-DNN) is fully connected and comprises two stages, each configured as an MLP. The architecture of the NN model is shown in Fig. 3.1.
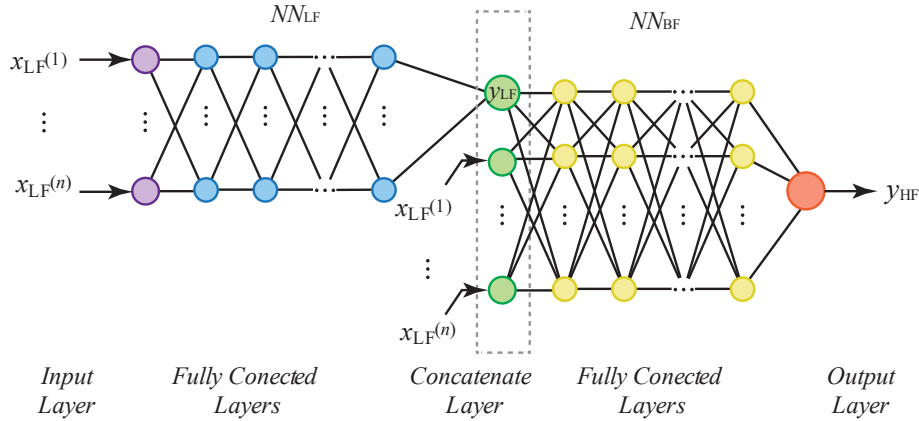
Figure 3.1 – Architecture of the composite neural networks for multi-fidelity modeling.

As depicted in Fig. 3.1, the BF-DNN model learns the unknown implicit function $F\left(\mathbf{x}, y_{\text{LF}}(\mathbf{x})\right)$ that maps the LF data to the HF level of Eq. 3.2 in the same structure. The DNN architecture consists of five levels trained simultaneously: input layer, LF hidden layers, concatenate layer, BF hidden layers, and output layer. In the Input Layer are n $\mathbf{x}_{tr,\text{LF}}$ input neurons (shown in purple). The LF Fully Connected Layers (blue) train the y_{LF} function considering a non-linear mapping. In the concatenate layer, there are $n+1$ input neurons (shown in green). They are $\mathbf{x}_{tr,\text{HF}}$, and $y_{\text{LF}}(\mathbf{x}_{\text{HF}})$. The BF Fully Connected Layers (yellow) train the y_{HF} function considering a non-linear mapping. The output neuron in the Output Layer

(orange) is $y_{HF}(\mathbf{x}_{HF})$. As an advantage of this BF-DNN model, the HF input locations should not be required to form a subset of the LF input locations.

According to Fig. 3.1, the first stage (NN_{LF}) learns $F_{LF}(\mathbf{x})$. Before that, it is possible to predict the values of the LF function at the training inputs $\mathbf{x}_{HF}$ of the HF data inputted in the concatenate layer denoted by $F_{LF}^{NN}(\mathbf{x}_{LF}) = \left\{ F_{LF}^{NN}(\mathbf{x}_{HF}^{(i)}) : 1 \leq i \leq n_{HF} \right\}$. Then a second stage ($NN_{BF}$) approximates the HF function $y_{HF}(\mathbf{x})$ based on the input data $\left(\mathbf{x}_{HF}, F_{LF}^{NN}(\mathbf{x}_{HF}) \right) = \left\{ \left(\mathbf{x}_{HF}^{(i)}, F_{LF}^{NN}(\mathbf{x}_{HF}^{(i)}) \right) : 1 \leq i \leq n_{HF} \right\}$ and the available HF output data $y_{HF}(\mathbf{x}_{HF}) = \left\{ \mathbf{y}_{HF}^{(i)} : 1 \leq i \leq n_{HF} \right\}$.

In the ANN, the neuron is a processing element with several inputs and one output. Each m neuron receives an input signal vector $\mathbf{x} = \left\{ \mathbf{x}^{(i)} : 1 \leq i \leq n \right\}$ from n input channels. The neurons are connected using connection weights w, each containing a bias and an activation. The weighted sum of x is calculated by multiplying each element x_c by a coefficient w_{mc}, indicating the importance of the input channel c. The activation a_m of the m-neuron is given by:

$$a_m = \sum_{c=1}^{n} w_{mk} x_n + b_m \tag{3.7}$$

where $b_m \in R$ is the bias, is a constant corrective term allowing a non-negative activation a_m. The output signal value y is calculated as a function of the activation. Table 3.1 presents some activation functions commonly applied in ANN.

Table 3.1 – Features of activation functions.

Activation Function	Function	Domain
Hyperbolic tangent	$f(a_m) = \left(e^{a_m} - e^{-a_m} \right) / \left(e^{a_m} + e^{-a_m} \right)$	$(-1,1)$
Sigmoid	$f(a_m) = 1 / \left(1 + e^{-a_m} \right)$	$(0,1)$
Rectified Linear Units (ReLU)	$f(a_m) = \begin{cases} a_m & a_m \geq 0 \\ 0 & a_m < 0 \end{cases}$	$[0,\infty)$
Exponential Linear Unit (ELU)	$f(a_m) = \begin{cases} a_m & a_m \geq 0 \\ \omega \left(e^{a_m} - 1 \right) & a_m < 0 \end{cases}$ with parameter $\omega = 1$	$(-\omega,\infty)$

The number of layers and neurons of each multi-layer NN strongly affects the predictive accuracy. It is important to emphasize that no specific constraint is imposed on the LF model. However, a convergence FE mesh study is necessary to evaluate the correlation between the LF and the HF model to guarantee that the resulting model is more accurate. The unknown NN parameters are decided by minimizing the following loss function using Root Mean Squared Error (RMSE) of the LF and HF D_{val} in association with weight regularization, used as a general technique to reduce overfitting of the D_{tr} and improve the generalization of the model as follow:

$$Loss = \sqrt{\frac{1}{n_{LF}}\sum_{i=1}^{n_{LF}}\left(\left|\hat{y}_{LF}^{(i)} - y_{LF}^{(i)}\right|\right)^2} + \sqrt{\frac{1}{n_{HF}}\sum_{j=1}^{n_{HF}}\left(\left|\hat{y}_{HF}^{(j)} - y_{HF}^{(j)}\right|\right)^2} + \tau\sum w_i^2, \tag{3.8}$$

where $\hat{y}_{LF}$, y_{LF}, $\hat{y}_{HF}$ and y_{HF} denote, respectively, the predicted value of the validation LF in the NN_{LF} stage, the real response of the LF point, the predicted value of the validation HF in the NN_{BF} stage, and the real response of the validation HF point; $\tau > 0$ is a penalty parameter, and w denotes weights and biases of neural networks. Once all training data are introduced, the backpropagation optimization is utilized to update the weights and biases of the model iteratively until the loss function reaches the minimum [111].

The loss function can be optimized using a method such as the Adaptive Moment Estimation (ADAM) method along with Xavier's initialization method [56, 111]. The ADAM is a method for efficient stochastic optimization that only requires first-order gradients with little memory requirement, is relatively easy to implement, computationally efficient, and is well suited for problems that are large in terms of data or parameters. Moreover, the method is designed to combine the advantages of two popular methods for gradient-based optimization: Adaptive Subgradient Methods , which works well with sparse gradients, and Root Mean Squared Propagation, which works well in online and non-stationary settings [117, 118].

3.2.2.1 Hyperparameter optimization based on Bayesian Optimization

A machine learning model A with a configuration space of the overall hyperparameters Λ performs best on a particular problem with an optimal combination of its hyperparameters λ^*. The space Λ can include both discrete and continuous dimensions. Given a dataset D_{tr}, the goal is to find the best configuration such that:

$$\lambda^* = \max_{\lambda \in \Lambda} E_{(D_{tr}) \sim D} L\left(A_\lambda, D_{tr}\right) \tag{3.9}$$

where L(A_λ, D_{tr}) denotes the validation performance of A_λ on dataset D.

Various optimization techniques can obtain an optimal combination of hyper-parameters. Traditionally, manual or automatic techniques (such as random and grid search) have been used as hyperparameter optimization [102]. The manual method tends to generate human errors and needs expertise. The accuracy in grid search decreases with the increase in the number of parameters in the optimization process. Moreover, random search is based on random distribution functions, which might miss optimal spots in the investigation. To overcome the difficulty in heuristics associated with machine learning models, Bayesian Optimization has been leveraged in the hyperparameters selection of DNN [102, 103, 119].

Bayesian optimization (BO), derived from the Bayes theorem, is an effective method for solving computationally demanding optimization without closed-form solutions [120, 121]. The BO algorithm adds sample points to the posterior distribution of an objective function L(A_λ,D_{tr}), approximated by a Gaussian process [101, 122]. The GP is a function where the variable is a Gaussian distribution:

$$f(h) \sim GP\big(m(h), k\big(h, h'\big)\big). \tag{3.10}$$

Here $m(h)$ is the distribution's mean function, and $k(h,h')$ is the covariance function of two tested points, h and h'. In this study, the Matérn 5/2 kernel can be helpful in the optimization setting [101]. This proposed study chooses the Expected Improvement as an acquisition function because it takes work to fall into the local optimum solution [102]. Considering an acquisition function, where high values of the acquisition function are associated with high values of the objective function, the maximum value of the objective function is obtained through an iterative process of maximizing the acquisition function at each step of Bayesian optimization. This process continues until the posterior distribution fits the actual distribution.

The BO algorithm as hyperparameter optimization requires few iterations and a high operation speed. In addition, a strong point of the BO algorithm is that it can optimize integers and categorical variables, typical in DNN hyperparameters. The details of the BO algorithm can be found in [101, 102, 121].

In this study, for all examples and both DNN stages, the hyperparameters: the number of hidden layers, neurons in each hidden layer, initial learning rate, and activation function, are tuned by BO. The hyper-parameters that required optimization (and their ranges) are given in Table 3.2.

42

Table 3.2 – The hyperparameters to be optimized with a defined search space.

Hyperparameter	Limits	Domain
Each Layer Size	[50,200]	Integer
Number of layers (NN_{LF})	[1,4]	Integer
Number of layers (NN_{BF})	[1,3]	Integer
Activation Function	[elu, relu, sigmoid, tanh]	Categorical
Learning rate	[5e-3,5e-2]	Real

3.2.3 Subset simulation for failure probability estimation

The performance function $G(\mathbf{x})$ provides the limit state function when $G(\mathbf{x})=0$, which separates the safe region $G(\mathbf{x})>0$ from the failure region $G(\mathbf{x})<0$. Mathematically, for a given structural system input parameter $\mathbf{x} = \left\{ \mathbf{x}^{(i)} : 1 \leq i \leq n \right\}$, the parameter P_f can be estimated by solving the multifold probability integral defined as:

$$P_f(\mathbf{x}) = P(F = G(\mathbf{x}) < 0) = \int_F f(\mathbf{x}) dx \tag{3.11}$$

where $f(\mathbf{x})$ is the joint probability density function of $\mathbf{x}$. Considering the statistical technique of Direct Monte Carlo Sampling (MCS), the failure probability can be estimated as follows:

$$P_f = \frac{1}{N} \sum_{i=1}^{N} I_F\left(\mathbf{x}^{(i)}\right) \tag{3.12}$$

where $\mathbf{x}^{(1)}, \ldots \mathbf{x}^{(N)}$ are i.i.d. samples from $f(\mathbf{x})$, and $I_F(\mathbf{x})$ stands for the indicator function, i.e.,

$$I_F(\mathbf{x}) = \begin{cases} 1 \text{ if } x \in F \\ 0 \text{ if } x \notin F \end{cases} \tag{3.13}$$

A brief introduction of the SS technique, largely following Ref. [13], will be provided. The Subset Simulation (SUS) is an adaptive Monte Carlo method proposed for estimating small failure probabilities in high-dimensional problems. The SUS is based on nested sets $F_1 \supset F_2 \supset \cdots F_M$, where $F_M=F$ is the failure event, and expresses the failure probability P_f as a product of conditional probabilities:

$$\hat{P}_f = P(F) = P_{F_1} \prod_{j=2}^{M} P(F_j \mid F_{j-1}) \tag{3.14}$$

where P_{F_1} is the first unconditional failure probability computed as the fraction of samples exceeding the current threshold level L_1, and $P(F_j \mid F_{j-1})$ are the subsequent conditional failure probabilities that are conditional on exceeding the prior intermediate thresholds in level L_{j-1}, and are computed as the fraction of samples exceeding the threshold level L_j.

In expressing $\hat{P}_f$ as a product of larger failure probabilities, SUS creates intermediate failure thresholds L before the required zero thresholds. As for SUS, the algorithm starts with a direct MCS to estimate P_{F_1}, while a Markov chain Monte Carlo is used to define $P_{F_j|F_{j-1}}$.

In SUS, the values $F(\mathbf{x}^{(k)})$, $k = 1, ..., N_n$, of the N_n samples are in the subset $S^{(1)}$. Crude MCS generates the samples independently identically distributed assuming the Nataf transformation, whereas samples in the subsequent subsets are generated by Markov Chain Monte Carlo (MCMC) simulation and correlated, considering the dependence of the seeds for each Markov chain. The samples falling between two subsequent intermediate failure thresholds L_{j-1} and L_j constitute a subset. The intermediate failure thresholds must be specified to estimate the intermediate failure probabilities. The range of the samples $N_s = p_0 N_n$ are retained in each step and serves as the seed for the Markov chains. The value $p_0 = 0.1$ is often used in the literature, which makes F_1 a relatively frequent event.

Au and Beck [13] proposed a Metropolis–Hastings algorithm, and this method is popular for simulating the conditional samples in $P_{F_j|F_{j-1}}$, but occasionally it can lead to degenerate sampling when dealing with models having geometrically complex performance functions. Papaioannou *et al*. [15] proposed a delayed rejection Markov chain Monte Carlo method to alleviate these issues, which is used in this study.

In addition, an approximate COV to estimate the intermediate failure probability in the subset is given by [16]:

$$\hat{\delta}_j = \begin{cases} \sqrt{\left(1 - P_{F_1}\right) / \left(P_{F_1} N_n\right)} & \forall \quad j = 1 \\ \sqrt{\left(1 - P_{F_1|F_{j-1}}\right) / \left(P_{F_1|F_{j-1}} N_n\right)} & \forall \quad 1 < j \le N_s \end{cases} \tag{3.15}$$

where N_n is the number of samples in a subset. The overall COV estimate over the required failure probability is:

$$COV_{\hat{P}_f} = \sqrt{\sum_{i=1}^{N_S} \hat{\delta}_i^2}. \tag{3.16}$$

The generalized reliability index β is expressed as:

$$\beta = -\ ^{-1}\ P_f \tag{3.17}$$

where $\ ^{-1}\ *$ is the inverse CDF function of the standard Gaussian distribution.

3.2.4 Reliability analysis using Subset simulation with the BF-DNN model

This section briefly overviews the BF-DNN model associated with SUS for reliability analysis of high-dimensional cases with complex performance functions with multi-fidelity modeling of non-linear fidelities correlations. A flowchart, shown in Fig. 3.2, summarizes it.
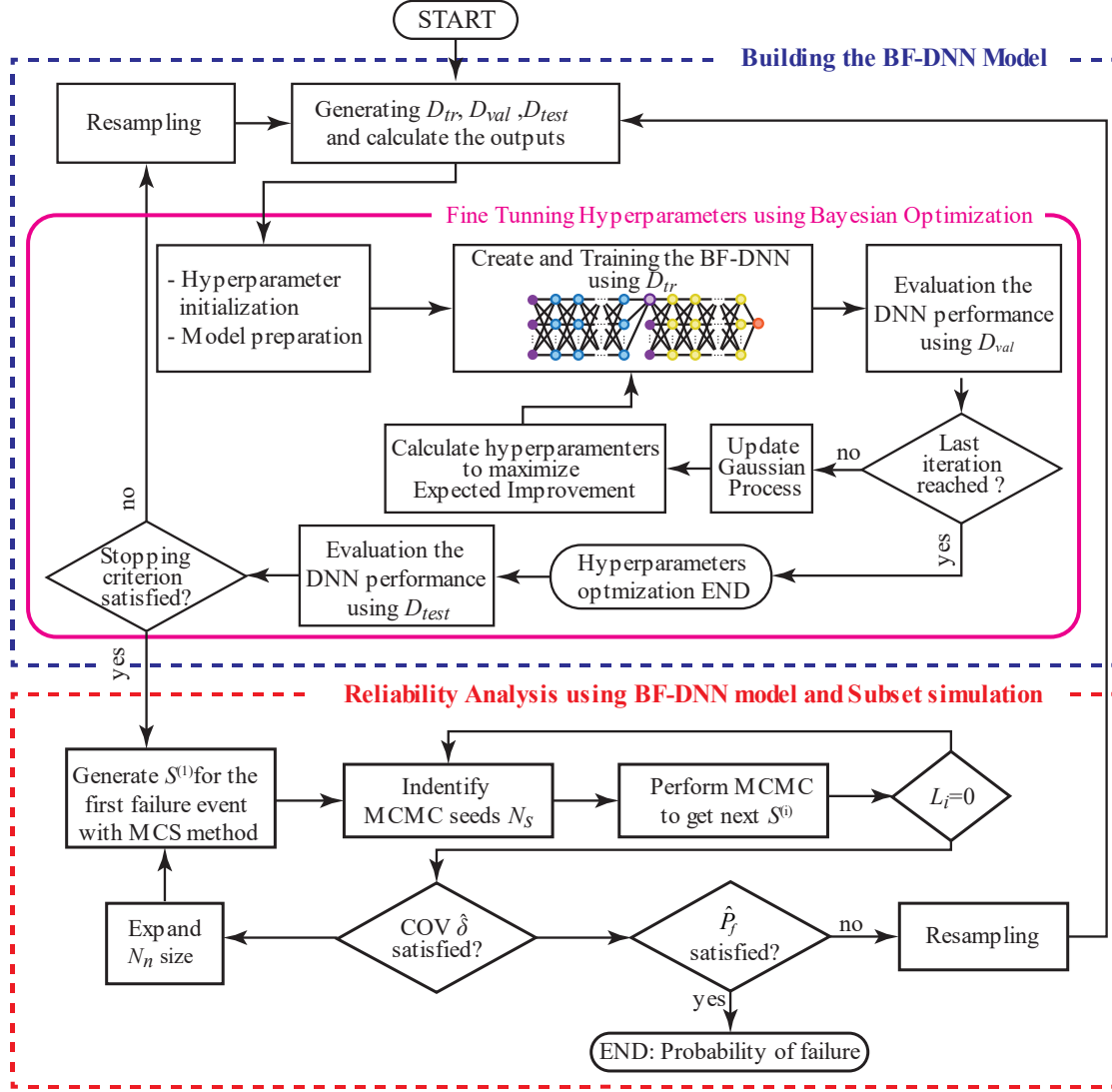
Figure 3.2 – Flowchart of the BF-DNN model and Reliability analysis.

Stage 01 develops the BF-DNN framework, and Stage 02 presents the structural reliability assessment using SUS. The steps of the proposed algorithm may be summarized as follows:

Stage 01 - Step 1: Generate the initial training database $D_0 = \{D_{tr}, D_{val}, D_{test}\}$ randomly by the Latin Hypercube sampling method and calculate the corresponding responses $\mathbf{y}_{tr}^{(0)} = \left[\mathbf{y}_{tr,\text{HF}}; \mathbf{y}_{tr,\text{HF}}; \mathbf{y}_{val}; \mathbf{y}_{test} \right]$ by the HF and LF functions.

Stage 01 - Step 2: Initialize the model and define the initial hyperparameters.

Stage 01 - Step 3 and Step 4: Create and train the BF-DNN model using the D_{tr} dataset. In each iteration, the training step is validated using Holdout Validation. Afterwards, evaluate the BF-DNN performance using D_{val}. If the last iteration is reached, the hyperparameters optimization ends and turns to Stage 03-Step 7. Otherwise, turn to Stage 01-Step 5.

Stage 01 - Step 5 and Step 6: Construct the Gaussian surrogate model and select new data points by optimizing the acquisition function, calculating the corresponding functions, and augmenting the data. Repeat Steps 3 and 4 with the latest data.

Stage 01 - Step 7: Evaluate the BF-DNN performance using D_{test}. If convergence criteria are satisfied, the BF-DNN construction is finished. To quantify the accuracy of the proposed BFMT-DNN model is used the sum of the average relative error (ε_{avg}) referring to the two stages of the BF-DNN is given as:

$$\varepsilon_{\mathrm{avg}} = \frac{1}{n_{test}} \sum\nolimits_{i=1}^{n_{test}} \left(\frac{\left| \hat{y}_{\mathrm{LF}}^{(i)} - y_{test,\mathrm{LF}}^{(i)} \right|}{y_{test,\mathrm{LF}}^{(i)}} \right) + \frac{1}{n_{test}} \sum\nolimits_{j=1}^{n_{test}} \left(\frac{\left| \hat{y}_{\mathrm{HF}}^{(j)} - y_{test,\mathrm{HF}}^{(j)} \right|}{y_{test,\mathrm{HF}}^{(j)}} \right) \qquad (3.18)$$

Where $y_{test,\mathrm{LF}}$ and $y_{test,\mathrm{HF}}$ denotes the LF and HF reference value of the D_{test}, $\hat{y}_{\mathrm{LF}}$ and $\hat{y}_{\mathrm{HF}}$, the LF and HF predicted values by BFMT-DNN. It is considered ε_{avg}=0.6% as the stop criterion.

Stage 02 – Step 1: Generate the first random sample data $S^{(1)}$ using the MCS method. The value of N_s depends on the different problems, and it is much smaller than the number required by the original MCS method. Approximately samples are needed to estimate a failure probability by MCS-based methods. In this paper, the initial N_n is generally set to 3×10^3, also of the other intermediate failure events size.

Stage 02 – Step 2 and Step 3: In ascending order, select the first 10% random failure points and corresponding minimum responses as seeds for each Markov chain. Use the MCMC method to generate the new random sample data $S^{(i)}$.

Stage 02 – Step 4: If $L_i \geq 0$, turn to Stage 03-Step 2. Otherwise, set L_i=0.

Stage 02 – Step 5: Calculate the coefficient of variation $COV_{\hat{P}_f}$ based on Eq. 3.19. If $COV_{\hat{P}_f} \leq 0.1$, turn to Stage 03-Step 6. Otherwise, turn to Stage 02-Step 7.

Stage 02 – Step 6: Expand the $N_n = N_n + 1 \times 10^3$ size of the random point and then go to Stage 03-Step 1.

Stage 02 – Step 7: Output the final $\hat{P}_f$. If $\hat{P}_f$ is satisfied, end the proposed method. Otherwise, resample the initial dataset D and turn to Stage 01-Step 1.

3.3   APPLICATION TO A MARITME INDUSTRY PROBLEM

The evaluation of the proposed bi-fidelity surrogate models based on an FE analysis for structural reliability analysis is performed based on an application problem of a fundamental problem of the shipbuilding industry. The problem studied consists of the reliability assessment of the ultimate strength of stiffened panel elements under axial compression, which are present in the deck structure of double-hull oil tanker ships as introduced by Gaspar *et al.* [3] and Chojaczyk *et al.* [2]. Figure 3.3(a) illustrates the model of the stiffened panel under the compressive load σ_z described by a non-linear FE structural model.
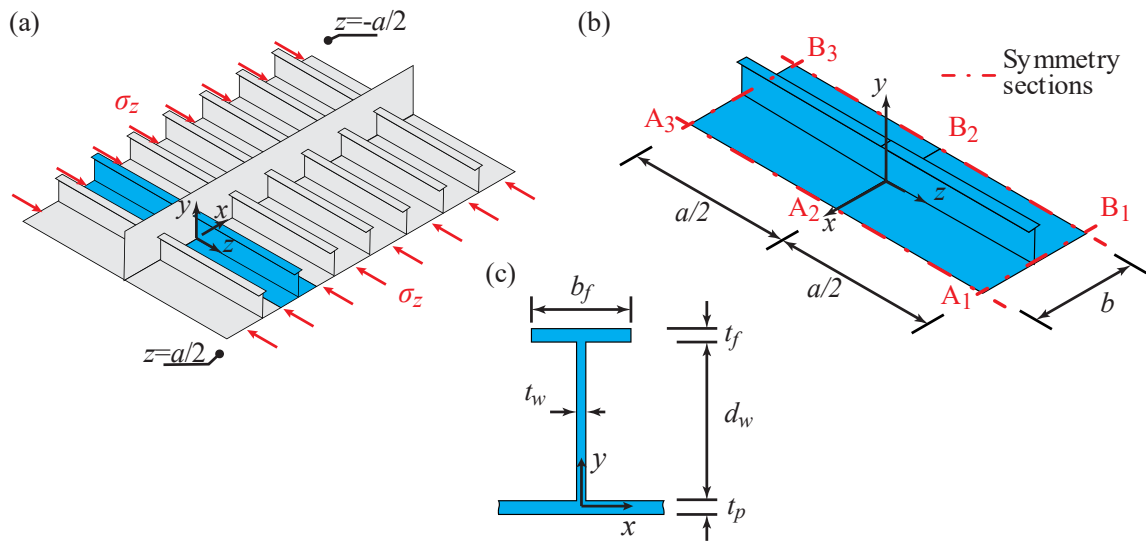
Figure 3.3 – Model of the stiffened panel (a) Deck under uniaxial compression in the longitudinal direction, (b) geometry of the FE model, and (c) cross-section dimensions.

3.3.1   Finite Element Model

The FE analysis of the present study consists of the following steps: (*i*) Definition of the model characteristics and choice of the finite element, (*ii*) definition of imperfections, boundary, and load conditions, (*iii*) modeling of the structure's geometry, (*iv*) simulation and (*v*) post-processing and analysis of results.

Considering the BF input values, an output database is obtained for the training points via non-linear FEM. These values comprise the ultimate strength results (σ_{zu}) for each sample $\mathbf{x}^{(i)}$. The FE mesh defines the fidelity for the high and low-fidelity models. A mesh convergence test determines the convergence of the FE mesh discretization with HF and LF. This test is relevant to define the mesh size that guarantees a non-significant reduction in the value of $\mathbf{y}_{LF}$, compared to $\mathbf{y}_{HF}$, but with a significant computational time reduction. In the present study, the stress-strain curves that describe the structural behaviour of the stiffened panel under uniaxial compression are simulated using the FE software ABAQUS. The non-

linear structural behaviour of the stiffened panel elements under axial compression is accounted for using the finite element S4R, a 4-node general-purpose, quadrilateral, stress/displacement shell element with reduced integration and large-strain formulation. The non-linear material behaviour is modeled using a linear-elastic and ideally-plastic material law, neglecting the strain-hardening effect. This material law is implemented using Bilinear Isotropic Hardening. Other constitutive models can be used depending on the nature of the application [22, 123].

The FE structural model adopted for the stiffened panel is a half-plus half-span (1/2+1/2) model, depicted in Fig. 3.3(b). The stiffened panel comprises steel plates, transverse girders, and T-type cross-section longitudinal stiffeners, as shown in Fig. 3.3(c). Since the panel is symmetrical, only one of the reinforcements of the reinforced panel with attached plating is analyzed, as shown in Fig. 3.3. This same model was also adopted in [3, 71].

In assessing the buckling strength of stiffeners or panels, stresses acting on the stiffeners or panels are used assuming that complete structures are in a 50% corroded state for assessment on the side of safety according to Common Structural Rules for Bulk Carriers and Oil Tankers by [69]. The design values for the corroded condition of the deck stiffened panel defined as deterministic variables are a=5450 mm for the stiffener span or spacing between transverse frames and b=900 mm for the spacing between stiffeners in the panel. The midship cross-section modulus at the deck for the corroded scantlings is Z_v=37.3 m^3. The plate and stiffeners geometry variables are t_p=20.5 mm for the plate thickness, d_w=400 mm for the web height, b_f=100 mm for the flange breadth, t_w=9.5 mm for the web thickness, and t_f=14.0 mm for the flange thickness.

The material of the plate and stiffeners is AH32 high-strength steel. This material has Young's modulus E=206000 MPa, yield stress σ_y=315.0 MPa, and Poisson's ratio v=0.3. The design value for σ_y corresponds to a 5% percentile characteristic value from the probability density function [74].

Symmetry boundary conditions are imposed at the attached plating longitudinal edges and the stiffener mid-span transverse sections. The displacement conditions of the stiffened panel are represented by [u_x, u_y, u_z, r_x, r_y, r_z] and applied as per Fig. 3.3(b). For the present analysis, consider the following boundary conditions, where "0" indicates translational or rotational constraint and "1" mean no constraint:

- A_1-A_3 and B_1-B_3 border (symmetric conditions): [0,1,1,1,1,0];

- A_1-B_1 border (symmetric condition): [1,1,0,0,0,1];

- A_3-B_3 border (symmetric condition): [1,1, d_z,0,0,1];

- A_2-B_2 for plate nodes: [1,0,1,1,1,1]

- A_2-B_2 for stiffener web nodes: [0,1,1,1,1,1]

where d_z indicates the displacement imposed at the transverse edge. The displacements of nodes on borders A_1-A_3 and B_1-B_3 along the x-direction are linked. This process safeguards uniform displacement even under a point force applied to one of the coupled nodes and avoids the local imperfections of the profile.

Some initial imperfections are also applied to the FE model to better represent the inevitable deformations of the ship steel structure's complex fabrication and welding process. The considered pattern is depicted in Fig. 3.4, and analytical functions of the form approximate the shapes of these initial imperfections [70]:

- Initial deflection of local plate panel:

$$w_{0p}\left(y,z\right) = w_{op} \sin\frac{m\pi z}{a} \sin\frac{\pi y}{b} \tag{3.19}$$

- Initial deflection of stiffener column-type imperfection:

$$w_{0c}\left(z\right) = w_{oc} \sin\frac{\pi z}{a} \tag{3.20}$$

- Initial deflection of the stiffener sideways imperfection:

$$w_{0s}\left(x,z\right) = w_{os} \frac{x}{d_w} \sin\left(\frac{\pi z}{a}\right) \tag{3.21}$$

where w_{op}, w_{oc}, and w_{os} are the following design values for the amplitudes given by Paik and Kim [71], with $w_{op}=b/200$ and $w_{oc}=w_{os}=a/1000$. The number of buckling half-waves in the longitudinal direction (m) is usually the integer of $m=a/b$, which is considered $m=7$.

(a) Local plate panel   (b) Stiffener column-type   (c) Stiffener sideways
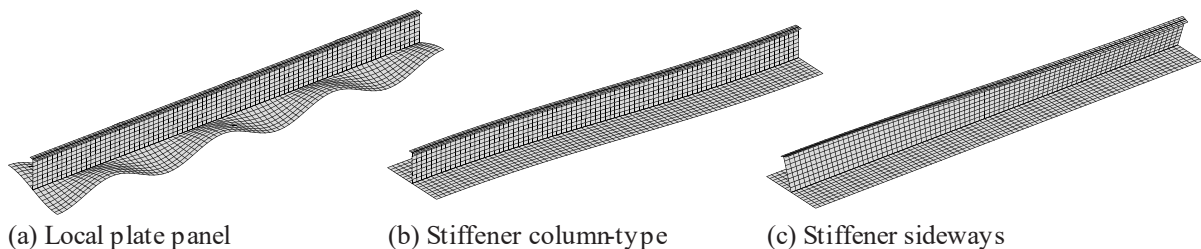
Figure 3.4 – Initial geometrical imperfections of the stiffened plate (mean amplitudes with scale factor 50×).

3.3.2 Performance function

Stiffened panels are subject to extreme axial compressive loads induced by ship hull girder bending moments, and therefore their ultimate compressive strength is an essential design requirement. The still water and vertical wave-induced bending moments are the primary loads that contribute to the longitudinal compressive stress of the bottom stiffened panel. The magnitude of these loads depends on several factors as the load condition of the ship, its operational profile, and uncontrollable environmental factors such as the sea state. Therefore, the safety margin associated with this hull beam failure mode can be described by a performance function of the form [8, 65] :

$$G(x) = \sigma_{zu} - \sigma_{za} \tag{3.22}$$

with σ_{zu} the ultimate compressive strength of the stiffened panel elements, given as follows:

$$\sigma_{zu} = \max\left(\sigma_z\left(\mathbf{x}^M, \mathbf{x}^G\right)\right) \tag{3.23}$$

where $\mathbf{x}^M$ is a vector of basic material variables, and $\mathbf{x}^G$ is a vector of basic geometric variables. The σ_{za} is the uniaxial compressive strength induced by the ship hull girder bending moments given by:

$$\sigma_{za} = \frac{M_{sw} + M_{wv}}{Z_v} \tag{3.24}$$

where M_{sw} and M_{wv} are the random vertical still water and wave-induced bending moments, respectively, and Z_v is the midship cross-section modulus at the deck.

3.3.3 Random variables definition

The ten basic random variables considered in the problem of Section 3.2 and the two random variables considered in the load components of the performance function in Eq. 3.22 are given in Table 3.3. Their corresponding probability density distribution (P.D.), statistical moments mean (μ), and standard deviation (*std*) are presented.

These basic random variables describe the stiffened panel's geometric and material properties and the welding-induced initial imperfection amplitudes. The probabilistic density distributions and moments in Table 3.3 are adopted for the structural dimensions, and material properties are established in the literature. They are commonly used in structural reliability analysis, e.g., [2, 8, 73, 124]. In this study, the probabilistic model for the amplitudes of the initial deflections is based on [65, 75]. The parameters of the lognormal distributions are

derived considering the design amplitudes as 95% percentile characteristic values, and the coefficient of variation (COV) is assumed to equal 0.50.

Table 3.3 – Random variables for strength and load definition.

Variable	Type	Units	P.D.	μ	std	COV (%)
t_p		mm	Normal	20.50	0.41	2.0
t_w		mm	Normal	9.50	0.19	2.0
t_f		mm	Normal	14.00	0.28	2.0
$\sigma_{yd,p}$		MPa	Lognormal	348.00	20.88	6.0
$\sigma_{yd,s}$	Strength	MPa	Lognormal	348.00	20.88	6.0
E_p		MPa	Lognormal	2.06e5	1.23e4	6.0
E_s		MPa	Lognormal	206e5	1.23e4	6.0
w_{op}		mm	Lognormal	2.30	1.15	50.0
w_{os}		mm	Lognormal	2.80	1.40	50.0
w_{oc}		mm	Lognormal	2.80	1.40	50.0
M_{sw}	Load	MNm	Normal	1483.70	430.30	29.0
M_{wv}		MNm	Gumbel	4603.15	414.30	9.0

The ship hull girder bending moments are the load basic random variables of the problem. The statistical parameters adopted for the hull girder vertical bending moments along a ship voyage are defined based on Hørte *et al.* [73] and as specified in IACS [69]. The still water bending moment (M_{sw}) is described by a normal distribution with $\mu=0.7 \cdot M_{sw,max}$ and $std=0.2 \cdot M_{sw,max}$, where $M_{sw,max}=2119.6$ MNm is the maximum still water bending moment specified in the ship loading manual.

Considering a reference period $T_r=1.0$ years for the extreme values of vertical wave-induced bending moment representative of a North Atlantic crossing, the stochastic model proposed by Guedes Soares *et al.* [58, 67] defines the wave-induced bending moment. The model is based on a two-parameter Weibull distribution to describe the peak values of vertical wave-induced bending moment at a random point in time with shape parameter $k=1$ and scale ξ parameter satisfying: $P\left[M_{wv} > M_{wv,max} \right]=10^{-8}$, where $M_{wv,max}$ is the maximum vertical wave-induced bending moment considered for design purposes [69]. The Gumbel distribution then describes the peak values over the T_r:

$$F_{M_{wve}}(M_{wv}) = \exp\left[-\exp\left(-\frac{M_{wv}-\varphi}{\zeta} \right) \right] \tag{3.25}$$

The Gumbel model parameters can be derived based on the initial Weibull distribution parameters φ and ζ, by Guedes Soares and Teixeira [76] as $\varphi = \xi(\ln n_c)^{1/k}$ and $\zeta = \xi/k\,(\ln n_c)^{1-k/k}$, where n_c corresponds to the mean number of load cycles expected over the operation period T_r [76]. The number of wave cycles corresponding to the return period is calculated considering an average wave period $T_w= 8$ s, which is the one that is applicable for

51

the areas of the North Atlantic. More details can be found in Hørte *et al.* [73] and Gaspar *et al.* [65].

3.4 CASES STUDIES

Next, three representative MF cases are adopted to verify the performance of the proposed algorithm, and the association with the SUS to reliability performance is presented. First, the effectiveness of this BF-DNN model is demonstrated on various artificial benchmarks and compared against commonly used MF methods. In addition, a real-world engineering problem experiment, presented in Section 3, validates the proposed BF-DNN surrogate model.

Examples 1 and 2 assume that the cost ratio between the HF and LF functions is 10:1, i.e., $c_{HF}=1$ and $c_{LF}=0.1$. N_{call} is the total cost, which equals the equivalent number of HF samples. The hyperparameters of the BF-DNN, for instance, the parameter to balance the Loss function terms (Eq. 3.8), are optimized by hold-out validation and BO. The HF data is split into 10% for the validation and 90% for the training. Each optimization iteration is stopped after reaching any of the following criteria: achieving a defined RMSE loss function or a fixed number of 30000 epochs. The BO algorithm stops after reaching the following: a fixed number of 50 iterations or a fixed time is defined as 7.2×10^5 seg. The regularization is fixed $\tau = 10^{-4}$. Therefore, the hyperparameters are optimized throughout the layers, and global optimization is executed across the validation set, thereby reducing the time taken and improving the model's performance. MATLAB is used to optimize the hyperparameters and build the BF-DNN model.

3.4.1 Example 1: 20-dimensional function approximation

The first example is the 20-*d* problem presented by Meng and Karniadakis [107], which is hard to approximate with other methods and is used to compare the proposed BF-DNN framework with another MF-DNN presented by the authors. The following high-dimensional functions define the MF setting in this example:

$$f_{HF}(\mathbf{x}) = (x_1 - 1)^2 + \sum_{i=2}^{20} \left(2x_i^2 - x_{i-1}\right)^2 , \tag{3.26}$$

$$f_{LF}(\mathbf{x}) = 0.8 y_{HF}(x) - \sum_{i=1}^{19} 0.4 x_i x_{i+1} - 50. \tag{3.27}$$

The random variables $\mathbf{x} = \{x_1, x_2, \ldots, x_{20}\}$ obey the uniform distribution in [-3,3]. The authors ([107, 125]) used the 3-step MF-DNN using $n_{HF}=5000$ and $n_{LF}=30000$, resulting in

N_{call}=8000. However, although a good prediction is depicted in [107, 125], the authors present no accuracy estimate for the test prediction.

The proposed BF-DNN method obtains good accuracy with n_{LF}=25000 and n_{HF}=3000, resulting in N_{call} =5500. The hyperparameters are tuned according to the search space presented in Table 3.2. The optimal parameter combinations of the BF-DNN after the BO are presented in Table 3.4. Figure 3.5(a) depicts the results prediction of a D_{test} dataset with n_{test}=10000 using only the available n_{HF}=3000 high-fidelity data that does not lead to an accurate function approximation. Figure 3.5(b) illustrates the excellent results prediction of the D_{test} dataset using the obtained BF-DNN. The average relative error $\varepsilon_{\hat{f},avg} = \left| \hat{f} - f_{ref} \right| / f_{ref}$ is used to quantify the accuracy of the predictions, with f_{ref} the reference value considered and $\hat{f}$ the predicted value. The selected BF-DNN surrogate model has an average relative error $\varepsilon_{\hat{f},avg} = 0.877\%$ for the HF prediction at D_{test}.

Table 3.4 – *Tunned hyperparameters by Bayesian Optimization in Example 01[a].*

n_{HF}	n_{LF}	Number of layers	Hidden Layer's Size	Act. Function	Learning rate ($\times 10^{-3}$)
3000	-	1	108	elu	0.87
3000	25000	(2, 2)	[(81 126), (62 55)]	(elu, elu)	(1.38, 0.91)

[a] The MF description references to (NN_{LF}, NN_{BF}) and [(NN_{LF}), (NN_{BF})].
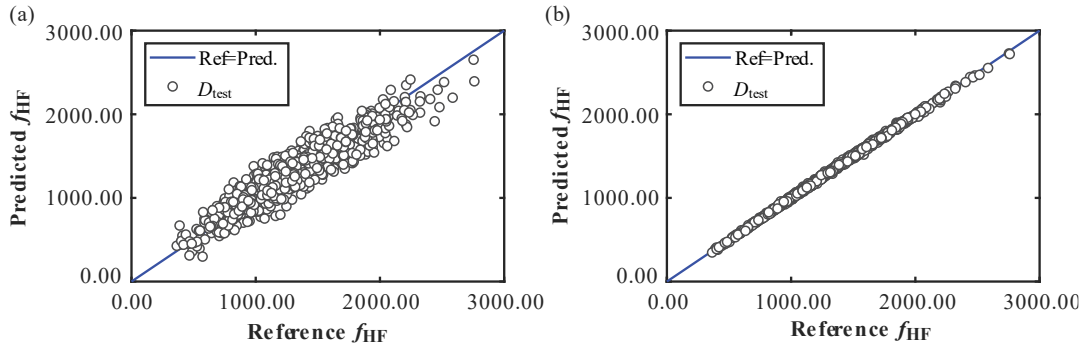
Figure 3.5 – *Approximations of the 20-d function at D_{test} using: (a) Single-fidelity DNN, and (b) BF-DNN.*

3.4.2 Example 2: Non-linear performance function

The proposed strategy for reliability analysis is illustrated with an adaptation of the Griewank function in a BF context. The high-dimensional non-linear MF-PF is as follows:

$$f_{HF}(\mathbf{x}) = \sum_{i=1}^{d}(2x_i) + \prod_{i=1}^{d} 1.25\left(\cos\frac{x_i}{5} \right) \qquad (3.28)$$

$$f_{LF}(\mathbf{x}) = f_{HF}(\mathbf{x}) + \sum_{i=1}^{d} \sin\left(\frac{5x_i}{11}\right), \tag{3.29}$$

All the random variables $\mathbf{x} = \{x_1, x_2, \ldots, x_d\}$ are the independent standard normal random variables. The reliability problem is defined as $R = \Pr\{f_{HF} < 0\}$. Table 3.5 lists the optimal parameter combinations of $d = \{5, 20, 30\}$.

Table 3.5 – Tunned hyperparameters by Bayesian Optimization in Example 02.

d	n_{HF}	n_{LF}	BO iteration	Number of layers	Hidden Layer's Size	Activation Function	Learning rate ($\times 10^{-3}$)
5	100	500	23	(1, 2)	[(51), (62 55)]	(tanh, elu)	(1.38, 0.91)
20	700	3500	19	(1, 2)	[(94), (64 92)]	(tanh, elu)	(1.32, 0.75)
30	2000	8000	39	(1, 2)	[(134), (89 123)]	(elu, elu)	(1.06, 0.92)

It can be noted that the best performances of the three models after hyperparameter optimization are with one hidden layer in the NN_{LF} stage. The three models' NN_{BF} stage has two hidden layers with ELU as the activation function. The results of the proposed method and other methods, including MCS, MCS+SUS, and the commonly used Co-Kriging method [8, 28, 29], associated with SUS are summarized in Table 3.6. The same training points are shared between the Co-Kriging and the proposed BF-DNN. Each SUS analysis is submitted to 50 simulation cycles and $\hat{P}_f$ predictions.

Table 3.6 – Failure probabilities comparisons in Example 02.

Dimension	Method	N_{call}	$\hat{P}_f$	$\hat{\beta}$	$\varepsilon_{\hat{\beta}}(\%)$	T_c (seg)
5	MCS (reference)	10^6	5.63×10^{-2}	1.586	-	-
	MCS+SUS	9000	5.64×10^{-2}	1.585	0.063	-
	Co-Krigagem+ SUS	150	5.63×10^{-2}	1.586	0.000	568
	BF-DNN+SUS (proposed)	150	5.64×10^{-2}	1.585	0.063	1620
20	MCS	10^6	2.30×10^{-3}	2.833	-	-
	MCS+SUS	12000	2.22×10^{-3}	2.845	0.423	-
	Co-Krigagem+ SUS	1050	2.78×10^{-3}	2.770	2.223	2019
	BF-DNN+SUS	1050	2.81×10^{-3}	2.772	2.153	1862
30	MCS	10^7	3.09×10^{-4}	3.423	-	-
	MCS+SUS	15000	3.14×10^{-4}	3.419	0.117	-
	Co-Krigagem+ SUS	2800	3.56×10^{-4}	3.385	1.110	4039
	BF-DNN+SUS	2800	3.52×10^{-4}	3.395	0.818	3117

As illustrated in Table 3.6, the necessary N_{call} to obtain accuracy reliability responses is significantly smaller than that of MCS and MCS+SUS for all d values. The average relative error $\varepsilon_{\hat{\beta}} = \left| \hat{\beta} - \beta_{ref} \right| / \beta_{ref}$ is used to quantify the accuracy of the reliability index predictions, with β_{ref} the reference reliability index considered and $\hat{\beta}$ the predicted reliability index. Regarding the estimated accuracy, the proposed BF-DNN+SUS performs with slight fluctuation than the MCS+SUS and Co-Kriging+SUS. The most critical case is when d=20, with $\varepsilon_{\hat{\beta}} = 2.153\%$.

The reference T_c in Table 3.6 is the total computational cost for optimizing the BF-DNN hyperparameters over BO 50 iterations and estimating the Co-Kriging model's hyperparameters using GA. For cases with d=5, the Co-Kriging performs in reduced T_c than the BF-DNN approach. As the problem's dimensionality increases, the T_c for the BF-DNN model increase at a lower rate compared to the Co-Kriging model in estimating the hyperparameters. For the case d=20, the T_c of the BF-DNN is slightly less than Co-Kriging. However, for the case d=30, the T_c of the BF-DNN is fewer than Co-Kriging to predict a rare event. In this case, the Co-Kriging scheme has a 10000×10000 covariance matrix, resulting in large computational schemes for the hyperparameters optimization. However, the DNN-BF model needed a DNN architecture with a similar size to the low-dimensional BF-DNN cases to obtain similar accuracy compared to Co-Kriging.

3.4.3   Example 3: Stiffened panels by non-linear FE analysis

This section addresses the real-world engineering problem experiment presented in Section 3 to validate the proposed BF-DNN surrogate model. Specifically, the first analysis investigates the mesh influence on building the HF and LF models. Next, it explor the effectiveness of the number of support points in the BF-DNN model training. Then, the success of the proposed method is examined to assess the reliability of a stiffened panel under axial compression. Also, the BF-DNN+SUS effectiveness in the reliability analysis is compared with the MCS, MCS+SUS, and Co-Kriging+SUS.

3.4.3.1 Bi-fidelity mesh description

The mesh convergence must be studied and refined for accurate and valid results. However, a trade-off between precision and computational time is needed to save computing time through mesh convergence. The dimensions used in discretization of the mesh consist of: N_x is considered the number of divisions along the *x*-axis of the plate, N_y is the number of

divisions in the stiffener web (y-axis) of the stiffener, N_z is the number of divisions along the z-axis of the plate and stiffener, and N_{ts} the number of divisions along the x-axis of the flange stiffener. In the present study, $N_x=N_y$ is adopted.

The stiffened panel with the mean values of the material and geometric variables presented in Table 3.3 is used for the mesh convergence study. The number of divisions' variation influence and their combinations are shown in Fig. 3.6(a). For all analyses, $N_{ts}=2$ is adopted. These time measurements include the time to preprocess the FE model, adjust the finite element mesh to match the initial imperfection condition and solve the FE analysis. Four sets of division numbers are selected as LF models to evaluate the fidelities correlation presented in Table 3.7.
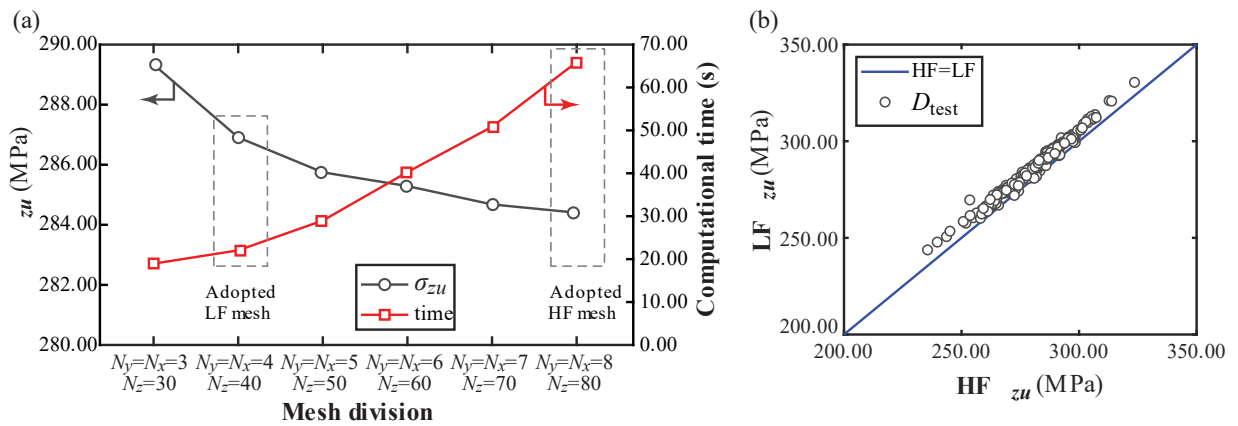
Figure 3.6 – Mesh assessment (a) convergence analysis and computational cost, and (b) Distance correlation between fidelities functions.

Table 3.7 – Fidelities correlation and cost functions.

Model Fidelity	$N_x=N_y$	N_z	Comp. Cost (s)	Ratio to HF	DIC
LFM1	2	20	16.200	0.2364	0.975
LFM2	3	30	17.300	0.252	0.983
LFM3	4	40	18.200	0.265	0.991
LFM4	5	50	27.500	0.400	0.992
High	8	80	68.500	1.000	-

The relationship of the mesh fidelities model is evaluated considering a test dataset D_{test} with $n_{test}=200$ to determine the DIC correlation parameter. Considering the information in Table 3.7, the LF models presented larger values of DIC, indicating a stronger relationship between each LF about the HF model. The LFM3 and LFM4 showed similar DIC values but different computational costs. Based on these presented results, the LFM2 is adopted as a reference LF model in the present study. Figure 3.6(b) shows the D_{test} points with the correlation between the HF and LF model. The cost of each fidelity model is the average ratio

of computational cost per sample given in Table 3.7, resulting in a cost function $c_{HF}=1.00$ and $c_{LF}=0.26$.

Figure 3.7(a) shows the stress-strain curve obtained by the mean value of the basic random variables for the HF and LF meshes adopted. The ultimate compressive strength, defined as the maximum value of the stress-strain curve considering the HF mesh, is $\sigma_{zu}=284.26$ MPa, and its deformed shape and von Mises stress distribution at collapse are shown in Fig. 3.7(b). The results obtained by the HF mesh in this paper adequately reflect the ones obtained by [2, 3, 8].
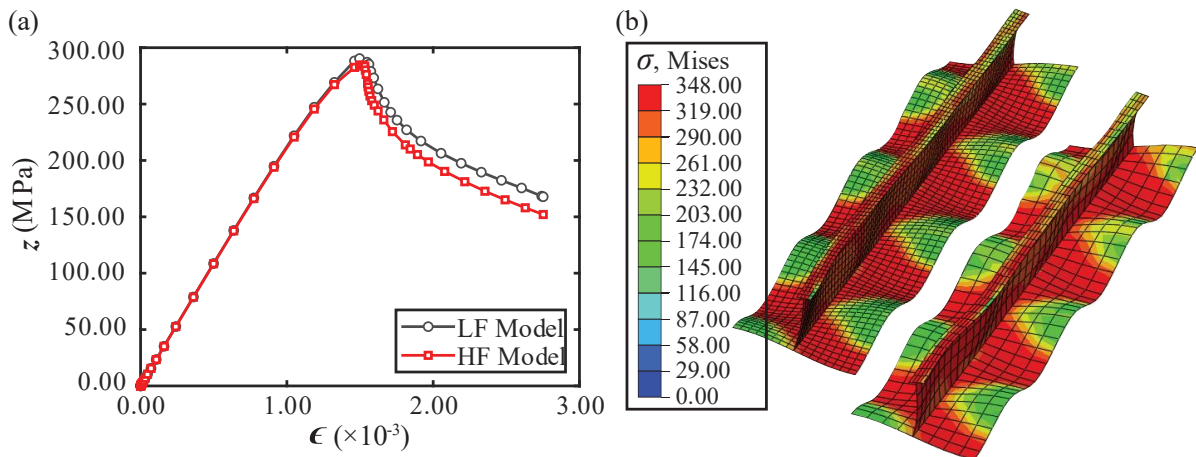
Figure 3.7 – FE results considering the mean value of the basic random variables: (a) Average stress-strain curve, and (b) von-Misses stress (MPa) for the HF and LF condition (displacement with scale factor 20×).

3.4.3.2 Accuracy and efficiency of the BF-DNN model

The predictive accuracy of the failure probability predictions depends on the σ_{zu} prediction of the stiffened panel elements. This accurate prediction, consequentially, depends on the number of support points and the characteristics of each multi-layer NN. Therefore, it is necessary to evaluate the training sample sizes n_{LF} and n_{HF} to obtain an accurate response surface BF approximation for this example. For the BF-DNN assessment, $n_{val}=200$ and $n_{test}=500$ are considered.

The n_{LF} value is the product of n_{HF} by a sample size ratio between HF and LF models (λ_{BF}) to obtain an accurate response surface approximation. The potential and impact of the LF sample points are evaluated with the HF samples fixed to $n_{HF}=500$, 750, and 1000 support points, while the infill strategy of LF samples is increased as the ratio λ_{BF} ranges from 2 to 6.

Figure 3.8 shows the influence of the sample size ratio on N_{call} to the performance function. By plotting the evolution of N_{call} of the performance function in terms of the sample ratio (Fig. 3.8(a)), it can be seen that for each value n_{HF}, the computational cost and N_{call}

increase at a linear rate of associated λ_{BF}. Each value of n_{HF} and comparing the extreme cases (λ_{BF}=2 and 6) N_{call} had a similar increase of 69% approximately.
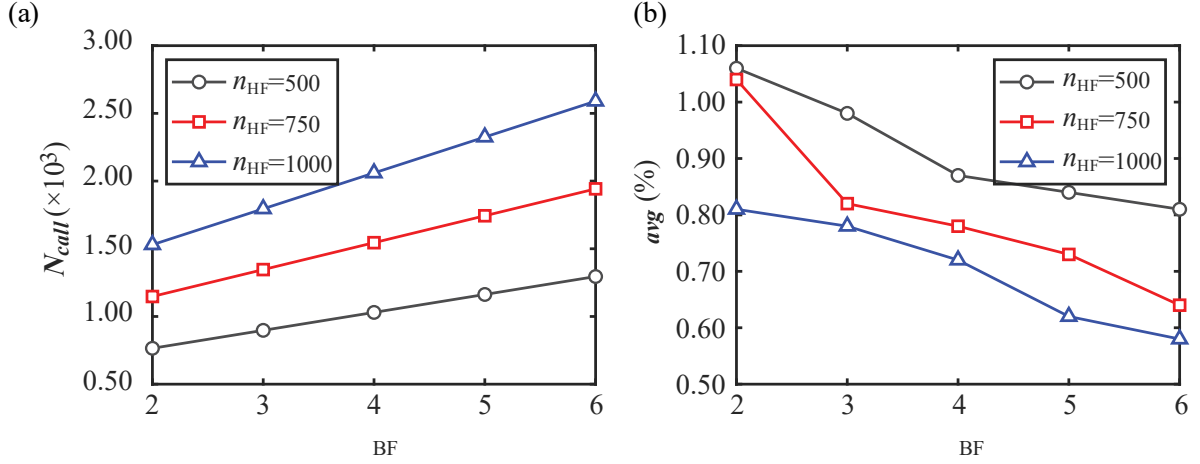
Figure 3.8 – The tendency of (a) the global N_{call} and (b) ε_{avg} accuracy for different sample ratios.

The accuracy metrics of the BF-DNN cases for performance function approximation are obtained with the sum of the metrics of each stage (NN_{LF} and NN_{BF}) of the BF-DNN model. The ε_{avg} (Eq. 3.18), presented in Fig. 3.8(b), is used as a fit measure of the BF-DNN surrogate model under different sample ratios λ_{BF}.

About the results presented in Fig. 3.8(b), considering the specific case n_{HF}=500, it is clear that by increasing LF samples for λ_{BF}>4, the BF-DNN models show a slight improvement in the value of the accuracy metrics. However, in BF models with n_{HF}=750 and 1000, it becomes apparent that the accuracy of the BF surrogate models improves rapidly when the rate λ_{BF} increase. The results show a trade-off between computational cost and accuracy.

The plots in Fig. 3.8 show that the computational cost and N_{call} of the cases n_{HF}=500 associated with high values of λ_{BF} and n_{HF}=750 related to low values of λ_{BF} are coinciding. It can also be seen when comparing the cases with n_{HF}=750 related to high values of λ_{BF} and n_{HF}=1000 linked to low values of λ_{BF}. Nevertheless, despite having cases with similar computational costs and N_{call}, cases with lower n_{HF} and higher values n_{LF} produced better accuracies. These results show the importance of the assessment of the relation λ_{BF}.

Considering the presented accuracy results, the case with n_{HF}=1000 and λ_{BF}=6 provides accurate metrics estimates and can be regarded as an appropriate BF-DNN surrogate model.

3.4.3.3 Comparison of the accuracy prediction with different BF surrogate

Computational experiments with another BF model are investigated to test the proposed method's performance. Results are compared mainly with Co-Kriging. The Co-Kriging comparison with BF-DNN considers the same initial training points (n_{HF}=1000 and λ_{BF}=6) shared between the two methods. The prediction for the test sample using the proposed BF-DNN and Co-Kriging is displayed in Figs. 3.9(a) and 3.9(b) and are linked with the parameters presented in Table 3.8. Based on the results, the proposed BF-DNN method showed a better adjustment when compared with the Co-Kriging method.
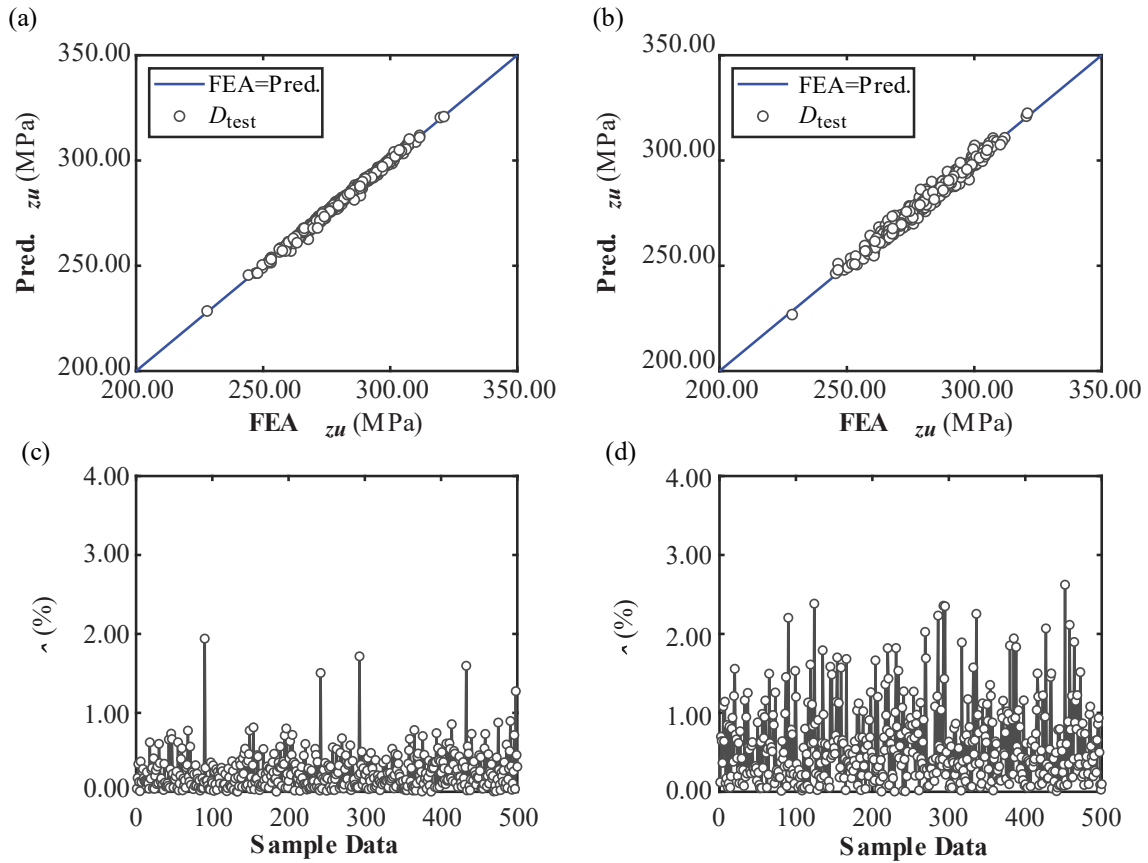
Figure 3.9 – Approximation of the σ_{zu} using (a) BF-DNN and (b) Co-Kriging; Relative errors $\varepsilon_{\hat{\sigma}}$ in the σ_{zu} using (c) BF- DNN and (d) Co-Kriging.

Table 3.8 – Reliability results and fit parameters to the best adjustments.

MF Model	R^2	$\varepsilon_{\hat{\sigma},avg}$ (%)	$\varepsilon_{\hat{\sigma},std}\left[\times10^{-3}\right]$	$\varepsilon_{\hat{\sigma},max}$ (%)	T_c (seg)
BF-DNN	0.9952	0.24	0.23	1.94	3813
Co-Kriging	0.9800	0.58	0.48	2.63	4283

Figure 3.9 and Table 3.8 report that the results obtained from BF models from these two methods present satisfactory results, but the result from the BF-DNN model is better than the Co-Kriging. This improvement in the accuracy metrics can be seen because the BF-DNN better fits the design points that predict low σ_{zu} values. Figures 3.9(c) and 3.9(d) illustrate the

relative errors for the test data sample by a graph. The selected BF-DNN surrogate model has an average relative error $\varepsilon_{\hat{\sigma},avg} = 0.24\%$, with a maximum prediction relative error of 1.94%. The low standard deviation means that the error of test predictions is clustered around the mean. For 99.0% of the test points, the relative error in the BF-DNN satisfies $\varepsilon_{\hat{\sigma}} = 1.00\%$. Distinctly, this error is satisfied for 84.0% of the test points in the Co-Kriging method. In the Co-Kriging results, the data outputs are more spread out with a small accuracy than the metrics obtained for the BF-DNN.

The reference T_c in Table 3.8 is the total computational cost considered to optimize the BF-DNN hyperparameters over BO 50 iterations and to estimate the hyperparameters in the Co-Kriging model. In this example, the HF and LF σ_{zu} function is highly non-linear, resulting in many samples for the surrogate model training. The BF-DNN T_c is slightly less than Co-Kriging. Due to many samples, the Co-Kriging scheme has a 7000×7000 covariance matrix, resulting in a high computational cost for the hyperparameters optimization.

3.4.3.4 Hyperparameter optimization process

For each combination between the proposed n_{HF} and λ_{BF}, the best DNN is determined by applying the BO for tuning and optimizing the hyperparameters. To create the BF-DNN model, the NN_{LF} and NN_{BF} stages must be trained to obtain HF outputs through a BF model. Therefore, two neural networks are trained simultaneously and combined into one through an element-wise operation of the layers in an intermediate position between the two stages. A summary of the best hyperparameter combinations for the BF-DNN models is shown in Table 3.9. Each hyperparameter is described for the two stages, NN_{LF} and NN_{BF}.

Table 3.9 – Tunned hyperparameters by Bayesian Optimization.

n_{HF}	λ_{BF}	Number of layers	Hidden Layer's Size	Activation Function	Learning rate [×10⁻³]
500	2	(3, 3)	[(175 53 163), (93 54 54)]	(elu, elu)	(7.5, 5.4)
	3	(3, 3)	[(167 129 195), (85 53 53)]	(elu, elu)	(17.8, 5.8)
	4	(3, 3)	[(112 85 196), (92 50 50)]	(elu, elu)	(17.4, 5.1)
	5	(2, 2)	[(187 161), (50 70)]	(elu, elu)	(13.7, 7.6)
	6	(3, 2)	[(51 52 157), (60 197)]	(elu, elu)	(19.9, 8.2)
750	2	(3, 2)	[(167 129 195), (54 127)]	(elu, elu)	(17.8, 6.1)
	3	(2, 2)	[(146 173), (175 67)]	(elu, elu)	(5.0, 13.1)
	4	(3, 2)	[(51 52 157), (131 50)]	(elu, elu)	(19.9, 5.1)
	5	(3, 3)	[(69 73 151), (196 61 61)]	(elu, elu)	(14.3, 8.2)
	6	(3, 2)	[(150 61 172), (53 113)]	(elu, elu)	(8.0, 5.6)
1000	2	(3, 2)	[(112 85 196), (200 156)]	(elu, elu)	(17.4, 5.0)
	3	(3, 2)	[(51 52 157), (61 122)]	(elu, elu)	(19.9, 5.0)
	4	(3, 3)	[(104 200 79), (51 73 73)]	(elu, elu)	(39.9, 5.1)
	5	(3, 3)	[(51 169 75), (111 65 127)]	(elu, elu)	(15.4, 8.7)
	6	(3, 2)	[(170 57 142), (115 199)]	(elu, elu)	(7.5, 5.1)

For each stage in the BF-DNN model, one activation function is optimized and used in all hidden layers present in these stages. In observed cases with Sigmoid and Hyperbolic tangent activation functions, the training is compromised because of the vanishing gradient problem. The ELU and RELU present similar performances, but the results obtained using ELU as an activation function are slightly better in all of the best observed feasible BF-DNN.

The selected BF-DNN model is developed by tuning its hyperparameters. The level of precision is used to determine the optimal model using the tuned parameters as given in Table 3.9. Figure 3.10 shows the progress of the BF-DNN hyperparameter optimization for the selected model. The plot in Fig. 3.10 corresponds to the minimum objective obtained for the D_{val} with the trained BF-DNN, considering the sum of the minimum value of the stage in each iteration corresponding to the minimum objective, as given by Eq. 3.8.
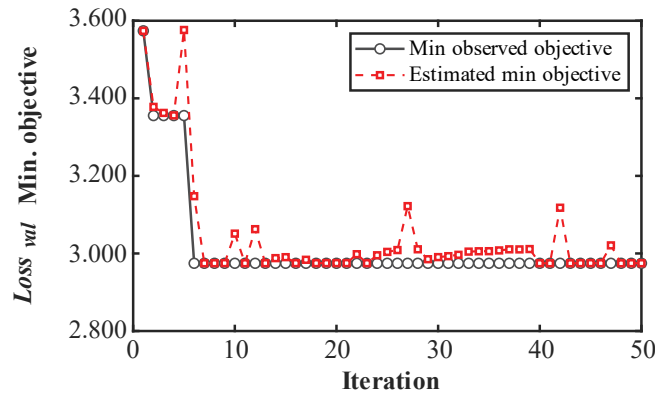
Figure 3.10 – The progress of Bayesian optimization for tuning hyperparameters of (a) NN$_{LF}$ and (b) NN$_{BF}$.

The red line in Fig. 3.10 corresponds to the estimated target values, i.e., the result of the surrogate model that determines where to evaluate next. The black line is the observed target value of the true objective function. Due to the agreement between the true and estimated values, it can be assumed that the estimated function is sufficiently compelling, attesting to BO's efficiency in choosing BF-DNN hyperparameters. The best workable point is where the observed objective value is the lowest. The score for the minimum objective of $Loss_{val}$=2.96 for the trained BF-DNN is kept at five iterations.

3.4.3.5 Reliability analysis using BF-DNN model

The BF-DNN model in the presented stiffened panel problem is also assessed with its ability to evaluate the failure probability based on the BF-DNN+SUS association. The FE analysis reproduces the implicit performance function (FE-PF). For the FE analysis in association with SUS (FEA+SUS), the initial N_s is set to 3×10^3 and $p_0 = 0.1$, also of the other intermediate failure events size. The process stops in the 7th iteration with $COV_{\hat{P}_f} = 0.092$. The values $P_{f,ref}$,

and β_{ref} are summarized in Table 3.10. The SUS analyses are submitted to 50 simulation cycles and $\hat{P}_f$ predictions. The BF-DNN+SUS and Co-Kriging+SUS results presented in Table 3.10 are discussed throughout this section.

Table 3.10 – Failure probabilities comparisons in Example 03.

Method	N_{call}	$\hat{P}_f$	$\hat{\beta}$	$\varepsilon_{\hat{\beta}}(\%)$
MCS+SUS (reference)	21000	6.77×10^{-6}	4.351	-
Co-Krigagem+ SUS	2590	5.826×10^{-6}	4.384	0.75
BF-DNN+SUS (proposed)	2590	6.932×10^{-6}	4.346	0.11

After each BF-DNN has been optimally tuned and successfully trained, the BF-DNN+SUS is performed to predict $\hat{\beta}$ value. Fig. 3.11(a) depicts $\hat{\beta}$ w.r.t. the sample ratio λ_{BF} for the n_{HF} variations. The average relative error $\varepsilon_{\hat{\beta},avg}$ is presented in Fig. 3.11(b).
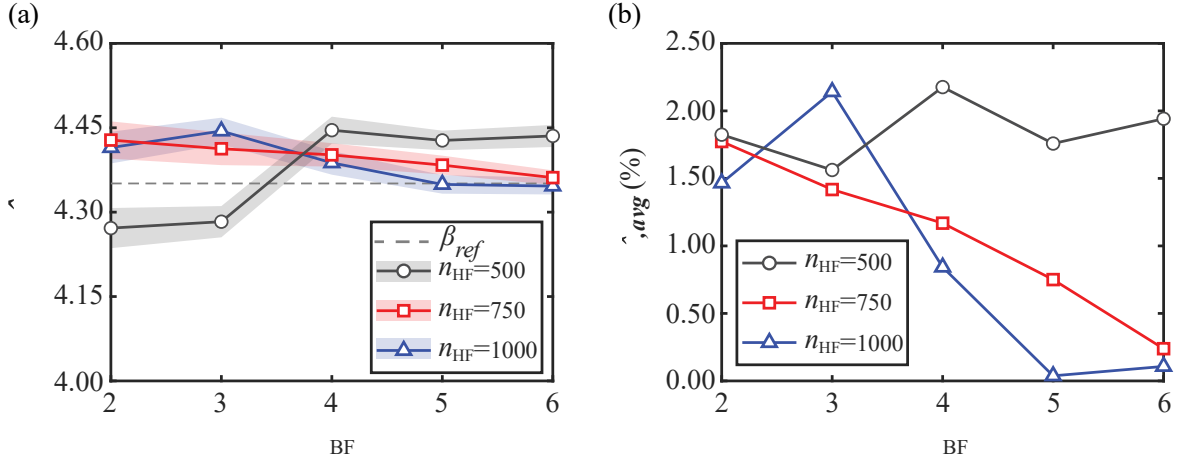
Figure 3.11 – Reliability index: (a) predicted β and (b) β relative error.

As stated in Fig. 3.11, the cases with n_{HF}=500 give dire predictions for β value when compared with cases with different values of n_{HF}. Despite slightly fluctuating $\hat{\beta}$ as the λ_{BF} rate increases, the models with n_{HF}=750 and 1000 drastically reduce the $\varepsilon_{\hat{\beta},avg}$ in the $\hat{\beta}$ prediction. However, it is essential to emphasize that although the BF-DNN model with n_{HF}=750 and λ_{BF}=6 presents good accuracy in the $\hat{\beta}$ prediction by SUS, it does not ensure satisfactory global accuracy results, as shown in Fig. 3.11(b). For the best circumstances, the best global σ_{zu} and PF prediction are obtained with n_{HF}=1000 and λ_{BF}=6, which offer a small variability $\varepsilon_{\hat{\beta},avg}=0.11\%$ compared to β_{ref}. As displayed in Fig. 3.11(a), CI becomes narrower as the number of λ_{BF} increases due to the effect of reducing the standard error of the failure probability estimations by BF-BFDNN+SUS.

The results of the proposed and other methods are shown in Table 3.10. The reduction of computing time is remarkable considering the proposed BF-Kriging+SUS and compared with the FEA-SUS. The proposed method reduces the N_{call} by about 88%, getting a similar value of $\hat{P}_f$. Comparing BF-DNN+SUS with BF-Kriging+SUS, the proposed method has a slightly better $\hat{P}_f$ prediction at similar N_{call} costs.

This approach combines the advantages of the BF-DNN model for approximating the true PF and the Subset Simulation for evaluating small failure probabilities. The BF-DNN inherits the advantage of the original DNN, but the evaluation of the response function using LF models in association with HF models reduces the computational cost.

3.5   CONCLUSIONS

This work presents a Bi-Fidelity Deep Neural Network to make explicit Performance Function for reliability analysis associated with Subset simulation. This framework's accuracy and computational efficiency are demonstrated by its application in reliability analysis in several scenarios.

The use of the low-fidelity sample in the bi-fidelity model has been proven. In the Finite Element problem, using the low-fidelity sample, there is a reduction of up to 75% of the computational cost in the FE analysis. The accuracy assessments are performed to define the best sample configuration that provides a fitted model that best represents the Performance Function.

Most surrogate-based procedures, as the compared Co-Kriging, are a victim of the curse of dimensionality that the computation effort increases dramatically since large populations are required to assess small probabilities or to fit non-linear Performance Function. Based on the results, the optimized BF-DNN solves this problem by simplifying the Neural Network architecture using a few Bayesian Optimization iterations.

Moreover, the BF-DNN+SUS association can express a small failure probability as a product of larger conditional failure probabilities of several intermediate failure events (See examples 2 and 3). However, even obtaining excellent results for low dimensions (See Example 2 with d=5), the time for network preprocessing can be considered expensive. This method for solving problems with low variables or non-complex PF should be considered with reservation. In addition, the proposed framework needs a large amount of data to train, becoming a limitation in a sparse high-fidelity dataset analysis.

These results suggest that the proposed bi-fidelity framework can be applied to the reliability analysis of complex structures, using few high-fidelity associated with many low-fidelity samples to predict high-fidelity values. This framework provides an acceptable level of accuracy provided by the Bayesian optimization of the hyperparameters and considerable computation time savings compared to the conventional methods that use only high-fidelity samples.

4 MULTI-TASK AND BI-FIDELITY DEEP NEURAL NETWORK ASSOCIATED WITH SUBSET SIMULATION FOR STRUCTURAL RELIABILITY ASSESSMENT

Abstract: A hyperparameter-optimized Bi fidelity and Multi-task Deep Neural Network (BFMT-DNN) using low-fidelity data samples added to the model to predict high-fidelity responses for structural collapse behaviour framework is presented in this paper. The assessment is realized in an offshore wind turbine in extreme conditions and described using non-linear Finite Element analysis to obtain multiple outputs. The low-fidelity model is accessed with two different models considering different FE elements. This multi-fidelity surrogate model efficiently reduces the high computational cost for highly non-linear and high dimensional structural reliability problems, and the multi-task framework allows training one model to predict multiple tasks. The proposed BFMT-DNN is associated with Subset Simulation to estimate rare events efficiently. The results show that the proposed framework, with its optimized hyperparameters using Bayesian Optimization, is an excellent strategy for reducing the number of samples used to construct the performance function's surrogate model. Moreover, the proposed framework can provide an accurate failure probability estimation with a lower computational cost in high non-linear, high dimensional, and rare events.

Keywords: Artificial Neural Networks; Multi-task; Multi-fidelity; Non-linear Finite Element Analysis; Wind Turbines.

4.1 INTRODUCTION

In 2021, the wind industry brought worldwide cumulative wind power capacity to 837 GW, showing year-over-year growth of 12%, only 1.8% behind the record year 2020 [126]. Even though the second year of the COVID-19 pandemic, the 93.6 GW of new installations added is a clear sign of the incredible resilience and upward trajectory of the global wind industry.

The main criteria, including the water depth, estimated environmental loads, the cost of production and installation, and the complexity of the design, are essential for choosing the types of support structures to accommodate the offshore wind turbine (OWT) [127]. Structural space frames, such as jackets, offer offshore sites a light and stiff option [128]. However, despite aleatoric and epistemic uncertainties in the dynamic loads and material properties, the OWT accurate life prediction is complex, and the computation analysis can be sensitive to small changes in the input parameters and, consequently, time-consuming [129].

Reliability analysis is of significant importance in structural engineering, as it aims to obtain the probability of failure of a system of interest under uncertainties that influence some event concerning some performance criterion [1-3, 7]. Engineers and researchers have well-recognized the importance of reliability in the past few decades, and its improvements have concentrated on making the performance functions more realistic, using proper strength and loading assessment methods based on numerical methods [4].

Small failure probabilities and complex failure boundaries involved in structural reliability problems can be computed efficiently using approximate methods such as the first-order or second-order reliability methods (FORM/SORM) (e.g. [7-9]) or using advanced simulation methods based on Monte Carlo simulation (MCS) with variance reduction techniques as MCS method with importance sampling (MCS-IS) (e.g. [10-12]), the directional IS (e.g. [11]), and the Subset Simulation (SUS) (e.g. [13-17]).

In recent years, many studies have been proposed to assess OWT reliability, such as [127, 130-132] . These studies show that the presence of uncertainty is a pertinent condition for evaluating the reliability of the OWT structure. Moreover, relevant limit states are presented in the OWT structural reliability analysis, such as fatigue, ultimate, and serviceability lifetime. In this context, more than one output can be necessary for the OWT reliability analysis.

Although surrogate models have been widely applied in designs to replace physical experiments or expensive simulations, they focus on single-fidelity models to reproduce the Performance (PF) function based on limited expansive calculations to obtain samples to reduce computational burden [8, 22, 85]. Are example of single-fidelity approaches: Polynomial Regression Models (e.g. [127, 128]), Kriging (e.g. [8, 85, 130]), Radial Basis Functions (RBF) (e.g. [22]), Polynomial Chaos Expansion (e.g. [21, 95]), Maximum Entropy (e.g. [133]) and Gaussian Processes (GP) (e.g. [96, 98]). Artificial Neural Networks (ANN) are amongst the most popular machine learning methods to generate the response surface and solve structural reliability problems. Chojaczyk *et al*. [2], Xu *et al.* [99], and Afshari *et al.* [100] summarize the main applications of single-fidelity ANN models in structural reliability and safety problems.

Hence, Multi-fidelity (MF) metamodel approaches have attracted significant attention recently for data regression which maximizes the use of cheaper low-fidelity (LF) models to predict high-fidelity (HF) outputs [28]. The Bi-fidelity (BF) strategy uses the correlation between an HF function and one LF function. The use of MF models has gained popularity in

structural engineering applications such as multidisciplinary, robust, and multi-objective optimization fields (e.g. [28, 29, 32, 56, 134]), uncertainty quantification fields (e.g. [34, 55, 112]), and structural reliability (e.g. [30, 33, 57]). Are examples of MF approaches: MF-Kriging (e.g. [28-30]), MF-RBF (e.g. [9, 31]), and MF-GP (e.g. [56, 134]).

The ability to handle large data sets and good generalization in precision have made ANNs a good candidate for MF regression. Several engineering fields aborded the MF-ANN to obtain HF responses, such as optimization (e.g.[56]); propagation (e.g. [125, 135]); and uncertainty quantification (e.g. [136, 137]).

The OWT reliability analysis eventually aims to confront situations where multiple tasks must be solved. The surrogates showed earlier obtain various outputs that are approximated one by one. In this context, Multi-tasking Learning solves tasks simultaneously [35-37]. Also, an ML model contains model parameters and hyperparameters. The model parameters are obtained by fitting the training data and defining the hyperparameters. Selecting an optimized model that archives the best performance on the data in a reasonable amount of time becomes a problem of optimizing these hyperparameters [101-103].

This paper aims to propose a Bi-Fidelity Multi-task Learning Model based on a Deep Neural Network (BFMT-DNN) for structural reliability analysis. The main contribution is developing a novel hyperparameter-optimized BFMT-DNN, considering the advantages of Bayesian Optimization, focusing on prediction accuracy, stability, and computational efficiency. As a numerical example, an offshore wind turbine with a jacket-tower structure sited in the south of Brazil is adopted and described using non-linear FE analysis to obtain the behaviour under extreme loads. The BFMT-DNN is used to predict the OWT's maximum stress and deflection. The LF model is constructed considering two scenarios varying the FE mesh size and element type. In addition, this paper presents a 50-year return period in extreme conditions of the offshore deep water over the south of Brazil shelf, based on WAVEWATCH III hindcast data with a real buoy measurements validation.

This paper is organized as follows: Section 2 introduces related concepts of Deep Neural Networks to elaborate the proposed BFMT-DNN model; Section 3 discusses the parametric FEA built for the OWT, highlighting the geometry, mesh, material, and boundary conditions used, as well as the validation of the FEA model; The extreme load estimation for the OWT through statistical extrapolation methods is also described in Section 3; In section 4, results obtained from the study are discussed extensively; Section 5 draws the main conclusions of the work conducted.

4.2 A BI-FIDELITY AND MULTI-TASK DEEP NEURAL NETWORK FRAMEWORK

This section outlines the computational framework of the Bi-fidelity Multi-task Deep Neural Network (BFMT-DNN) to assess structures' uncertainty quantification and failure probability. This network intends to employ a small amount of HF data and a large amount of LF data, both generated by FEA, to train the surrogate model. The BFMT-DNN is proposed to predict the maximum stress and the deflection in wind turbine support structures, learning in a two-stage method. The problem has been chiefly modeled concerning a single-task learning framework.

4.2.1 Bi-fidelity surrogate model

The variable fidelity analysis is essential for tackling computational costs using a combination of LF and HF data. Although LF samples are noisy and somewhat inaccurate, they usually correlate strongly with HF samples [78]. As such, it may be possible to leverage this correlation to avoid total reliance on HF data correcting the HF model using LF information.

The combinations of the fidelities can be categorized into three groups [115]: (*i*) Adaptation approaches that use adaptation to enhance LF models with information from HF models while the computation proceeds; (*ii*) Fusion strategies that evaluate LF models and HF models and then combine information from all outputs and (*iii*) Filtering approaches that use the HF model if the LF model is inaccurate, or when the candidate point meets some criterion.

Currently, there are several available correction methods, which can be briefly divided into three categories [116]: (*i*) Additive and multiplicative corrections that construct a surrogate model of the difference or the ratio between the HF and the LF models; (*ii*) Comprehensive corrections, where both corrections (additive and multiplicative) are used in the same bi-fidelity surrogate model; and (*iii*) Space mapping (input correction) where instead of correcting the output of the LF model, it is also possible to correct the input variables.

Suppose an *n*-dimensional random vector $\mathbf{x} \in \mathbb{R}^d$ is mapped through a model to obtain a desired output $y(\mathbf{x})$. Let $y_{LF}(\mathbf{x})$ and $y_{HF}(\mathbf{x})$ denote the approximated values of $y(\mathbf{x})$ by an LF and HF computational model, respectively. In the current study, a fusion MF with a comprehensive correction model adopted is expressed as follows:

$$y_{HF}(\mathbf{x}) = \rho(\mathbf{x}) \cdot y_{LF}(\mathbf{x}) + z(\mathbf{x})$$
<div align="right">Equation Chapter (Next) Section 1(4.1)</div>

where $\rho(\mathbf{x})$ is a scaling factor that quantifies the correlation between [y_{HF}, y_{LF}], and $z(\mathbf{x})$ the corresponding noise. However, one of the main disadvantages of this scheme is that it can

only handle linear correlation between two fidelity data. In contrast, many interesting cases follow non-linear relationships between LF and HF data [107].

4.2.1.1 The Fidelities correlation considering Distance Correlation (DIC)

Although LF samples are noisy and skewed, they usually correlate strongly with HF samples [78]. As such, leveraging this correlation to avoid total reliance on HF data is possible. The empirical Distance Correlation (DIC) is introduced by Székely *et al*. [81, 82] to measure associations between two random vectors, not necessarily in the same dimension. As an advantage over the classical Pearson product-moment correlation, the DIC has a significant benefit for characterizing a non-linear relationship of random vectors. Moreover, DIC satisfies $0 \leq \mathrm{DIC}\left(y_{\mathrm{HF}}, y_{\mathrm{LF}}\right) \leq 1$ and equals zero if independence holds. Given a sample $\left(y_{HF}^{(i)}, y_{LF}^{(i)}\right)_{i=1}^{N}$ of N observations of the joint random vector (Y_{HF},Y_{LF}), is defined the double-centered distance matrices *A* and *B* as follows:

$$A_{i,j} = a_{i,j} - \frac{1}{N}\sum_{l=1}^{N} a_{il} - \frac{1}{N}\sum_{k=1}^{N} a_{kj} + \frac{1}{N^2}\sum_{k,l=1}^{N} a_{kl}, \text{ and} \tag{4.2}$$

$$B_{i,j} = b_{i,j} - \frac{1}{N}\sum_{l=1}^{N} b_{il} - \frac{1}{N}\sum_{k=1}^{N} b_{kj} + \frac{1}{N^2}\sum_{k,l=1}^{N} b_{kl}, \tag{4.3}$$

where $a_{ij} = \left\| y_{HF}^{(i)} - y_{HF}^{(j)} \right\|_p$ and $b_{ij} = \left\| y_{LF}^{(i)} - y_{LF}^{(j)} \right\|_q$. Then, the sample dcov is the square root of

$$\mathrm{dcov}^2 = \frac{1}{N^2}\sum_{i,j=1}^{N} A_{i,j}B_{i,j}. \tag{4.4}$$

Similarly, the sample distance correlation is the standardized sample covariance given by:

$$DIC^2\left(y_{\mathrm{LF}}, y_{\mathrm{HF}}\right) = \begin{cases} \dfrac{\mathrm{dcov}^2\left(y_{\mathrm{LF}}, y_{\mathrm{HF}}\right)}{\sqrt{\mathrm{dcov}^2\left(y_{\mathrm{LF}}, y_{\mathrm{LF}}\right)\mathrm{dcov}^2\left(y_{\mathrm{HF}}, y_{\mathrm{HF}}\right)}} & \text{if } \mathrm{dcov}^2\left(y_{\mathrm{LF}}, y_{\mathrm{LF}}\right)\mathrm{dcov}^2\left(y_{\mathrm{HF}}, y_{\mathrm{HF}}\right) > 0 \\ 0 & \text{if } \mathrm{dcov}^2\left(y_{\mathrm{LF}}, y_{\mathrm{LF}}\right)\mathrm{dcov}^2\left(y_{\mathrm{HF}}, y_{\mathrm{HF}}\right) = 0. \end{cases} \tag{4.5}$$

4.2.2 Artificial neural networks for multi-task learning

Machine Learning (ML) uses mathematical algorithms that can recognize patterns and similarities in data and use their information to be trained and make predictions about new data they have never seen before. As a subset of ML, Deep Learning (DL) recreate the brain's structure by creating a network of neurons. As a successful branch of DL, Artificial Neural Networks (ANN) apply these concepts to solve and simulate real-world problems and

scenarios teaching computers to model an unknown function $f(x)$ from available datasets and algorithms.

The number of layers, neurons, and activation functions specifies the ANN. The neurons consist of internal model parameters weights and biases, the values of which are optimized by repeated learning of data [2, 138]. Single Task Learning (ST) updates the neural networks internal model parameters using input sequence from only one task involving one dataset. Multi-tasking (MT) learning aims at solving more than one problem simultaneously.

Given t learning tasks $\left\{T^{(i)}\right\}_{i=1}^{t}$ where all the tasks or a subset of them are related. In the MT Soft Parameter Sharing, each task has its model with some sharable parts. In the MT Hard Parameter Sharing approach, the model shares the hidden layers across all tasks.

This study uses the multi-layer feed-forward perceptron (MLP) to train the data [2, 100]. The Bi-fidelity Multi-Task Deep Neural Network (BFMT-DNN) is a customized MT Hard Parameter Sharing fully connected structure to capture the relationship between the different fidelities and tasks. Considering a multi-task scenario, the non-linear relationship between LF and HF data can be considered as follows:

$$\mathbf{y}_{\text{HF}}\left(\mathbf{x}_{\text{HF}}\right) = F\left(\mathbf{x}_{\text{HF}}, \mathbf{y}_{\text{LF}}\left(\mathbf{x}_{\text{HF}}\right)\right) \tag{4.6}$$

where $F(\cdot)$ is an unknown function that maps the LF data to the HF level that is based on the input data $\left\{\left(\mathbf{x}_{\text{HF}}^{(i)}, \mathbf{y}_{\text{LF}}\left(\mathbf{x}_{\text{HF}}^{(i)}\right)\right) : 1 \leq i \leq n_{\text{HF}}\right\}$, $\mathbf{y}_{\text{LF}}(\mathbf{x})$ is the LF vector data obtained by T_{LF} outputs, and $\mathbf{y}_{\text{HF}}(\mathbf{x})$ is the HF data vector obtained by T_{HF} outputs. The BFMT-DNN learns the unknown implicit function $F\left(\mathbf{x}_{\text{HF}}, \mathbf{y}_{\text{LF}}\left(\mathbf{x}_{\text{HF}}\right)\right)$ that maps the LF data to the HF level of Eq. 4.1 in the same structure, as depicted in Fig. 4.1.

The DNN architecture consists of five levels trained simultaneously: input layer, LF hidden layers, concatenate layer, BF hidden layers, and output layer. To capture the multi-task learning scenario, MT-DNN aims to help improve the learning of a model for task $T^{(i)}$ by using the knowledge in some or all of the tasks [36, 37]. According to Fig. 4.1, the first stage (NN_{LF}) learns $F_{\text{LF}}(\mathbf{x}_{\text{LF}})$. The $F_{\text{LF}}(\mathbf{x}_{\text{LF}})$ function outputs t learning LF tasks $\left\{T_{\text{LF}}^{(i)}\right\}_{i=1}^{t}$. Before that, it is possible to predict the values of the LF function for each T_{LF} at the training inputs $\mathbf{x}_{\text{HF}}$ of the HF data inputted in the concatenate layer denoted by $F_{\text{LF}}^{NN}\left(\mathbf{x}_{\text{HF}}^{(i,t)}\right) : 1 \leq i \leq n_{\text{HF}}$. Then the second stage (NN_{BF}) approximates the HF outputs $\mathbf{y}_{\text{HF}}(\mathbf{x})$ based on the input data

$$\left(\mathbf{x}_{\text{HF}}, F_{\text{LF}}^{NN} \left(\mathbf{x}_{\text{HF}} \right) \right) = \left\{ \left(\mathbf{x}_{\text{HF}}^{(i)}, F_{\text{LF}}^{NN} \left(\mathbf{x}_{\text{HF}}^{(i)} \right) \right) : 1 \leq i \leq n_{\text{HF}} \right\}$$ and the available HF output data

$$\mathbf{y}_{\text{HF}} \left(\mathbf{x}_{\text{HF}} \right) = \left\{ \mathbf{y}_{\text{HF}}^{(i,t)} : 1 \leq i \leq n_{\text{HF}} \right\}.$$
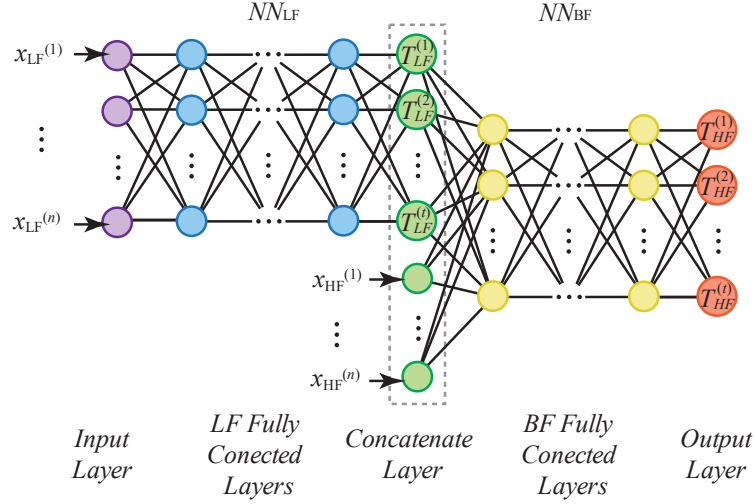
Figure 4.1 – Architecture of the composite neural networks for the BFMT-DNN model.

In the ANN, the neuron presented in Fig. 4.1 is a processing element with several inputs and one output. Each m neuron receives an input signal vector $\mathbf{x} = \left\{ \mathbf{x}^{(i)} : 1 \leq i \leq n \right\}$ from n input channels. The neurons are connected, the connection weights w, and each neuron contains a bias and an activation. Next, the weighted sum of x is calculated by multiplying each element x_c by a coefficient w_{mc}, demonstrating the importance of input channel c. The activation a_m of the m-neuron is given by:

$$a_m = \sum_{c=1}^{n} w_{mc} x_n + b_m \tag{4.7}$$

where $b_m \in R$ is the bias, is a constant corrective term which allows having a non-negative activation a_m. The output signal value y is calculated as a function of the activation. Table 4.1 presents some activation functions commonly applied in ANN. The code described is used as a reference in the hyperparameters fine-tuning.

4.2.2.1 Hyperparameter optimization based on Bayesian Optimization

The optimization of the hyperparameters is employed to increase the bounds of the fitting ability of the machine learning model. Therefore, the main optimization courses mainly include the structure of the model, the model training mode, and the hyperparameters of the model aspects [139].

Table 4.1 – Features of activation functions.

Activation Function	Function	Domain
Hyperbolic tangent	$f(a_m) = \left(e^{a_m} - e^{-a_m} \right) \big/ \left(e^{a_m} + e^{-a_m} \right)$	$(-1, 1)$
Sigmoid	$f(a_m) = 1 \big/ \left(1 + e^{-a_m} \right)$	$(0, 1)$
Rectified Linear Units (ReLU)	$f(a_m) = \begin{cases} a_m & a_m \geq 0 \\ 0 & a_m < 0 \end{cases}$	$[0, \infty)$
Exponential Linear Unit (ELU)	$f(a_m) = \begin{cases} a_m & a_m \geq 0 \\ \omega\left(e^{a_m} - 1 \right) & a_m < 0 \end{cases}$ with parameter $\omega = 1$	$(-\omega, \infty)$

Let *A* denote a machine learning algorithm with a configuration space of the overall hyperparameters *Λ*. The space *Λ* can include both discrete and continuous dimensions. Given a dataset D_{tr}, the goal is to find the best configuration such that:

$$\lambda^* = \max_{\lambda \in \Lambda} E_{(D_{tr}) \sim D} L\left(A_\lambda, D_{tr} \right) \tag{4.8}$$

where $L(A_\lambda, D_{tr})$ denotes the validation performance of A_λ on the dataset D.

The unknown hyperparameters are decided by minimizing the following loss function using Root Mean Squared Error (RMSE) of the LF and HF D_{val} as follow:

$$Loss = \sum_{i=1}^{t} \left[\sqrt{\frac{1}{n_{LF}} \sum_{i=1}^{n_{LF}} \left(\left| \hat{y}_{LF}^{(i)} - y_{LF}^{(i)} \right| \right)^2} + \sqrt{\frac{1}{n_{HF}} \sum_{j=1}^{n_{HF}} \left(\left| \hat{y}_{HF}^{(j)} - y_{HF}^{(j)} \right| \right)^2} \right], \tag{4.9}$$

where $\hat{y}_{LF}$, y_{LF}, $\hat{y}_{HF}$, and y_{HF} denotes, respectively, the predicted value of the validation LF in NN_{LF} stage, the real response of the LF point , the predicted value of the validation HF in NN_{BF} stage, and the real response of the validation HF point of the *t-th* task. Once all training data are introduced, the backpropagation optimization is utilized to update the weights and biases of the model iteratively until the loss function reaches the minimum [111].

The loss function can be optimized using a method such as the ADAM method along with Xavier's initialization method [56, 111]. The ADAM is a method for efficient stochastic optimization that only requires first-order gradients with little memory requirement, is relatively easy to implement, computationally efficient, and is well suited for problems that are large in terms of data or parameters. Moreover, the method is designed to combine the advantages of two popular methods for gradient-based optimization: Adaptive Subgradient Method, which works well with sparse gradients, and Root Mean Squared Propagation, which works well in online and non-stationary settings [117, 118].

Various optimization techniques can obtain an optimal combination of hyper-parameters. Traditionally, manual or automatic techniques (such as random and grid search) have been used as hyperparameter optimization [102]. The manual method tends to human error and needs expertise. The accuracy in grid search decreases with the increase in the number of parameters in the optimization process. Moreover, random search is based on random distribution functions, which might miss optimal spots in the investigation. To overcome the difficulty in heuristics associated with machine learning models, Bayesian Optimization has been leveraged in the hyperparameters selection of DNN [102, 103, 119].

Bayesian optimization (BO), derived from the Bayes theorem, is an effective method for solving computationally demanding optimization without closed-form solutions [120, 121]. The BO algorithm adds sample points to the posterior distribution of an objective function $L(A_\lambda, D_{tr})$, approximated by a Gaussian process [101, 122]. The GP is a function where the variable is a Gaussian distribution:

$$f(h) \sim GP\left(m(h), k\left(h, h'\right)\right). \tag{4.10}$$

Here $m(h)$ is the distribution's mean function, and $k(h, h')$ is the covariance function of two tested points, h and h'. In this study, the Matérn 5/2 kernel can be helpful in the optimization setting [101]. This proposed study chooses the Expected Improvement as an acquisition function because it takes work to fall into the local optimum solution [102]. Considering an acquisition function, where high values of the acquisition function are associated with high values of the objective function, the maximum value of the objective function is obtained through an iterative process of maximizing the acquisition function at each step of Bayesian optimization. This process continues until the posterior distribution fits the actual distribution.

The BO algorithm as hyperparameter optimization requires few iterations and a high operation speed. In addition, a strong point of the BO algorithm is that it can optimize integers and categorical variables, typical in DNN hyperparameters. The detail of the BO algorithm can be found in [101, 102, 121].

In this study, for all examples and both DNN stages, the hyperparameters: the number of hidden layers, neurons in each hidden layer, initial learning rate, and activation function, are tuned by BO. The hyper-parameters that required optimization (and their ranges) are given in Table 4.2.

Table 4.2 – The hyperparameters to be optimized with a defined search space.

Hyperparameter	Limits	Domain
Each Layer Size	[20,300]	Integer
Number of layers (NN_{LF})	[1,4]	Integer
Number of layers (NN_{BF})	[1,3]	Integer
Activation Function	[elu, relu, sigmoid, tanh]	Categorical
Learning rate	[5e-3,5e-2]	Real

4.2.3 Subset simulation for failure probability estimation

The performance function $G(\mathbf{x})$ provides the limit state function when $G(\mathbf{x})=0$, which separates the safe region $G(\mathbf{x})>0$ from the failure region $G(\mathbf{x})<0$. Mathematically, for a given structural system input parameter $\mathbf{x} = \left\{ \mathbf{x}^{(i)} : 1 \le i \le n \right\}$, the parameter P_f can be estimated by solving the multifold probability integral defined as:

$$P_f(\mathbf{x}) = P\left(F = G(\mathbf{x}) < 0 \right) = \int_F f(\mathbf{x})\,dx \tag{4.11}$$

where $f(\mathbf{x})$ is the joint probability density function of $\mathbf{x}$. Considering the statistical technique of Direct Monte Carlo Sampling (MCS), the failure probability can be estimated as follows:

$$P_f = \frac{1}{N} \sum_{i=1}^{N} I_F\left(\mathbf{x}^{(i)} \right) \tag{4.12}$$

where $\mathbf{x}^{(1)}, \ldots \mathbf{x}^{(N)}$ are i.i.d. samples from $f(\mathbf{x})$, and $I_F(\mathbf{x})$ stands for the indicator function, i.e.,

$$I_F(\mathbf{x}) = \begin{cases} 1 \text{ if } x \in F \\ 0 \text{ if } x \notin F \end{cases} \tag{4.13}$$

A brief introduction of the SS technique, largely following Ref. [13], is expressed. The Subset Simulation (SUS) is an adaptive Monte Carlo method proposed for estimating small failure probabilities in high-dimensional problems. The SUS is based on nested sets $F_1 \supset F_2 \supset \cdots F_M$, where $F_M = F$ is the failure event, and expresses the failure probability P_f as a product of conditional probabilities:

$$\hat{P}_f = P(F) = P_{F_1} \prod_{j=2}^{M} P(F_j \mid F_{j-1}) \tag{4.14}$$

where P_{F_1} is the first unconditional failure probability computed as the fraction of samples exceeding the current threshold level L_1, and $P(F_j \mid F_{j-1})$ are the subsequent conditional failure probabilities conditional on exceeding the prior intermediate thresholds in level L_{j-1}, and are computed as the fraction of samples exceeding the threshold level L_j.

In expressing $\hat{P}_f$ as a product of larger failure probabilities, SUS creates intermediate failure thresholds L before the required zero thresholds. As for SUS, the algorithm starts with a direct MCS to estimate P_{F_1}, while a Markov chain Monte Carlo is used to define $P_{F_j|F_{j-1}}$.

In SUS, the values $F(\mathbf{x}^{(k)})$, $k = 1, ..., N_n$, of the N_n samples are in the subset $S^{(1)}$. Crude MCS generates the samples independently identically distributed assuming the Nataf transformation, whereas samples in the subsequent subsets are generated by Markov Chain Monte Carlo (MCMC) simulation and correlated, considering the dependence of the seeds for each Markov chain. The samples falling between two subsequent intermediate failure thresholds, L_{j-1} and L_j, constitute a subset. The intermediate failure thresholds must be specified to estimate the intermediate failure probabilities. The range of the samples $N_s = p_0 N_n$ is retained in each step and serves as the seed for the Markov chains. The value $p_0 = 0.1$ is often used in the literature, which makes F_1 a relatively frequent event.

The authors Au and Beck [13] proposed a Metropolis-Hastings algorithm, and this method is popular for simulating the conditional samples in $P_{F_j|F_{j-1}}$, but occasionally it can lead to degenerate sampling when dealing with models having geometrically complex performance functions. Papaioannou *et al.* [15] proposed a delayed rejection Markov chain Monte Carlo method to alleviate these issues and is used in this study.

In addition, an approximate COV to estimate the intermediate failure probability in the subset is given by [16]:

$$\hat{\delta}_j = \begin{cases} \sqrt{\left(1 - P_{F_1}\right)\big/\left(P_{F_1} N_n\right)} & \forall \quad j = 1 \\ \sqrt{\left(1 - P_{F_1|F_{j-1}}\right)\big/\left(P_{F_1|F_{j-1}} N_n\right)} & \forall \quad 1 < j \leq N_s \end{cases} \tag{4.15}$$

where N_n is the number of samples in a subset. The overall COV estimate over the required failure probability is:

$$COV_{\hat{P}_f} = \sqrt{\sum_{i=1}^{N_S} \hat{\delta}_i^2}. \tag{4.16}$$

The generalized reliability index β is expressed as:

$$\beta = -\ ^{-1}\ \textbf{\textit{P}}_f \tag{4.17}$$

where $^{-1}\ *$ is the inverse CDF function of the standard Gaussian distribution.

4.2.4 Limit state design criteria using BFMT-DNN+SUS

A novel BFMT-DNN framework is developed in Stage 01, according to Fig. 4.2. For the development of the BFMT model for stress and deflection prediction, a surrogate metamodel is associated with an FEA. Stage 02 presents the structural reliability assessment. The steps of the proposed algorithm may be summarized in the described steps.
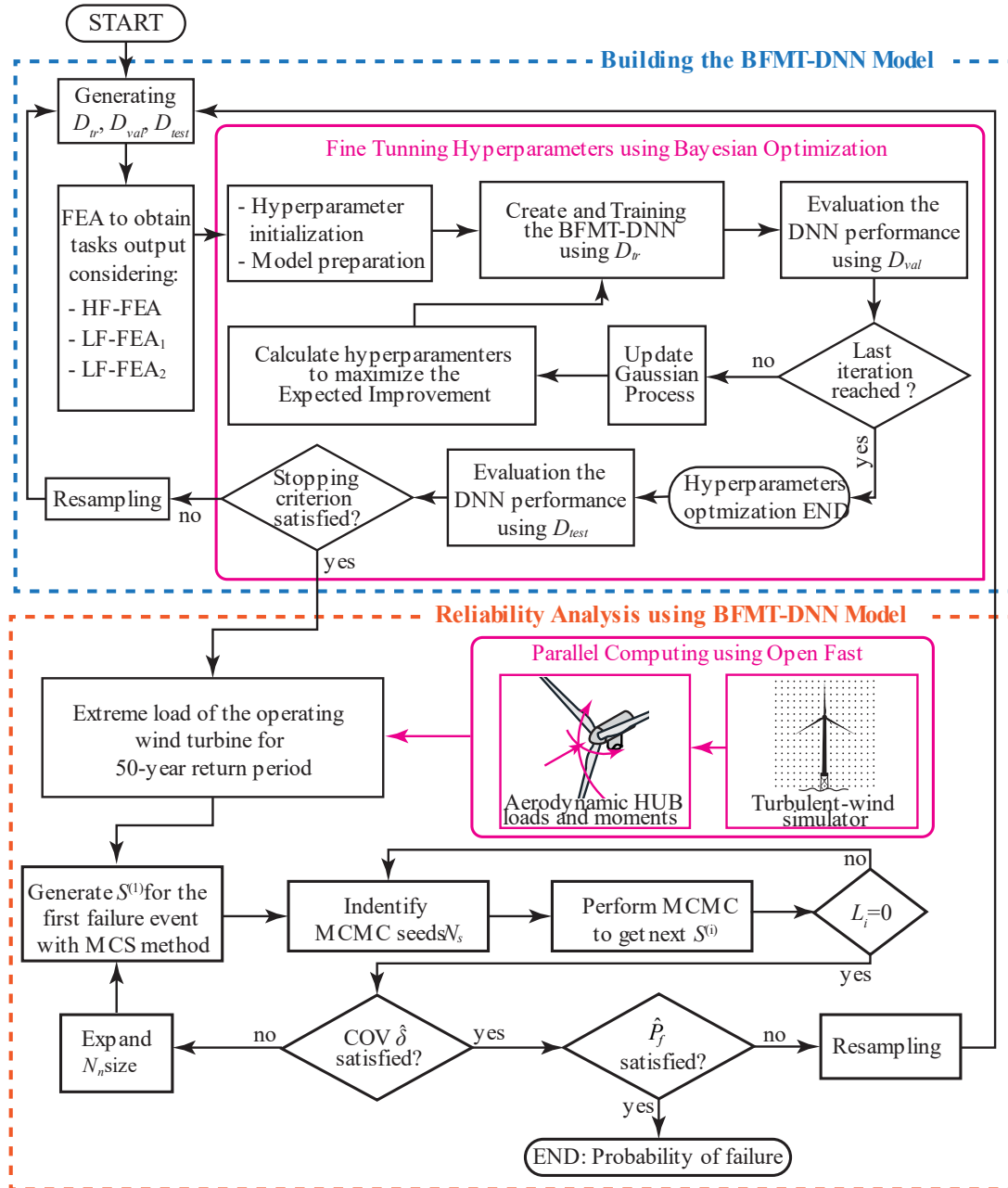

Figure 4.2 – Flowchart of the BFMT-DNN model and optimization process.

Stage 01 - Step 1: Three different datasets $D = \{D_{tr}, D_{val}, D_{test}\}$ are adopted. Namely, a training dataset $D_{tr,\text{HF}} = \left\{ \left(\mathbf{x}_{tr,\text{HF}}^{(i)}, \mathbf{y}_{tr,\text{HF}}^{(i)} \right) : 1 \leq i \leq n_{\text{HF}} \right\}$ and $D_{tr,\text{LF}} = \left\{ \left(\mathbf{x}_{tr,\text{LF}}^{(i)}, \mathbf{y}_{tr,\text{LF}}^{(i)} \right) : 1 \leq i \leq n_{\text{LF}} \right\}$ is used to estimate the gradients during the training of the neural networks. A holdout validation dataset $D_{val} = \left\{ \left(\mathbf{x}_{val}^{(i)}, \mathbf{y}_{val,\text{HF}}^{(i)}, \mathbf{y}_{val,\text{LF}}^{(i)} \right) : 1 \leq i \leq n_{val} \right\}$ is applied to optimize (fine tuning) the model's hyperparameters, and a test dataset $D_{test} = \left\{ \left(\mathbf{x}_{test}^{(i)}, \mathbf{y}_{test,\text{HF}}^{(i)}, \mathbf{y}_{test,\text{LF}}^{(i)} \right) : 1 \leq i \leq n_{test} \right\}$, never seen during training and validation, estimates the performance of the final tunned model. The database is generated randomly by Latin Hypercube sampling.

Stage 01 - Step 2: Initialize the model and define the initial hyperparameters.

Stage 01 - Step 3 and Step 4: Create and train the BFMT-DNN model using the D_{tr} dataset. In each iteration, the training step is validated using Holdout Validation. After, evaluate the BFMT-DNN performance using D_{val}. If the last iteration is reached, the hyperparameters optimization ends and turns to Stage 03-Step 7. Otherwise, turn to Stage 01-Step 5.

Stage 01 - Step 5 and Step 6: Construct the Gaussian surrogate model and select new data points by optimizing the acquisition function, calculating the corresponding functions, and augmenting the data. Repeat Steps 3 and 4 with the latest data.

Stage 01 - Step 7: Evaluate the BF-DNN performance using D_{test}. If convergence criteria are satisfied, the BF-DNN construction is finished. To quantify the accuracy of the proposed BFMT-DNN model is used the sum of the average RMSE (ε_{avg}) referring to the two stages of the BF-DNN is given as:

$$\varepsilon_{\text{avg}} = \sum_{i=1}^{t} \left[\sqrt{\frac{1}{n_{test}} \sum_{i=1}^{n_{test}} \left(\left| \hat{y}_{\text{LF}}^{(i)} - y_{test,\text{LF}}^{(i)} \right| \right)^2} + \sqrt{\frac{1}{n_{test}} \sum_{j=1}^{n_{test}} \left(\left| \hat{y}_{\text{HF}}^{(j)} - y_{test,\text{HF}}^{(j)} \right| \right)^2} \right], \tag{4.18}$$

where $y_{test,\text{LF}}$ and $y_{test,\text{HF}}$ denotes the LF and HF reference value of the D_{test}, $\hat{y}_{\text{LF}}$ and $\hat{y}_{\text{LF}}$, the LF and HF predicted values by BFMT-DNN. It is considered $\varepsilon_{\text{avg}}=1.0\%$ as the stop criterion.

Stage 02 – Step 1: Define the 50-year extreme values of the wave and wind conditions. In addition, are defined the extreme loads of the OWT hub.

Stage 02 – Step 2: Generate the first random sample data $S^{(1)}$ using the MCS method for each task. The value of N_n depends on the different problems, and it is much smaller than the number required by the original MCS method. Approximately samples are needed to estimate a failure probability by MCS-based methods. In this paper, the initial N_s is generally set to 1×10^3, also of the other intermediate failure events size.

Stage 02 – Step 3 and Step 4: In ascending order, select the first 10% random failure points and corresponding minimum responses as seeds for each Markov chain. Use the MCMC method to generate the new random sample data $S^{(i)}$.

Stage 02 – Step 5: If $L_i \geq 0$, turn to Stage 03-Step 3. Otherwise, set $L_i=0$.

Stage 02 – Step 6: Calculate the coefficient of variation $COV_{\hat{P}_f}$ based on Eq. 4.16. If $COV_{\hat{P}_f} \leq 0.1$, turn to Stage 03-Step 7. Otherwise, turn to Stage 02-Step 8.

Stage 02 – Step 7: Expand the $N_n = N_n+1\times10^3$ size of the random point and then go to Stage 03-Step 2.

Stage 02 – Step 8: Output the final $\hat{P}_f$. If $\hat{P}_f$ is satisfied, end the proposed method. Otherwise, resample the initial dataset D and turn to Stage 01-Step 1.

4.3 OFFSHORE WIND TURBINE IN EXTREME CONDITIONS

Offshore wind turbines (OWTs) are exposed to harsh marine environments. The complex and high-demanding environment with non-linear interactions and high variability of loading conditions exposes the OWT to uncertainties in the environmental loads. The successful deployment of OWTs widely depends on the correct stochastic load's effects estimation and the accurate prediction of the components' integrity throughout their service life.

4.3.1 Site and measurements

The region of this study is an offshore deep-water site near Rio Grande, a city located in Rio Grande do Sul (RS) state in Brazil. The location has a NE-SW coastline orientation (Fig. 4.3), and it is characterized by a large shelf (from 110 km to 170 km) with a smooth slope [140].
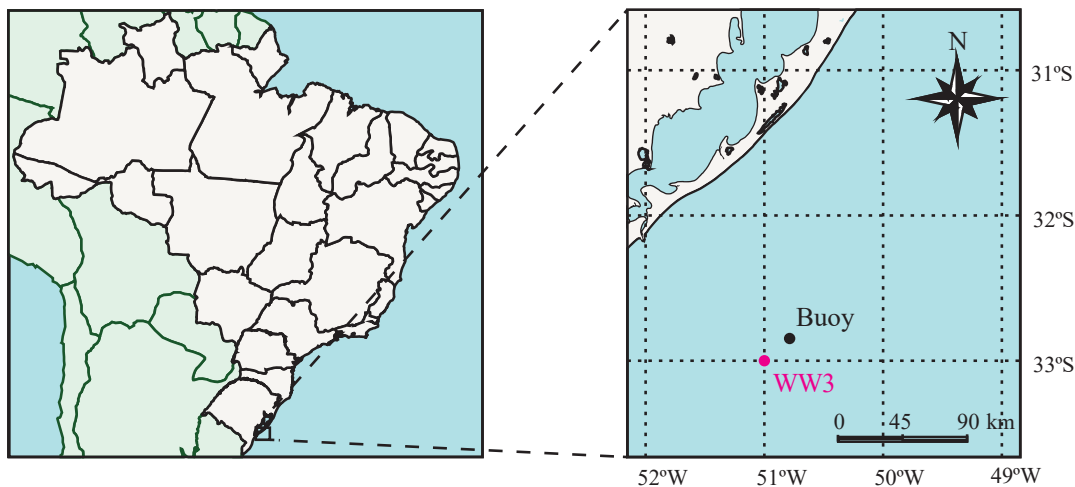

Figure 4.3 – Map of Brazil showing the location of the measurement.

The buoy in Fig. 4.3 is a directional Argus oceanographic buoy deployed at 100 m depth at 32°54'S and 50°48'W. The specified data collected between June/2002 and May/2003 validate the model in Section 4.1.

Extreme wind and wave conditions can be estimated from a comprehensive database measured by in situ buoys or by remote sensing. Nevertheless, in the South Atlantic Ocean, the absence of uninterrupted long-term wave data still is a significant obstacle for characterizing severe events in these regions and for extrapolating return levels with long return periods. With the significant advancements in model physics and numerical resources, wind and wave hindcasts generated by numerical wave models offer an alternative source for the data and have been widely adopted for wave analysis and ocean engineering when validated with measurement data [141-143].

One of the most referred and successful projects is the Wavewatch III (WW3) third-generation implemented by NOAA (National Oceanic and Atmospheric Administration) [141], which uses a WAVEWATCH III framework driven by NCEP Climate Forecast System Reanalysis and Reforecast wind data and operationally releases multi-scale forecast and hindcast waves in a global model domain. The NCEP Climate Forecast System Reanalysis (CFSR) entails a coupled reanalysis of the atmospheric, oceanic, sea-ice, and land data from 1979 to 2010 and a forecast run with this reanalysis afterward. This wind and wave hindcast is one of the most prevalent methods to forecast spatial wind and wave characteristics by wind using wave action conservation theory to characterize the complex temporal-spatial relationship between wind and wave [142].

Wind-wave evolution in third-generation models is described by the wave action balance:

$$\frac{\partial N}{\partial t} + \nabla c_g \, N = \frac{S_{tot}}{\sigma} \tag{4.19}$$

which considers all energy fluxes and source terms, represented by physical processes that contribute to wind-wave growth. The term c_g is the group velocity, $N(k,\theta) = F(k,\theta)/\sigma$ is the action density spectrum, F is the variance density spectrum, and σ is the relative radian frequency. The function F depends on the wavenumber k, direction θ, time t and space x,y. The first term of the left-hand side of Eq. 4.19 represents the rate of net change of wave action. The total source term S_{tot} on the right-hand side of the action balance (Eq.4.19) is all considerable energy fluxes contributing to wind-wave evolution. For deep water it is generally accepted that the total source terms is based on three main physical processes [144]:

$$S_{tot} = S_{in} + S_{ds} + S_{nl},\tag{4.20}$$

pointed out that these individual terms have to be further partitioned, where S_{in} is the atmospheric input, S_{ds} is the wave dissipation, S_{nl} is the nonlinear wave-wave interactions.

The outputs of the Wavewatch III hindcast are the u-component of input wind, the v-component of input wind, the significant height of combined wind waves and swell (H_s), the primary wave mean period (T_p), and the primary wave direction every 3 hours from 1979 to 2009 (31-year). Figure 4.4 shows the directional wave histograms of the means of H_s, T_p, and the directional wind histograms of velocity V_w obtained by Wavewatch III data. The predominance of ENE and S incident wave directions can be observed. The wind has high intensity and variability. The prevailing wind direction is NE, and the highest intensity is in the SW direction.
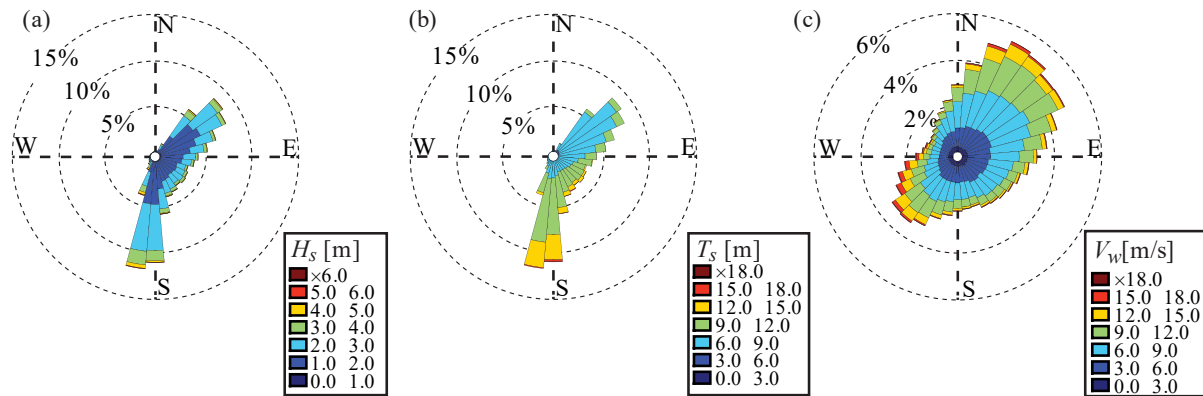
Figure 4.4 – Wave and wind directional histograms: (a) H_s, (b)T_p, and (c) V_w.

4.3.2 OWT case study structural modeling

This section presents the Finite Element Analysis (FEA) development of the jacket offshore wind turbine composed of the Offshore Code Comparison Collaboration Continuation (OC4) reference jacket structure and the National Renewable Energy Laboratory (NREL) 5 MW baseline turbine [145, 146]. The offshore wind turbine device comprises a tower, transition piece, support structure, and piles. The first step in FEA modeling is to define all geometrical parameters of the model, such as diameters, structural member thicknesses, and other geometric data, as described in Fig. 4.5.

(a) Offshore wind turbine structure

(b) Rotor Nacelle Assembly

(c) Tower cross sectional properties

global height	z [m]	outer diameter [m]	thickness [mm]	point mass [t]
Tw_0	20.15	5.600	32	1.9
Tw_1	21.15	5.577	32	No
Tw_2	32.15	5.318	30	No
Tw_3	42.15	5.082	28	No
Tw_4	54.15	4.800	24	1.4
Tw_5	64.15	4.565	22	No
Tw_6	74.15	4.329	20	No
Tw_7	83.15	4.118	30	No
Tw_8	88.15	4.000	30	1.0

(d) Properties of jacket members

Leg color	outer diameter [m]	thickness [mm]
Blue	1.200	35
Red	1.200	50
Orange	0.800	20

(e) Jacket positions

global height	z [m]	global height	z [m]
L_0	-44.001	L_5	-8.992
L_1	-43.127	L_6	-1.958
L_2	-33.373	L_7	+4.378
L_3	-24.614	L_8	+10.262
L_4	-16.371	L_9	+15.615

Figure 4.5 – The OC4 reference OWT.

As descript by Jonkman *et al*. [146] and Vorpahl *et al*. [145], the OC4 is designed for a reference site with a water depth of 50 m. The conical tower is 68m high and connected to the Transition Piece by bolts and flanges. The support structure is a jacket structure with mud braces and four central concrete piles with a penetration depth of 45m being grouted to mud. The transition piece between the jacket and tower is a block of concrete penetrated by the upper parts of four jacket legs. As the height increases, the tower's diameter, and wall thickness decrease. The turbine characteristics are shown in Table 4.3.

Table 4.3 – Properties for the NREL 5-MW baseline wind turbine.

Type	Parameter
Rating	5 MW
Rotor orientation, configuration	Upwind, 3 blades
Rotor, hub diameter	126m, 3m
Hub height	90.55m
Cut-in, cut-out wind speed, rated	3 m/s, 25 m/s, 11.4 m/s
Cut-in, rated rotor speed	6.9 rpm, 12.1 rpm
Rated tip speed	25 m/s
Rotor mass	110,000 kg
Nacelle mass	240,000 kg

The primary parts of the jacket support structure are mainly made up of steel materials. The material of the plate and stiffeners is AH32 high-strength steel. This material has Young's modulus E=206000 MPa, density ρ_s=8500 kg/m^3, yield stress σ_y=315.0 MPa, and Poisson's ratio υ=0.3. The transition piece and the pilling are a steel-concrete configuration. This material has Young's modulus E=70000 MPa, density ρ_c =2400 kg/m^3, Poisson's ratio υ=0.18, compressive strength σ_c=200.0 MPa and tensile strength σ_t=10.0 MPa.

4.3.3 OWT Finite Element Model

Two different FE models are considered to reproduce the fidelities functions concepts. The whole offshore wind turbine structure is modeled by ABAQUS software as shown in Fig. 4.6. For the first model (FEM1) depicted in Fig. 4.6(a), the jacket structure is modeled by B31 beam element, and tower, transition peace and piles are modeled by C3D87 solid element. For the second model (FEM2) depicted in Fig. 4.6(b), the jacket and tower structures are modeled by the B31 beam element and C3D87 solid element model transition peace and piles.

The rotor-nacelle assembly, the nacelle, and the tower equipment's mass are introduced in the FE model as concentrated point masses. The mesh size is crucial to define the MF model and is discussed in the next chapter.

Constraints are imposed to reproduce the iterations in the different structure's parts. A tie constraint between the transition piece and the tower (Fig. 4.6(c)) ties the surfaces together, so there is no relative motion between them. In addition, an embedded region constraint is also applied between the transition piece and the jacket structure (Fig. 4.6(d)) that allows the Transition Peace to act as a host region of the embedded jacket.
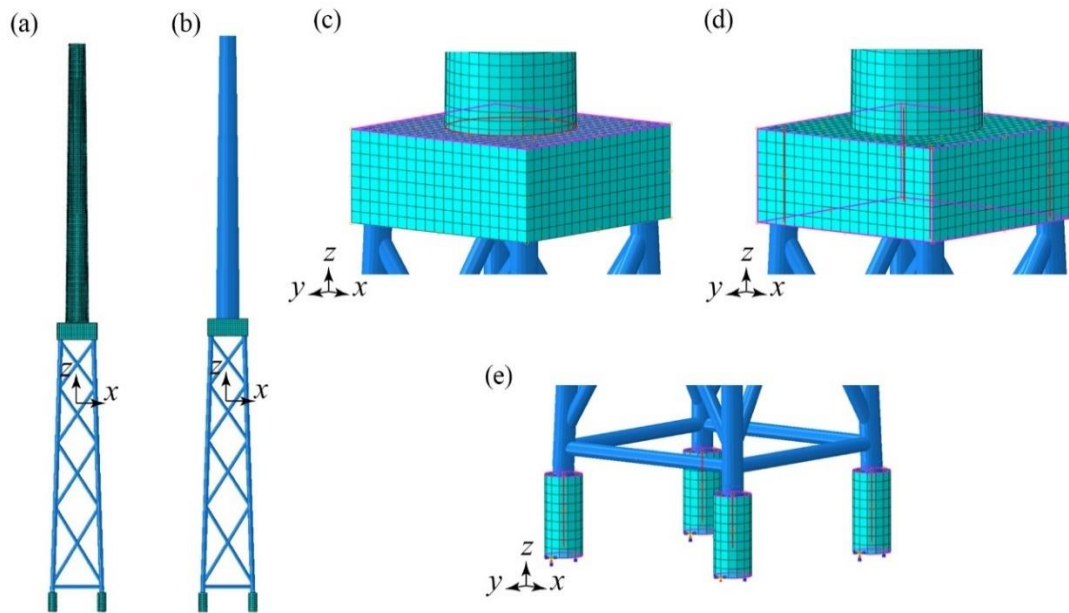
Figure 4.6 – FE model construction of OC4 Jacket Substructure.

The jacket is also embedded in the piles (Fig. 4.6(d)). Previous studies (e.g. [127, 128]) incorporated the soil-structure interaction, in order to map its response under varying input conditions. However, for simplicity, OWT bottom-fixed support structures are used in OWT assessment (e.g. [131, 147-149]), where the jacket is fixed to the seabed by grouting connected to piles. In the present study, the calculations are assuming that the piles are truncated at the seabed, where they are assumed as perfectly clamped at the bottom. The grouted connection range between the piles and the jacket is 4.5m above the seabed.

The non-linear structural analysis is divided into two steps. In the first step, a gravity load is applied to define a uniform acceleration in the z direction in the whole model to evaluate the inertia and gravitational effects in the structure. The structural behaviour obtained by the first step is used by initial geometry for the second step analysis. The aerodynamics and environmental loads are applied to get the maximum stress and deflection in the OWT.

4.3.4 Loads for offshore structures

According to standard codes DNV–OS–J101 [150] and IEC 61,400–3 [151], the relevant loads imposed on OWT support structures can roughly be classified into (*i*) aerodynamic loads transferred from the rotor; (*ii*) wind loads on the tower; (*iii*) inertia loads; and (*iv*) wave loads. The scheme of the resulting loads acting in the OWT structure is shown in Fig. 4.7.
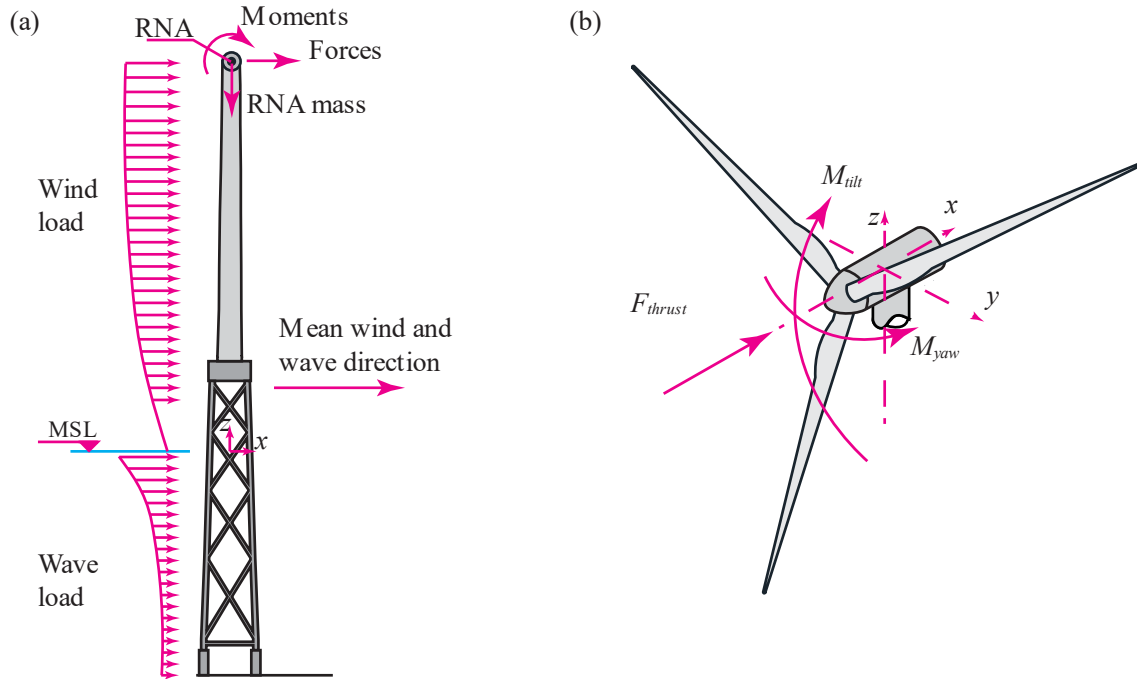
Figure 4.7 – Environmental actions over the (a) Jacket and tower and (b) the RNA.

Due to the mass of the support structure, equipment, and the Rotor nacelle assembly (RNA) mass at the top of the tower, inertia and gravitational loads are static and dynamic loads that can significantly affect the modal frequencies of the OWT support structure and contribute to buckling. Therefore they should be considered in the structural analysis of support structures [151].

Wind loads contain two parts: one is acted on the tower, and the other is acted on the blades. Aerodynamic loads are static and dynamic loads that occur due to the airflow and its interaction with the stationary and moving parts of the wind turbine. The magnitude of the load is dependent upon the average wind speed and turbulence across the rotor plane, the rotational speed of the rotor, the density of the air, and the aerodynamic shapes of the wind turbine components and their interactive effects, including aeroelastic effects [128, 151, 152].

As a result of Aerodynamic analysis, final loads and moments are considered, as depicted in Fig. 4.7(b) [128]. The thrust (F_{thrust}) load was imposed on the support structure and applied at the assumed RNA Centre of Mass (CM) in the same direction as the diagonal of the base of the jacket. The notable aerodynamic moments (M_{tilt} and M_{yaw}) represent the hub mechanical loading in two orthogonal quantities that can be calculated from parameters such as the yaw error and tilt angle. Since the RNA coupled with the blades for the turbine is usually large, considerable eccentricity is created due to the lateral span of this assembly imposing these moments.

84

Environmental loads are those caused by environmental phenomena, i.e., the set produced by wave force and wind force. Wave, wind loads, and rotor thrust vectors are applied in the same direction as the jacket base diagonal to produce maximum compression and tension effects on the opposite legs [128, 148].

Wind load pressure acting on the tower structure is among the essential load sources to be considered in the design of a wind turbine support structure [127, 152]. Wind loads on a structure result from drag between the structure above sea level and the wind in a given field. The wind load is dependent on the mean wind velocity $V_w(z)$ mathematically expressed assuming wind speed variation with height above still water level as a power law profile [151]:

$$V_w(z) = V_{w,r} \left(\frac{z}{z_r} \right)^{\alpha_r} \tag{4.21}$$

where $V_{w,r}$ is the wind speed at the reference height z_r, and α_r is the roughness coefficient. For offshore locations, it is recommended to apply an exponent $\alpha_r=0.14$ [150]. Wind loads along the tower are then determined from the following:

$$F_{tower}(z) = \frac{1}{2} \rho_a C_w D(z) V_w^2(z) \tag{4.22}$$

where ρ_a is the air density, C_w is the drag coefficient of the tower, taken as 1.0 from [127, 128, 152], D is the external diameter of the tower at height z.

A recognized wave theory for representing wave kinematics shall be applied to calculate wave loads. In the present analysis, the kinematics of regular waves is represented by Linear Wave Theory (Airy theory) for small-amplitude deep water waves [150, 151]; by this theory, the wave profile is represented by a sine function. Morison's Equation can be applied to calculate the wave loads for slender structures, such as jacket structure components and monopile structures. Morison's Equation assumes that the total wave force exerted on a structure can be calculated by the linear superimposition of the drag and inertia forces, assuming that:

$$F_{wave}(z) = \frac{1}{4} \rho_w \pi D^2 C_m \dot{u}(z,t) + \frac{1}{2} \rho_w D C_d u(z,t) . |u(z,t)| \tag{4.23}$$

where ρ_w is the density of seawater, C_m and C_d are the coefficient of inertia and drag of the piles respectively, taken as 1.6 and 1.0 respectively [150], while $u(z,t)$ and $\dot{u}(z,t)$ are respectively horizontal velocity and acceleration of water particles, which can be obtained from linear/Airy wave theory.

4.3.5 Ultimate Load Case

The Design load cases (DLC) should cover a set of design situations considering the most severe conditions that an OWT is likely to be subjected to, combining extreme or standard external conditions with operational states of the wind turbine or other operating modes. Design standard IEC61400-3 [151] defines 32 load cases for the structural design of OWTs, covering all the operating conditions of an OWT categorized into two groups, i.e., ultimate and fatigue. For simplicity, the ultimate load case in extreme events is considered in this study.

One of the standard IEC test cases is the IEC 61400–3 DLC 6.1 [151], which considers the turbine in an extreme wind scenario when the turbine is parked. In this study, it is assumed that the critical load case for Ultimate Limit State (ULS) and Serviceability limit state (SLS) is defined under the 50-year EWM (extreme wind model) with 50-year RWH (reduced wave height). The OWT modeled parked was used as a reference in ULS and SLS in previous reliability assessments, e.g., [127, 128, 131, 132, 153-155].

The desired turbulent wind is initially specified to model the NREL 5-MW baseline wind turbine as an idling rotor. Disabling aerodynamic induction, dynamic stall, and pitch control are necessary, which is not valid at large angles of attack. The blade's pitch angle is set to a feathered position (90°), and the rotor speed to 0 rpm. Additionally, the condition of the generator never starts is to enable.

The EWM is considered turbulent, and the 10 min average wind speeds as functions of z with recurrence periods of 50 years, respectively, shall be given by Eq. 4.21 with longitudinal turbulence standard deviation $\sigma_1 = 0.11 V_{hub}$ [151]. The RWH defines the extreme event with a specified return period. Load safety factors for gravitational and other loads (such as environmental) are 1.1 and 1.35, respectively.

Sophisticated numerical models in the time domain have been developed, some under code collaborations, to predict the response of OWTs. In this stud, the OpenFAST code [156] developed by the National Renewable Energy Laboratory (NREL) is used with the aero-hydro-servo-elastic simulation code to obtain the Aerodynamic loads in the Hub. A linear finite element model with Timoshenko beam elements represents the substructure of the OpenFast. The software TurbSim [156], based on the Kaimal Wind Turbulence Model, produces the wind turbulence map and the spectra of three orthogonal wind speed components [157].

The ULS accounts for the ability of the support structure to resist plastic deformation. The equivalent stress is generally determined using the von Mises stress theory for an OWT jacket support structure. The performance function of the ultimate limit state based on the von Mises criterion is given by:

$$G_u(x) = \chi_s \sigma_{allow} - \chi_{site} \chi_{aero} \chi_{dyn} \chi_{mat} \chi_{wind} \chi_{stm} \sigma_{max} \tag{4.24}$$

where σ_{allow} is the allowable stress and σ_{max} is the maximum von-Mises stress. The allowable stress is the ratio $\sigma_{allow} = \sigma_y / \gamma_s$, considering σ_y the yield strength of the material, and γ_m the safety factor for the material. The factor χ_s is the stress capacity uncertainty factor, χ_{site} is the uncertainty related to the site and atmospheric conditions, χ_{aero} is the uncertainty in the aerodynamic properties, χ_{dyn} is the model uncertainty associated with the structural dynamics, χ_{mat} is the uncertainty due to variations in material and geometrical properties, χ_{wind} is the model uncertainty in the wind model, and χ_{sim} is the statistical uncertainty associated with the design process of sampling wind conditions with a limited number of simulations.

The deflection tolerance criteria influence the SLS of the OWT. To avoid the uncertainties introduced by large deformations, the maximum total deflection measured at the top of the tower is $\delta_{allow} = L / \gamma_d$, where L represents the length of the support structure, and γ_d is the safety factor for the structure. Considering δ_{max} the maximum deflection, the performance function of deflection limit state design can be expressed as:

$$G_d(x) = \chi_d \delta_{allow} - \chi_{site} \chi_{aero} \chi_{dyn} \chi_{mat} \chi_{wind} \chi_{stm} \delta_{max} \tag{4.25}$$

where the variable χ_d is the displacement calculation uncertainty.

4.3.5.1 Blocks Maxima Method to obtain 50-year extreme condition

In this study, the load cases used in the structural design of offshore wind turbines are the ultimate load under 50-year extreme conditions identified by the Block (or Annual) Maxima (BMM) and are associated with the use of the generalized extreme value (GEV) distribution. This association is widely used for extreme value theory studies [158-160].

The Block Maxima, Annual, or Gumbel method is described as follows. Assume that the given data y_i are maxima, $y_i = \max(x_{i,1}, \ldots, x_{i,m})$ where the $x_{i,j}$ may not be observable. If the $x_{i,j}$ can be observed, then taking maxima out of certain subperiods is another possibility for extracting extreme upper values from a data set. Then, the annual maximum can be regarded as the maximum of, e.g., monthly, or specific seasonal maxima.

The GEV distribution is a family of continuous probability distributions developed within extreme value theory. It combines the Gumbel, Fréchet, and Weibull extreme value

distributions. The PDF ($\phi(.)$) and CDF ($\Phi(.)$) for the generalized extreme value distribution with location parameter η, scale parameter ξ, and shape parameter k are:

$$\phi(x) = \frac{1}{\xi} \cdot t(x)^{k+1} \cdot \exp(t(x)),$$

$$\Phi(x) = \exp(-t(x)), \text{ where}$$
(4.26)

$$t(x) = \begin{cases} \left[1+\left(\dfrac{x-\eta}{\xi}\right)\right]^{-\frac{1}{k}} & \text{if } k \neq 0 \\[2em] \exp\left(-\dfrac{x-\eta}{\xi}\right) & f\, k = 0 \end{cases}$$
(4.27)

When $k<0$, the GEV is Type III (Weibull). When $k>0$, the GEV distribution is Type II (Fréchet). If the limit as k approaches 0, the GEV is the mirror image of Type I (Gumbel). The T–Year return level z_p is given by $(1-p)$–quantile of the $\Phi(.)$, with return period $p=1/T$ resulting in:

$$z_p = \begin{cases} \eta - \dfrac{\xi}{k}\left[1 - y_p^{-k}\right], & \text{if } k \neq 0 \\[1.5em] \eta - \xi \log(y_p), & \text{if } k = 0 \end{cases}$$
(4.28)

where $y_P = -log(1-p)$. More precisely, z_p is exceeded by the annual maximum in any particular year with probability p. The inference for the return levels is made by the Delta method that provides uncertainty bounds with an assumption that the distribution of return levels obtained from the MLE estimator follows a Gaussian distribution. The variance of the return level is given by [159]:

$$\text{Var}(z_P) \approx \nabla z_P^T V \nabla z_P$$
(4.29)

where V is the variance-covariance matrix of (σ, μ, ξ) and

$$\nabla z_P^T = \left[\frac{\partial z_P}{\partial \eta}, \frac{\partial z_P}{\partial \xi}, \frac{\partial z_P}{\partial k}\right] = \left[1, -k^{-1}\left(1 - y_P^{-\xi}\right), \xi k^{-2}\left(1 - y_P^{-k}\right) - \xi k^{-1} y_P \log y_P\right]$$
(4.30)

4.3.6 Stochastic models for basic random variables

The successful deployment of OWTs largely depends on accurately estimating the effects of stochastic loads acting on the asset and accurately predicting the components' integrity throughout their service life. The basic random variables considered in the OWR reliability problem are shown in Table 4.4, both with their corresponding probability density distribution (PDF) and statistical moments mean (μ) and standard deviation (*std*) of the associated distribution functions.

Table 4.4 – Stochastic models of the basic random variables.

Variable	Units	P.D.	μ	COV (%)	std
χ_s	-	Lognormal	1.0	5.0	0.05
χ_d	-	Lognormal	1.0	5.0	0.05
χ_{site}	-	Lognormal	1.0	10.0	0.10
χ_{aero}	-	Lognormal	1.0	10.0	0.10
χ_{dyn}	-	Lognormal	1.0	5.0	0.05
χ_{mat}	-	Lognormal	1.0	5.0	0.05
χ_{wind}	-	Lognormal	1.0	10.0	0.10
χ_{sim}	-	Lognormal	1.0	5.0	0.05
σ_y	MPa	Lognormal	348.00	6.0	20.88
E_p	MPa	Lognormal	206000	6.0	12360
$H_{s,50}$	m		See Results Section 4.2		
$T_{p,50}$	s		See Results Section 4.2		
$V_{w,50}$	m/s		See Results Section 4.2		
Ft_{thrust}	kN		See Results Section 4.3		
Mt_{ilt}	kNm		See Results Section 4.3		
My_{aw}	kNm		See Results Section 4.3		

The stochastic models for the uncertainties factor variables are based on IEC-CD-TS-61400-9 [161, 162]. The authors describe a lognormal distribution with μ=1.0 and COV=5% to represent the capacity calculation uncertainty factors χ_s and χ_d. Similarly, the uncertainty factors χ_{site}, χ_{aero}, *and* χ_{wind} follow a lognormal distribution with μ=1.0 and COV=5%. The uncertainty factors χ_{dyn}, χ_{mat}, and χ_{sim} follow a lognormal distribution with μ=1.0 and COV=10%.

The design value considered for the steel yield stress σ_y=315 MPa corresponds to the probability density function's 5% percentile characteristic value of the lognormal distribution with COV=6% and lower limit=0, resulting in a mean of μ=348 MPa [74]. Young's modulus corresponds to the mean characteristic value of the lognormal distribution with COV=6% and lower limit=0 [74]. The steel Poisson's ratio can be treated as a deterministic variable.

The expected extreme wind speed $V_{w,50}$, significant wave height $H_{s,50}$, and primary wave mean period $T_{p,50}$ with a return period of 50 years shall be calculated considering the return level of the GEV distributions applied in the extrapolation fitting functions. The fitted GEV for the peak responses of V_w, H_s, and T_p are estimated by Block Maxima Method. The $V_{w,50}$, $H_{s,50}$, and $T_{p,50}$ parameters follow a normal distribution with means equal to the expected values and standard deviation calculated by the Delta Method presented in Results Section 4.2.

The IEC standard [151] recommends the 10-minute load samples to estimate the probability distribution of the 50-year load of wind turbines through the statistic extrapolation method. The detailed statistical procedures are not specified in the IEC standard, so the Block Maxima Method is used for the peak extraction. The F_{tilt}, M_{tilt}, and M_{yaw} loads estimation follow GEV distributions presented in Results Section 4.3.

4.4   ANALYSIS OF RESULTS

This section firstly shows the results of validating the ocean wind and wave model-WW3, clarifying the model tendency to the study area, and quantifying the model performance in an extreme loads scenario. After that, it presents the results of the extreme value analysis based on the BMM and GEV methods. Then, a validation of the FE OWT model is realized. In sequence, the trained BFMT-DNN model results are presented. Later, an OWT reliability analysis considering the BFMT-DNN model is presented.

4.4.1   Hindcast validation

The wave and wind hindcasts based on the NCEP/NCAR Reanalysis and WAVEWATCH III are validated considering a one-year interval of data collected by the buoy in situ. Figure 4.8 shows time series comparisons of blended wind speeds and H_s, T_p, and V_w measurements. Table 4.5 shows the basic statistic parameters, mean μ and standard deviation sd, compared between the two scenarios for the same period.
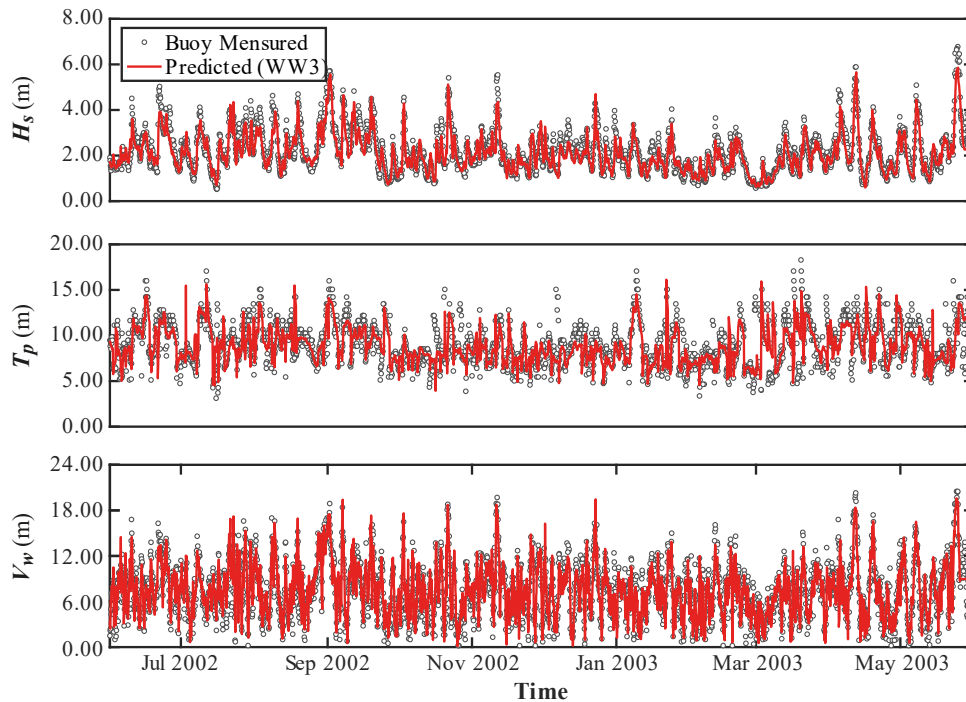

Figure 4.8 – Numerically and experimentally obtained (a) H_s, (b) T_p, and (c) V_w time series.

90

Table 4.5 – Basic statistical parameters for the Buoy and WW3 measurements.

Variable	Buoy Measured		Predicted (WW3)	
	μ	sd	μ	sd
H_s (m)	2.2081	0.9662	2.1074	0.8301
T_p (s)	9.1142	2.3169	8.9510	1.9758
V_w (m/s)	7.5388	3.7210	7.5501	3.3974

The results presented in Fig. 4.8 and Table 4.5 show good agreement between numerical and experimental results for the H_s, T_p, and V_w parameters. Figure 4.9 shows density scatter plots between the measured buoy and numeric WW3 outputs for H_s, T_p, and V_w. The black lines are the expected equality between the measured buoy and numerical WW3 output. Scatter density diagrams generally show good agreement between numerical and experimental results, most notably for H_s and V_w. The resulting scattering shape for T_p emphasizes bimodal structures in the ocean wave spectra, in which two dominant frequencies and two dominant directions are found, also discussed in [140]. In these cases, differences can be observed between measurement and numerical results related to the position, sea, or swell parts, where the highest energy is concentrated.
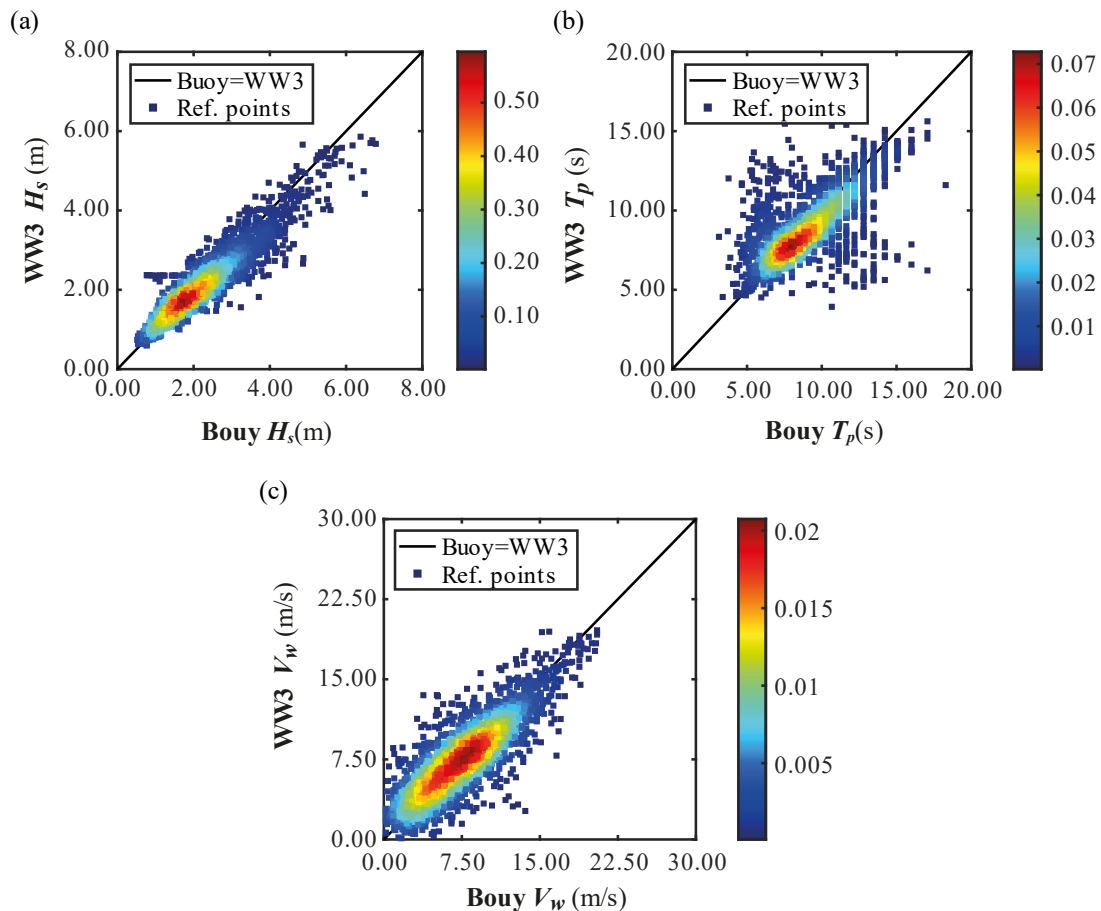

Figure 4.9 – Scatter plot density diagrams of (a) H_s, (b) T_p, and (c) V_w time series.

4.4.2 Extreme events on the southern coast of Brazil

The Block Maxima method with a reference subperiod of six months is used to collect the extreme occurrence of H_s, and V_w parameters during the 31 years hindcast. The period T_p, the return period corresponding to H_s. After that, each parameter is fitted by the GEV function, with the best probability distribution results presented in Table 4.6. The return levels associated with the 50-year return period are estimated from the best GEV-fitted distribution for each parameter. Figure 4.10 shows the return period estimation.

Table 4.6 – Best fitted GEV distributions for the wave and wind parameters.

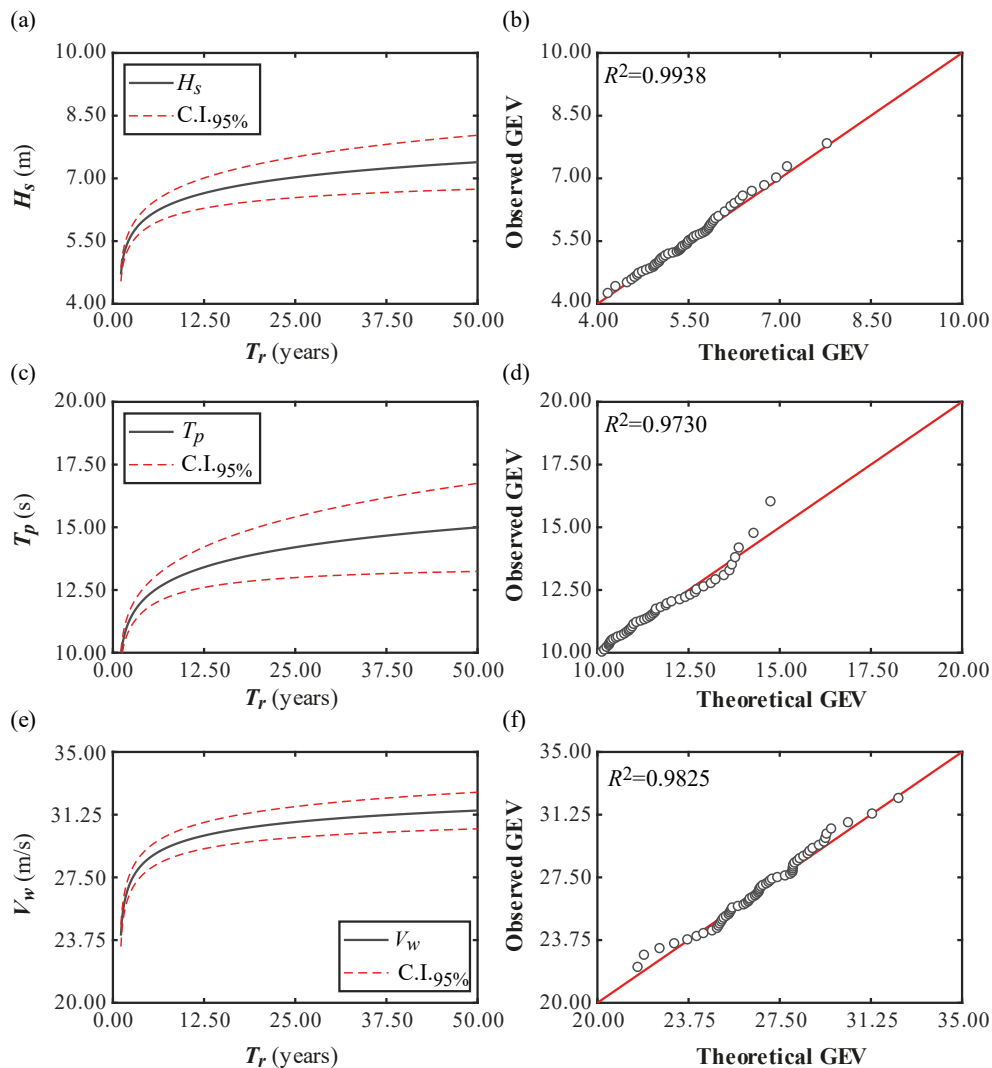| Variable	Unit	GEV type	k	ξ	η
H_s	m	Weibull	-0.043	0.600	5.235
T_p	s	Fréchet	0.015	1.065	10.710
V_w	m/s	Weibull	-0.211	1.519	19.120

Figure 4.10 – Return level estimates based on fitting the GEV distribution to BMM. On the left side, the (a) H_s, (c) T_p, and (e)V_w return level. On the right side, the Q-Q plot of (b) H_s, (d) T_p, and (f)V_w.

The black line is the return level over the years, and the dashed red line is the confidence interval obtained by the Delta variance method over the years. The small confidence intervals for extreme return levels show enough information to predict return periods over 50 years. Considering the results presented in Fig. 4.10, it is possible to determine the uncertainty under the 50-year parameters. Assuming that the MLE estimator follows a Gaussian distribution by the Delta method, the unexpected results in the distributions are presented in Table 4.7.

Table 4.7 – Uncertainty of 50-year wind and wave parameters.

Variable	Unit	P.D.	μ	sd	COV
$H_{s,50}$	m	Normal	7.400	0.108	0.014
$T_{p,50}$	s	Normal	15.000	0.790	0.052
$V_{w,50}$	m/s	Normal	31.500	0.168	5.336e-3

4.4.3 OWT loads scenario in the extreme conditions

To realize the OWT aerodynamic simulation, the Extreme Wind Model conditions are analyzed for 10 min considering as reference the mean value of $H_{s,50}$=32.5m/s with longitudinal turbulence standard deviation σ_1=3.25%. The time step of wind speed history is 0.05 s. The time history of three orthogonal wind speed components is input in a 17×17 mesh within an area of 200×200 m, as shown in Fig. 4.2, which is sufficient to cover the wind turbine rotor and is used to calculate the aerodynamic load of the wind turbine. The original 10-min simulation is run for 650 s, and the data of the first 50 s is removed to eliminate startup transients. The stacked plot in Fig. 4.11 shows the EWM state and load variation over the 10-min time series.

The 10-min peak loads are extracted from the load time history, as shown in Fig. 4.11. The BMM is used to extract the peak loads in a subperiod of 5 s. The peak values are multiplied by the load safety factors of value 1.35. The GEV fitting functions are used to model the extreme load, and the performance of these fitting functions is evaluated through peak loads. The probability density functions of the extreme loads are shown in Fig. 4.12, and their parameters are presented in Table 4.8.
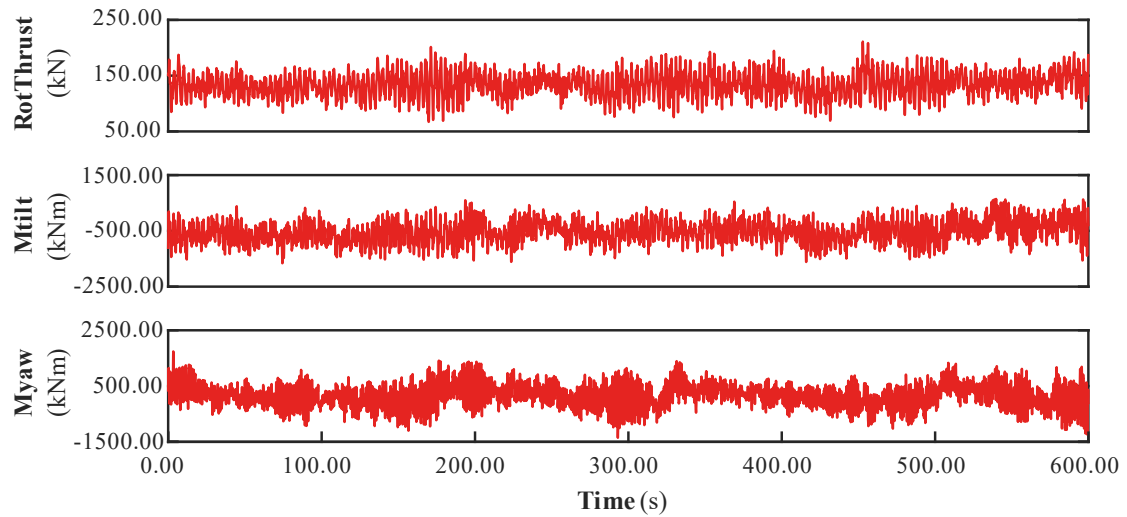
Figure 4.11 – The 10-min time series of wind speed, rotor thrust, tilt and yaw moments, and power generated.
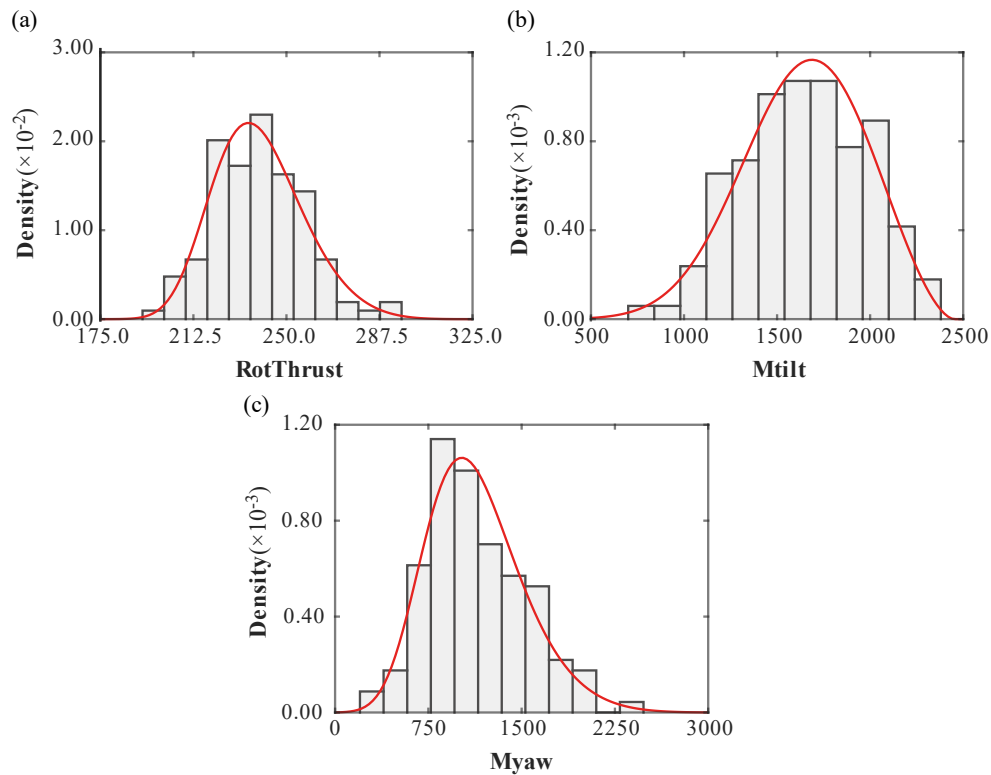
Figure 4.12 – RNA extreme loads probability distribution.

Table 4.8 – Best fitted GEV distributions for the extreme loads.

Load	Unit	Direction	GEV type	k	ξ	η
Thrust Force	kN	$(+)z$	Weibull	-0.168	16.9	231.6
Tilt Moment	kNm	$(-)x$	Weibull	-0.375	342.1	1539.4
Yaw Moment	kNm	$(+)y$	Weibull	-0.125	349.1	972.2

94

4.4.4 Validation of the OWT FE models

The validation of the proposed OWT FE models FEM1 and FEM2 offered in Section 3.3 is realized considering the benchmarking study of the NREL 5 MW OWT OC4 jacket support structures that consider the deflection and the modal frequency analysis by references Jonkman *et al*. [146], and Damiani *et al*. [148]. The proposed benchmark used to validate OWT models has been reported in previous research (e.g., [127, 163]).

The deflection scenario assesses the total deflection of the OWT support structure in static analysis. In the reference, the case study applies a thrust of 2 MN in the RNA. The RNA and tower base displacements are measured and presented in Table 4.9. The modal analysis accesses the natural frequencies of the model. The modal frequencies calculated from the present FEA model are compared with the reference, and the comparison results are presented in Table 4.10.

Table 4.9 – Deflection validation on OC4 jacket structure in m.

Site	Reference	FEM1	%Diff.	FEM2	%Diff.
tower top	1.209	1.187	1.820	1.164	3.722
tower base	0.137	0.139	1.460	0.140	2.189

Table 4.10 – Modal analysis validation on OC4 jacket structure in Hz.

Site	Reference	FEM1	%Diff.	FEM2	%Diff.
1st Fore-aft	0.319	0.325	1.881	0.317	0.627
1st Side-to-side	0.319	0.325	1.881	0.317	0.627
2nd Fore-aft	1.194	1.138	4.690	1.146	4.020
2nd Side-to-side	1.194	1.138	4.690	1.146	4.020

Reasonable agreement is achieved in all validation sets, with a maximum relative difference of 4.69% observed for the 2nd fore-aft and side-to-side frequency modes in the FEM1. The maximum relative difference for the FEM2 is 4.02% for the 2nd fore-aft and side-to-side frequency modes. These results confirm the validity of the developed FE models.

4.4.5 The bi-fidelity model approach

This subsection focuses on the MF functions in which fidelities are combined inside the surrogate model. One of the fundamental steps of the design under the MF context is the ability to efficiently define the HF and LF models and catch the correlation between the different fidelities functions that may have some impact on the performance of MF surrogate models. The OWT FE model introduced in Subsection 3.3 is used to construct HF and LF functions. Figure 4.13 show the mesh convergence for each output.
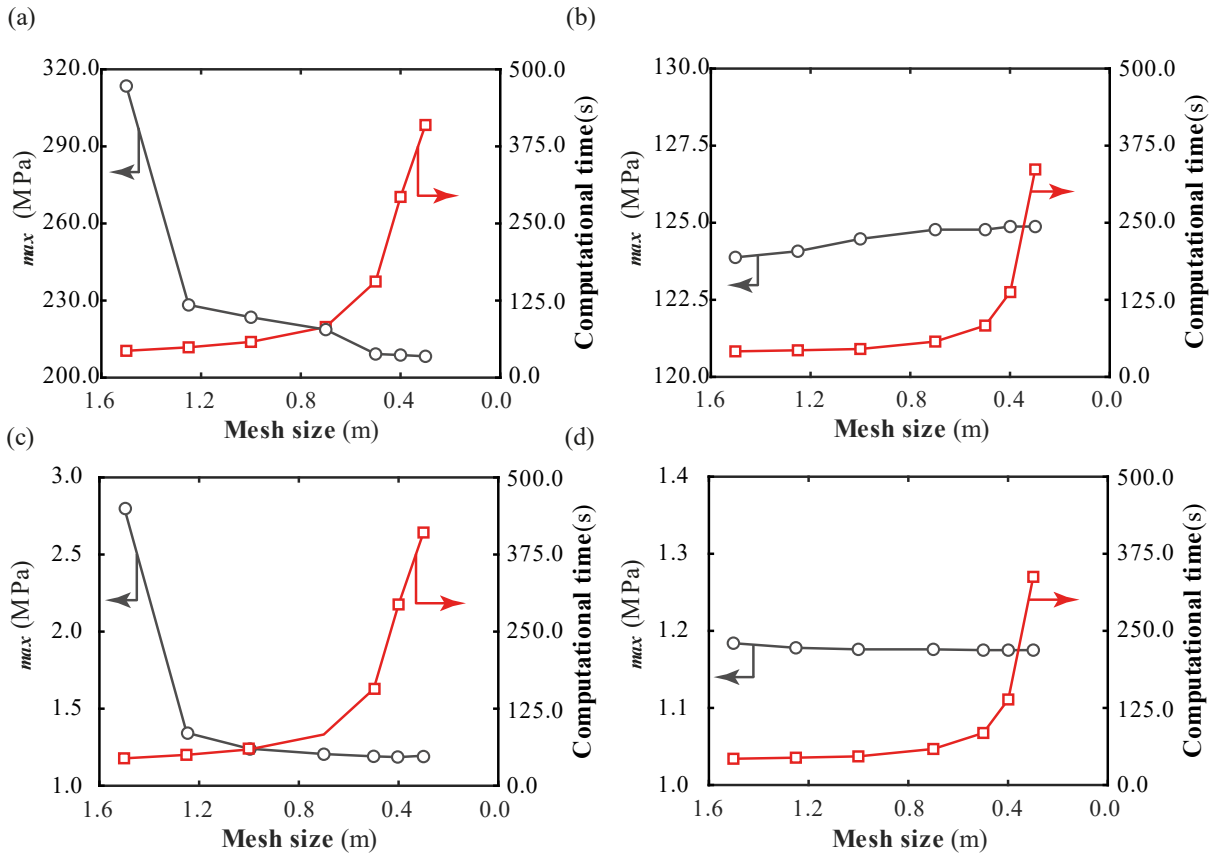
Figure 4.13 – Mesh convergence analysis and computational cost. On the left side, the (a) σ_{max} and (c) δ_{max} convergence for the FEM1. On the right side, the (b) σ_{max} and (d) δ_{max} convergence for the FEM2.

Before HF and LF functions definition, the influence of mesh size (*ms*) must be studied in the FEM1 (Fig. 4.13 (a)) and FEM2 (Fig. 4.13(b)), refining it enough to provide accurate and valid results. The OWT reference used in the FE validation (Subsection 4.4) is also used for the mesh convergence study. The *ms* influence in the OWT σ_{max} is shown in Fig. 4.13(a) and 4.13(b) for the FEM1 and FEM2, respectively. For the OWT δ_{max}, the results are depicted in Fig. 4.13(c) and 4.13(d) for the FEM1 and FEM2, respectively. All mesh convergence analysis is associated with computational time.

Considering the results presented in Fig. 4.13 is noticed that FEM1 and FEM2 have similar computational time for the most significant *ms* values because, in these cases, the large mesh size and, consequently, the number of finite elements is not predominant in the computational cost. However, for refined mesh values (*ms*<0.8 m), the time for both models increases considerably, but FEM1 is slower because the number of elements is considerably increased, and consequently, the number of constraints and boundary conditions in the FE model.

The HF function is defined considering the FEM1 with an accurate mesh. Two different LF functions are utilized to identify the advantages and disadvantages of using a quicker but less accurate mesh in the FEM1 and a FE simplification in the FEM2 analysis. To evaluate the relationship of the mesh fidelities model, a test dataset D_{test} is used to determine the DIC correlation parameter. Figure 4.14 depicts the correlation between the HF functions and each LF function for the σ_{max} and δ_{max} outputs considering D_{test}. For the fidelity's correlation, a dataset with n_{test}=200 is considered. Table 4.11 summarize the MF functions and their FE characteristics, computational costs, and DIC parameter between the HF and LF functions.
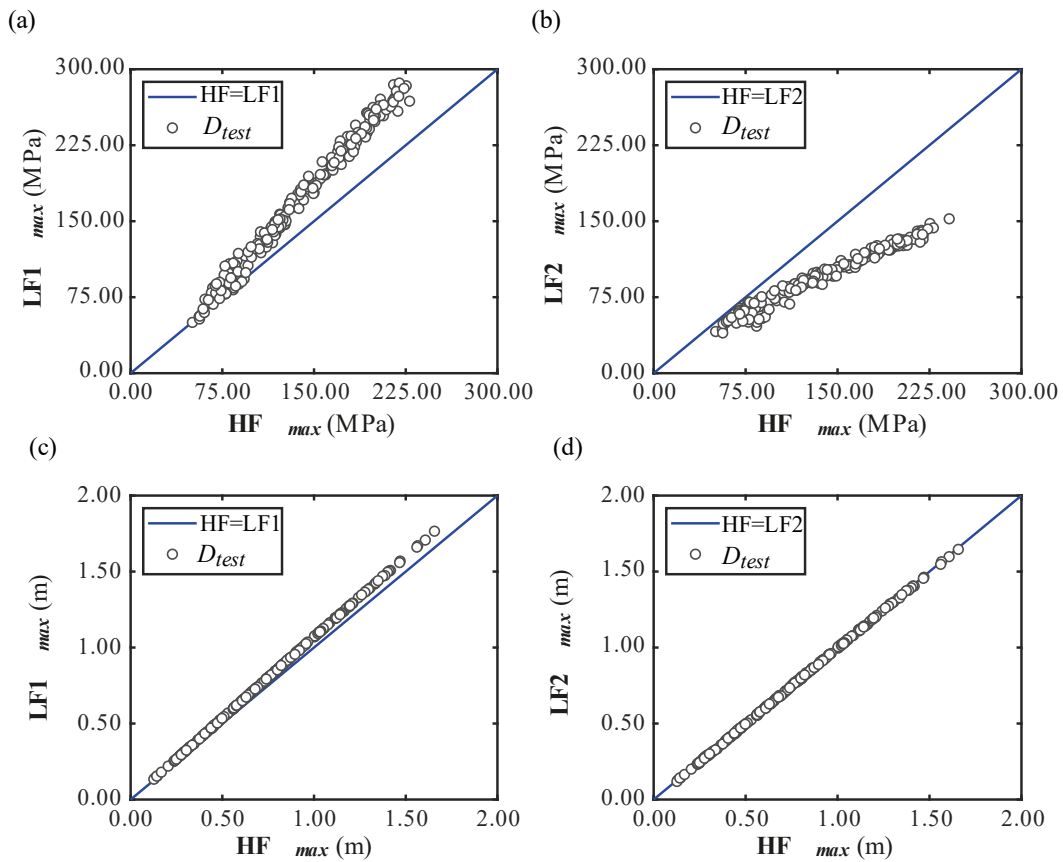

Figure 4.14 – Correlation between the HF and LF models. On the left side, the (a) σ_{max} and (c) δ_{max} correlation for the FEM1. On the right side, , the (a) σ_{max} and (c) δ_{max} correlation for the FEM2.

Table 4.11 – Fidelities correlation and cost functions.

Fidelity Function	Mesh size [m]	Comp. Cost [s]	Ratio to HF	DIC	
				σ_{max}	δ_{max}
LF1	1.00	59.00	0.198	0.993	0.999
LF2	0.70	58.00	0.198	0.984	0.999
HF	0.40	294.00	-	-	-

The computational cost presented in Table 4.11 is the average time consumed by the D_{test} to run the non-linear FE analysis. The two LF functions share a similar computational cost for a fair comparison. Among these two LF functions, both presented a linear correlation about δ_{max}, although the LF2 function presented similar δ_{max} values compared to the HF function values. About σ_{max}, the compared two different LF functions present different correlation behaviors drastically, and the LF1 presented a better DIC correlation value, although the DIC is close between the LF functions.

4.4.6 Accuracy and efficiency of the BFMT-DNN model

For future analyses, two Bi-fidelity models, BF1 and BF2, are referenced, built through each association of the LF1 and LF2 models with the HF model. Before choosing the BFMT-DNN model with hyperparameters optimized, it is essential to define the number of support points n_{LF} and n_{HF} in the HF and LF training datasets $D_{tr,HF}$, and $D_{tr,LF}$. For the BFMT-DNN hyperparameter optimization process, a validation dataset D_{val} with n_{val}=200 minimizes the loss function over the iterations. The same D_{test} dataset used in the DIC is also used to test the optimized BFMT-DNN model.

The n_{LF} value is determined by multiplying n_{HF} by a sample size ratio between HF and LF models (λ_{BF}) to obtain the improvement level with potential LF sample points. The potential and impact of the LF sample points are evaluated with the HF samples fixed to n_{HF} =500 support points, while the infill strategy of LF samples is increased as the ratio λ_{BF} ranges from 2 to 6. The total cost is given by the number of calls to the performance function N_{call} that, based on the cost functions, equals the equivalent number of HF samples given in Table 4.11. Figure 4.15(a) shows the influence of the sample size ratio on N_{call} to the performance function.
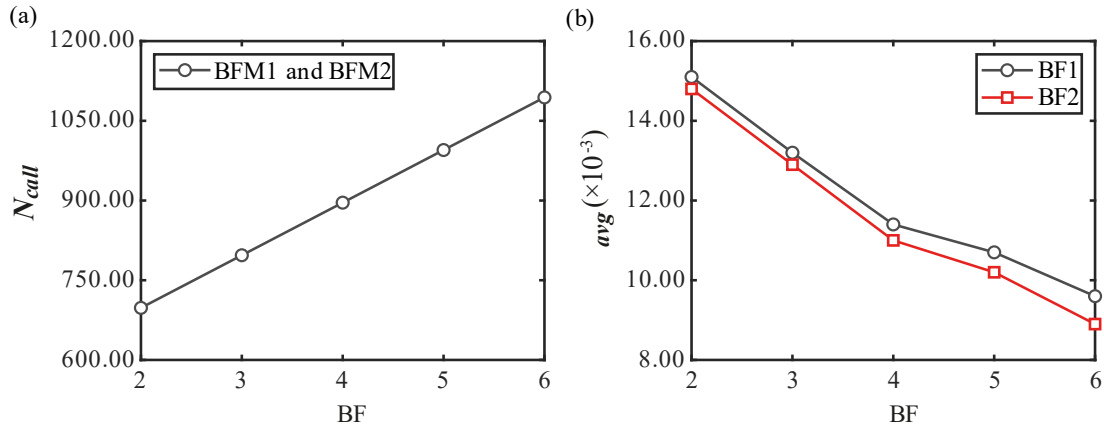

Figure 4.15 – The tendency of (a) the global N_{call}, and (b) MAE accuracy for different sample ratios.

By plotting the evolution of the N_{call} of the performance function in terms of the λ_{BF}, it can be seen that the N_{call}, and consequently the computational cost, increase at a linear rate of associated λ_{BF}. For each increment of λ_{BF}, the value of N_{call} increases by 99.0, and comparing the extreme cases (λ_{BF}=2 and 6), N_{call} is increased by 40 and 118% approximately, when compared with the N_{call} of only the HF samples.

The accuracy metrics of the BFMT-DNN cases are obtained with the sum of the metrics of each stage (NN_{LF} and NN_{BF}) and for each output (δ_{max} and σ_{max}) of the BFMT-DNN model as described in the loss definition (Eq. 4.9). The RMSE presented in Fig. 4.15(b) is used as a fit measure of the BF-DNN surrogate model under different sample ratios λ_{BF}. For the n_{HF}=500 , it is clear that the accuracy line presents two slopes for both BF1 and BF2 models, changing when the λ_{BF}=4. For the cases with $\lambda_{BF}\leq 4$, the precision increases on a larger scale compared to the cases with λ_{BF}>4, where the accuracy improves with the increase of the λ_{BF} ratio , but more conservatively when compared with the first slope.

For all cases, the BF2 presented a better accuracy when compared with the BF1. Also, comparing the two different slopes in Fig. 4.15(b), for the cases with $\lambda_{BF}\leq 4$, the accuracy difference between the BF1 and BF2 models is slight. For the cases with λ_{BF}>4, the difference between the accuracy values of the two models remains small, but the BF2 model is highlighted compared to the BF1 model. These results show the importance of the assessment of the relation λ_{BF}. Considering the presented accuracy results, the case with n_{HF}=500 and λ_{BF}=6 provides accurate metrics estimates and can be viewed as an appropriate BFMT-DNN surrogate model for BF1 and BF2 models.

Figure 4.16 depicts a scatter plot of the OWT problem outputs (σ_{max} and δ_{max}) predicted by the HF FEA model and the BFMT-DNN model at the D_{test} dataset showing the accuracy of the BF1 and BF2 models approximation. Figure 4.17 illustrates the relative errors for the D_{test} dataset.

The average relative error $\varepsilon_{\hat{f},avg} = \left|\hat{y} - y_{ref}\right| / y_{ref}$ is used to quantify the accuracy of the predictions, with y_{ref} the reference value considered and $\hat{y}$ the predicted value. The BF1 surrogate model has a mean absolute error $\varepsilon_{\hat{\sigma},avg} = 0.76\%$ and $\varepsilon_{\hat{\delta},avg} = 0.56\%$ for the σ_{max} and δ_{max} output, respectively.
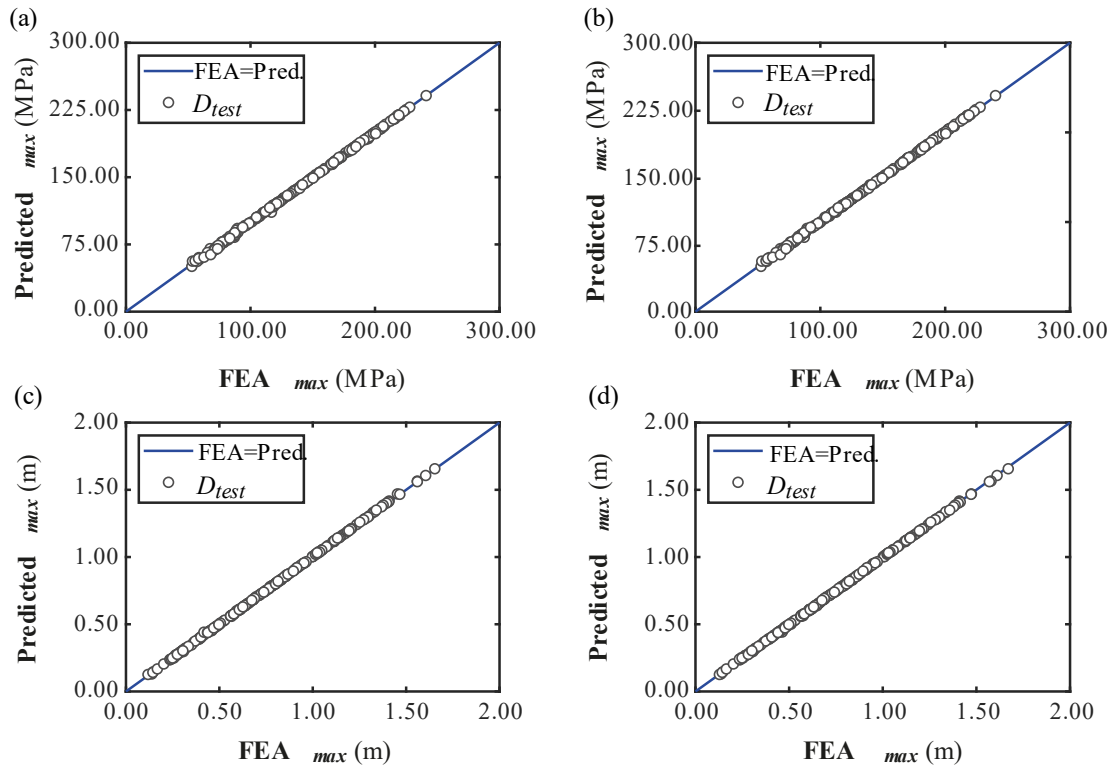
Figure 4.16 – Approximation of the OWT problem outputs in D_{test}. On the left side, the (a) σ_{max} and (c) δ_{max} predictions using the BF1 model. The (a) σ_{max} and (c) δ_{max} predictions using the BF2 model are on the right side.

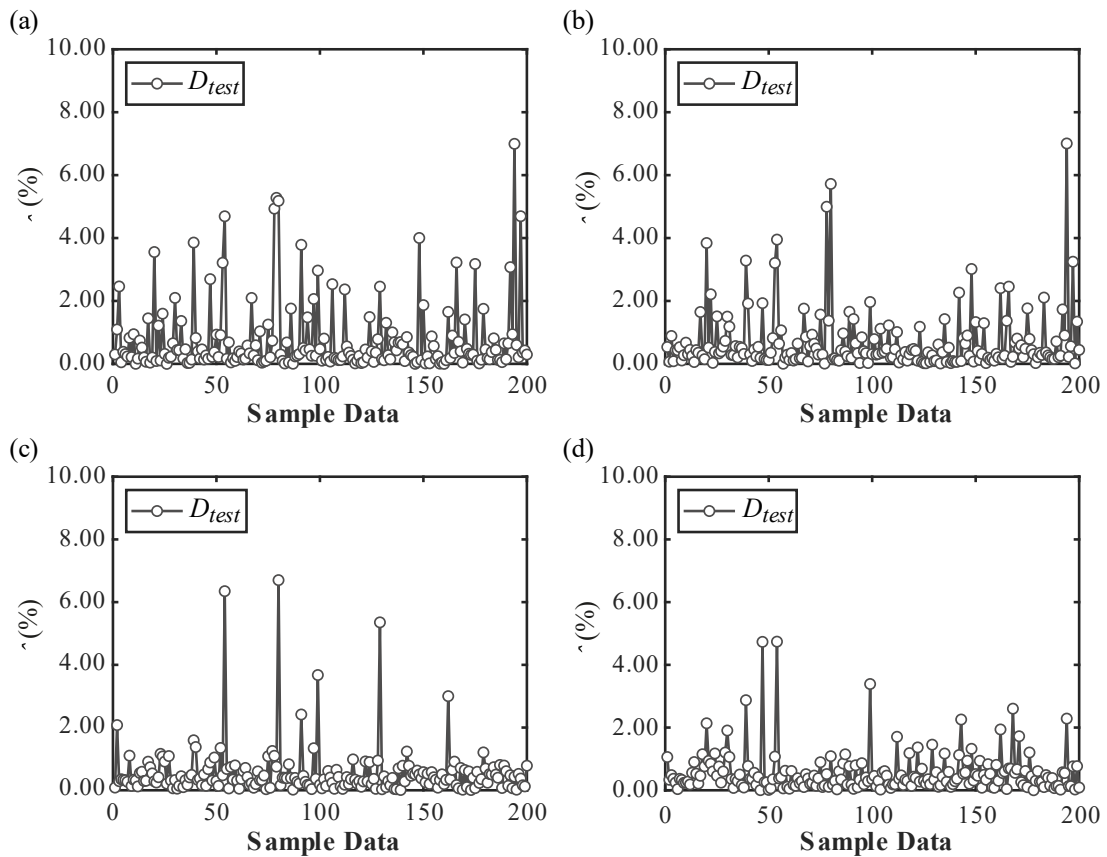*Figure 4.17 – Relative errors of the OWT problem outputs in D_{test}. On the left side, the (a) σ_{max} and (c) δ_{max} predictions using the BF1 model. On the right side, the (a) σ_{max} and (c) δ_{max} predictions using the BF2 model.*

The BF2 surrogate model has a mean absolute error $\varepsilon_{\hat{\sigma},avg} = 0.69\%$ and $\varepsilon_{\hat{\delta},avg} = 0.56\%$ for the σ_{max} and δ_{max} output, respectively. For both surrogate models and each output (δ_{max} and σ_{max}) of the BFMT-DNN model, the coefficient of determination is $R^2 > 0.999$. In the trained BF1 for 88.5% and 96.5% of the D_{test} dataset, the absolute error satisfies $\varepsilon_{\hat{\sigma}} \leq 2.0\%$ and $\varepsilon_{\hat{\delta}} \leq 2.0\%$ for the σ_{max} and δ_{max} output, respectively. For the trained BF2, 93% and 96.0% of the D_{test} dataset satisfy the absolute error condition $\varepsilon_{\hat{\sigma}} \leq 2.0\%$ and $\varepsilon_{\hat{\delta}} \leq 2.0\%$. The maximum absolute error in the D_{test} dataset is $\varepsilon_{\hat{\sigma}} = 7.1\%$ and $\varepsilon_{\hat{\sigma}} = 6.7\%$, in the trained BF1. For the trained BF2, the maximum absolute error in the D_{test} dataset is $\varepsilon_{\hat{\sigma}} = 7.0\%$ and $\varepsilon_{\hat{\sigma}} = 4.8\%$. The results presented show that the BF1 and BF2 have similar accuracy predictions in the HF multi-task outputs, and these results imply that these BFMF-DNN surrogate models ensure sufficient accuracy in the global projection of the D_{test} data.

4.4.7 Hyperparameter optimization process

For each combination between the proposed n_{HF} and the λ_{BF} values, the best DNN is determined by applying the BO for tuning and optimizing the hyperparameters. To create the BFMT-DNN model, the NN_{LF} and NN_{BF} stages must be trained to obtain HF outputs through a BF model. Therefore, two neural networks are trained simultaneously and combined into one through an element-wise operation of the layers in an intermediate position between the two stages. In both stages, the hyperparameters: the number of hidden layers, neurons in each hidden layer, initial learning rate, and activation function, are tuned by BO. The category or the lower and upper bounds for real and integer-valued variables are given in Table 4.12.

Table 4.12 – DNN Hyperparameters limits and domain.

Hyperparameter	Limits	Domain
Each Layer Size	[20,200]	Integer
Number of layers (NN_{LF})	[1,3]	Integer
Number of layers (NN_{BF})	[1,3]	Integer
Activation Function	['tanh', 'sigmoid', 'relu', 'elu']	Categorical
Learning rate	[1e-4,5e-3]	Real

Each optimization iteration is stopped after reaching any of the following criteria: achieving a defined loss function *Loss*=0.004 for the training D_{tr} prediction or a fixed number of 30000 epochs for each stage. The BO algorithm stops after reaching the following: a fixed number of 50 iterations or a fixed time is defined as 7.2×10^5 seg. Therefore, the hyperparameters are optimized throughout the layers, and global optimization is executed across the validation set, thereby reducing the time taken and improving the model's

performance. A summary of the best hyperparameter combinations out of 50 iterations for the BFMT-DNN models is shown in Table 4.13. Each hyperparameter is described for the two stages NN_{LF} and NN_{BF}.

Table 4.13 – Tunned hyperparameters by Bayesian Optimization.

Model	λ_{BF}	Number of layers (NN_{LF}, NN_{BF})	Hidden Layer's Size $[(NN_{LF}), (NN_{BF})]$	Activation Function (NN_{LF}, NN_{BF})	Learning rate $[\times 10^{-3}]$ (NN_{LF}, NN_{BF})
BF1	2	(1,2)	[(162 0), (85 31)]	(Relu, Relu)	(0.21, 0.32)
	3	(1,2)	[(178 0), (98 25)]	(Relu, Elu)	(0.23, 0.72)
	4	(1,2)	[(125 0), (171 36)]	(Relu, Elu)	(0.88, 0.17)
	5	(2, 2)	[(36 50), (102 34)]	(Relu, Elu)	(0.24, 0.92)
	6	(2, 2)	[(42 54), (91 22)]	(Relu, Elu)	(0.28, 2.40)
BF2	2	(1,2)	[(142 0), (63 25)]	(Relu, Elu)	(0.55, 2.10)
	3	(1,2)	[(166 0), (54 20)]	(Relu, Elu)	(0.64, 4.00)
	4	(2, 2)	[(102 132), (80 36)]	(Elu, Elu)	(2.70, 1.20)
	5	(2, 2)	[(83 164), (121 40)]	(Elu, Elu)	(2.05, 0.81)
	6	(2, 2)	[(51 184), (104 63)]	(Elu, Elu)	(2.70, 1.20)

Based on Table 4.13 results, it is noted that no optimized DNN reached the upper limit of 3 layers for the stages. Furthermore, the LF stages have one layer in the BF1-optimized models with $\lambda_{BF}\leq3$. For each stage in the BFMT-DNN model, one activation function is optimized and used in all hidden layers present in these stages. In observed cases with Sigmoid and Hyperbolic tangent activation functions, the training is compromised because the number of iterations necessary to achieve the defined loss typically exceeds the fixed number of epochs for each stage. On the contrary, they occur for cases with ELU and RELU that presented similar performances and alternated in their choices during optimization.

The selected BFMT-DNN model is developed by tuning its hyperparameters. The level of precision is used to determine the optimal model using the tuned parameters as given in Table 4.13. Figure 4.18 shows the progress of the BF-DNN hyperparameter optimization for the selected model. The plot in Fig. 4.18 corresponds to the minimum objective obtained for the D_{val} with the trained BFMT-DNN, for the BF1 (Fig. 4.18(a)) and BF2 (Fig. 4.18(b)) models considering the loss function.

The red line in Fig. 4.18 corresponds to the estimated target values, i.e., the result of the surrogate probability model that determines where to evaluate next. The black line is the observed target value of the true objective function. Due to the agreement between the true and estimated values, we can assume that the estimated function is sufficiently compelling, attesting to BO's efficiency in choosing BFMT-DNN hyperparameters. The best workable point is where the observed objective value is the lowest. For the BF1, the score for the minimum objective observed of $Loss_{val}=9.62\times10^{-3}$ for the trained BFMT-DNN in the 8th

iteration. For the BF2, the score for the minimum objective observed of $Loss_{val}$=8.90×10^{-3} for the trained BF-DNN in the 24th iteration.
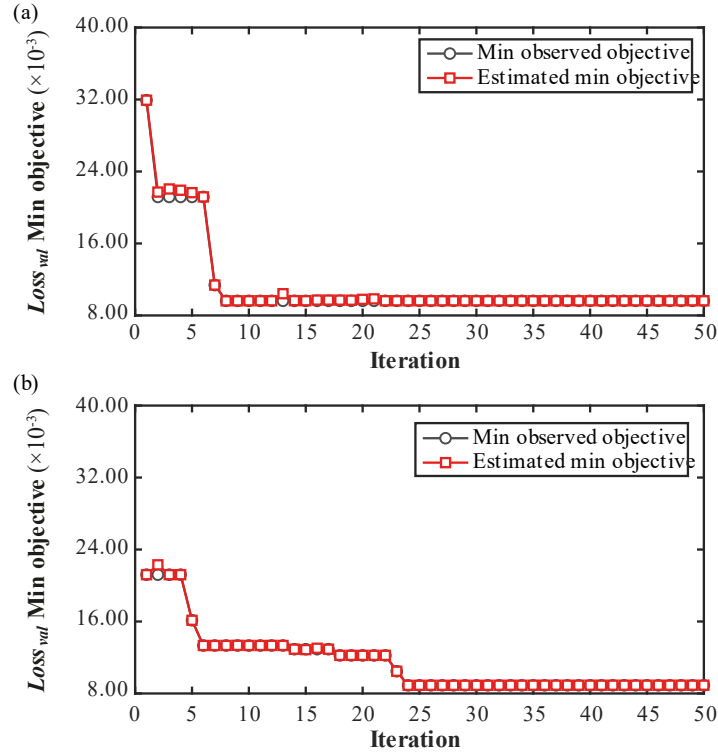

Figure 4.18 –Bayesian optimization progress for tuning hyperparameters of (a) BF1 and (b) BF2.

4.4.8 Reliability Analysis using BF-DNN model

The assessment of efficiency, accuracy, and fidelity of the BFMT-DNN surrogate model in the offshore engineering case presented in this paper is also performed through a systematic comparison of their accuracy in the approximation of the reliability index of the application problem based on SUS. The target $P_f = 10^{-4}$, corresponding to a reliability index for OWT support structures typically β=3.71 [150]. Hence, in this example, the P_f predictions for the σ_{max} and δ_{max} design constraints are based on the safety factors γ_s and γ_d.

The results obtained from the non-intrusive $\hat{P}_f$, and consequently $\hat{\beta}$ predictions, using the BFMT-DNN+SUS over the safety factors variation in each design constraint are depicted in Fig. 4.19. The SUS analyses are submitted to 50 simulation cycles resulting in the averages predictions of $P_{f\,avg,\sigma}$ and $P_{f\,avg,\delta}$ for the σ_{max} and δ_{max} design constraints $\hat{P}_f$ predictions respectively. The averages predictions of $\beta_{avg,\sigma}$ and $\beta_{avg,\delta}$ for the σ_{max} and δ_{max} design constraints $\hat{\beta}$ predictions respectively. The $COV_{\hat{P}_f} \leq 0.1$ is acquired when N_s is set to 3×10^3 and p_0= 0.1, also of the other intermediate failure events size.
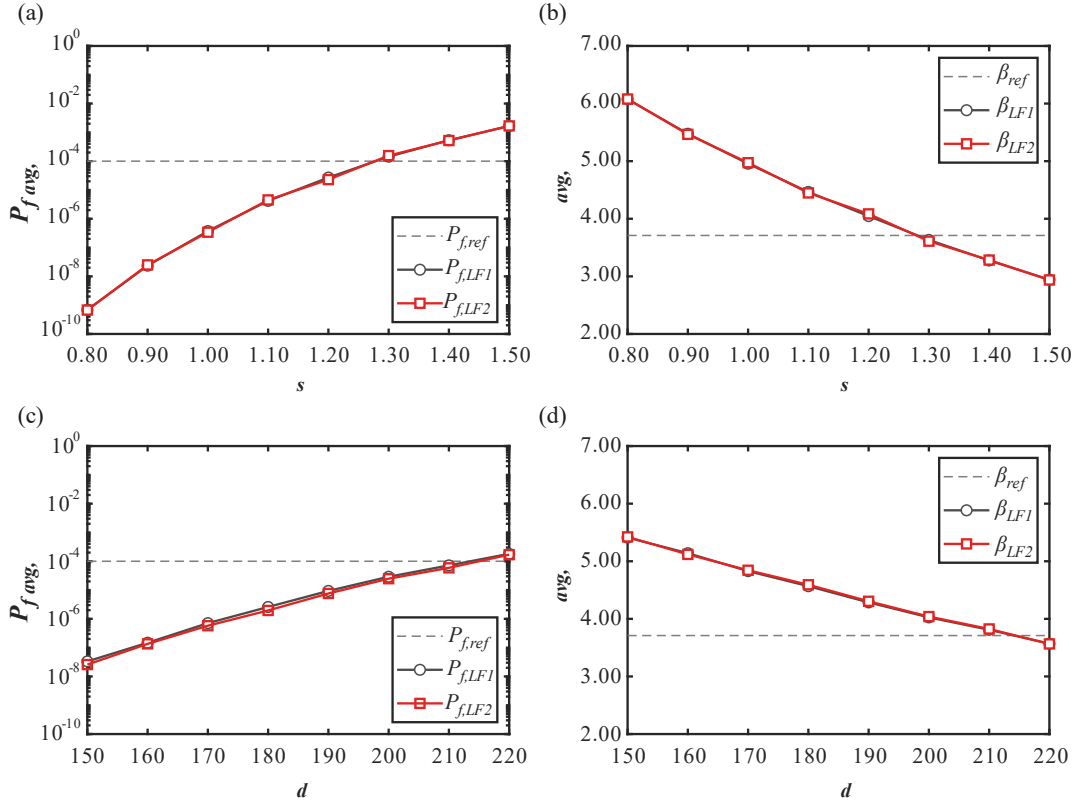
Figure 4.19 – Reliability results. On the left side, the average P_f for the (a) σ_{max} and (c) δ_{max} PF. On the right side, the average β of the (b) σ_{max} and (d) δ_{max} PF.

For the stress constraint case (ULS), is considered a range for the factor of safety $\gamma_s \in [1.48, 1.52]$, that presents a considerable variation in the $P_{f\,avg,\sigma}$ and $\beta_{avg,\sigma}$ values. According to Figs. 4.19(a) and 4.19 (b), the BFMT-DNN+SUS reliability analysis considering the LF1 and LF2 models presented similar approximations for the $P_{f\,avg,\sigma}$ and $\beta_{avg,\sigma}$ values. It is adopted the reference value of γ_s=1.1 [150] for the present ULS analysis.

Considering the BFMT-DNN+SUS case with LF1, the average probability for the maximum displacement $P_{f\,avg,\sigma}(\sigma_{max} > 286.36\,\text{MPa}) = 4.17 \times 10^{-6}$ with the 95% confidence interval $\text{CI}_{95\%} = [3.55 \times 10^{-6}, 4.19 \times 10^{-6}]$. The corresponding reliability index value is $\beta_{avg,\sigma} = 4.45$. For the BFMT-DNN+SUS case with LF2, the average probability for the maximum displacement $P_{f\,avg,\sigma}(\sigma_{max} > 286.36\,\text{MPa}) = 4.50 \times 10^{-6}$ with the 95% confidence interval $\text{CI}_{95\%} = [3.83 \times 10^{-6}, 5.28 \times 10^{-6}]$, and the corresponding reliability index value is $\beta_{avg,\sigma} = 4.43$.

However considering the reference value of γ_s=1.1 [150] for the present ULS analysis, the SUS could not produce results within the region of σ_{max}>286.36 MPa, which the value of the P_f of the maximum equivalent stress is expected to be neglected.

The deformation constraint case (SLS), is considered a range for the factor of safety $\gamma_d \in [150.0, 220.0]$, that present a considerable variation in the $P_{f\,avg,\delta}$ and $\beta_{avg,\delta}$ values. As depicted in Figs. 4.19(c) and 4.19 (d), the BFMT-DNN+SUS reliability analysis considering the LF1 and LF2 models presented similar approximations for the $P_{f\,avg,\sigma}$ and $\beta_{avg,\sigma}$ values. It is adopted the reference value of γ_d=200.0 [150] for the present SLS analysis.

Considering the BFMT-DNN+SUS case with LF1, the average probability for the maximum displacement $P_{f\,avg,\delta}(\delta_{max} > 0.69\text{m}) = 2.89 \times 10^{-5}$ with the 95% confidence interval $\text{CI}_{95\%} = [2.62 \times 10^{-5}, 3.16 \times 10^{-5}]$. The corresponding reliability index value is $\beta_{avg,\delta} = 4.02$. For the BFMT-DNN+SUS case with LF2, the average probability for the maximum displacement $P_{f\,avg,\delta}(\delta_{max} > 0.69\text{m}) = 2.73 \times 10^{-5}$ with the 95% confidence interval $\text{CI}_{95\%} = [2.47 \times 10^{-5}, 2.98 \times 10^{-5}]$, and the corresponding reliability index value is $\beta_{avg,\delta} = 4.03$.

4.5   CONCLUSIONS

This work presented a Bi-Fidelity Multi-task Deep Neural Network to make explicit Performance Function for reliability analysis associated with Subset simulation. This framework's accuracy and computational efficiency are demonstrated by its application in an offshore wind turbine reliability analysis in several Limit State scenarios. The reliability of the structure was investigated by defining the uncertainties of the statistical parameters for the structural design variables. Statistical analysis between the model and measurements highlighted the acceptable representation of the wind and wave fields by Wavewatch 3 numerical model.

The use of the low-fidelity samples in the bi-fidelity framework has been proven. In the Finite Element problem, using both proposed LF1 and LF2 models in the low-fidelity scenario allowed similar accuracy in the implicit performances functions predictions and, consequently, in the reliability indicators. In addition, using the same BF-DNN surrogate model to create a multi-task scenario to predict many outputs simultaneously demonstrated the main advantage of the proposed framework. For practical applications, a dataset using

around 500 high and 3000 low-fidelity Finite Element Analysis can be an initial estimative to accurately fit BF-MT models to assess the reliability of the wind turbine with similar design variables and boundary conditions.

Most surrogate-based procedures are a victim of the curse of dimensionality that the computation effort increases dramatically since large populations are required to assess small probabilities or to fit non-linear Performance Function. Based on the results, the optimized BFMT-DNN solves this problem by simplifying the Neural Network architecture using a few Bayesian Optimization iterations.

These results suggest that the proposed bi-fidelity framework can be applied to the reliability analysis of complex structures under rare events, using few high-fidelity associated with many low-fidelity samples to predict HF values. This framework provides an acceptable level of accuracy provided by the Bayesian optimization of the hyperparameters and considerable computation time savings compared to the conventional methods that use only high-fidelity samples.

5   CONCLUSIONS

As introduced in Chapter 1, the need for reducing the computational cost of the reliability problems that use high-fidelity non-linear systems was addressed. In particular, it was focused on real-world shipbuilding and offshore wind turbine structural reliability problems. The limit state-based capacity and structural behavior assessment is a complex process, and simulating such processes are typically a computationally intensive task owing to the non-linear behavior. Moreover, forward propagation of uncertainties through those complex systems requires a large number of simulations and thus is usually computationally prohibitive in a realistic scenario. To address such challenges in the complex process, the main objective of the work presented in this thesis was to develop a novel multi-fidelity and multi-task deep neural network surrogate model to predict multiple outputs in reliability assessment. According to the main objective of this thesis, multi-fidelity modeling formulations were formulated to reduce the computational cost associated with high-fidelity finite element analysis. As mentioned in Chapter 1, the primary objective of the work presented in this thesis broke down into five sub-objectives. The conclusions for these sub-objectives are summarized in this Section.

Chapter 2 and Chapter 3 addressed the issue related to the cost of solving the reliability assessment of the stiffened panel in the deck under axial load. The mesh convergence analysis made it possible to evaluate the correlation between the accuracy of the ultimate strength of each model with the average computational time for the studied meshes. Chapter 4 presented an offshore wind turbine model where two low-fidelity models were evaluated, considering a variation in the mesh size and a simplification of the 3d elements for linear elements of beams. Distance Correlation, frequently used for non-linear variables correlation, was innovatively imported for constructing spatial correlation functions to estimate the relationship between the high and low-fidelity models. Assessing the correlation between the fidelities reduces the complexity of the high and low-fidelity models' choice.

In Chapter 2, the Bi-Fidelity Kriging surrogate model is used to become explicit a highly non-linear stiffened panel performance function to facilitate the reliability analysis. The BF-kriging was combined with the Subset Simulation, allowing the rare event reliability analysis. In the numerical results, after the low-fidelity model selection, the reduction in the computational time provided by the proposed BF-Kriging has been up to 75% of the computational cost in the Finite Element analysis. It was also shown that the BF-kriging method can be cast in the Bayesian framework, and the regression results naturally include an

uncertainty estimation, which is usually desirable. Moreover, it is suitable for many applications as a non-parametric regression tool. The BF-kriging framework results suggest that this surrogate model can be applied to make the performance function explicit in the reliability analysis of structures, using high-fidelity associated with low-fidelity samples to predict HF values, providing an acceptable level of accuracy and considerable computation time savings compared to the conventional methods that use only high-fidelity samples.

Nevertheless, the Co-Kriging could be more suitable for high-dimensional drawbacks due to the curse of dimensionality. Thus, Chapter 3 focused on implementing a BF-DNN for the solution of high dimensional and high non-linear problems. The BF-DNN was used to make explicit PF for reliability analysis. The Subset Simulation association allowed the rare event reliability analysis estimation using conditional probabilities. The BF-DNN solves the curse of dimensionality problem by simplifying the neural network architecture using a few Bayesian Optimization iterations. The generalizability of the BF-DNN model and Subset Simulation association is accessed in two academic case studies and on reliability assessment of the ultimate compressive strength of plate elements typical of the deck structure of double-hull oil tanker ships described by a non-linear Finite Element analysis. These applications showed excellent results for high dimensional and non-linear problems considering rare events. However, as a disadvantage, this method for solving problems with low variables or non-complex Performance Functions should be regarded with reservation and become a limitation in a sparse high-fidelity dataset analysis.

Chapter 4 presented a BFMT-DNN to make explicit PF for reliability analysis associated with Subset simulation. This framework's accuracy and computational efficiency were demonstrated by its application in a wind turbine reliability analysis in two limit-state scenarios. Two low-fidelity frameworks were used to predict the maximum von Misses stress and displacement. They obtained similar results in the reliability assessment of the offshore wind turbine under different loads. In addition, using the same BF-DNN surrogate model to create a multi-task scenario to predict many outputs simultaneously demonstrated the main advantage of the proposed framework. Moreover, the reliability of the structure was investigated by defining the uncertainties of the statistical parameters for the structural design variables. Statistical analysis between the model and measurements highlighted the acceptable representation of the wind and wave fields by the WaveWatch 3 numerical model.

These results suggest that the proposed BFDNN and BFMT-DNN frameworks can be applied to the reliability analysis of complex structures under rare events, using a few HF

associated with many low-fidelity samples to predict high-fidelity values. These frameworks provide an acceptable level of accuracy provided by the Bayesian optimization of the hyperparameters and considerable computation time savings compared to the conventional methods that use only HF samples.

5.1 FUTURE WORK

Some possible future research works are summarized as follows:

- Considering the stiffened panel example in the multi-fidelity framework, the aspect ratio of the plate is about six. In this case, for future analysis, 1/2+1+1/2 span model with periodical boundary condition can be used to reproduce the plate buckling in the high-fidelity model.

- The current design of offshore wind turbines follows mainly the IEC 61400-3 standard. The list of DLCs implied for this standard is comprehensive. The multi-task scenario can also be analyzed considering DLCs focused on power production. The buckling ultimate limit state of the structure, the fatigue limit state, and the Frequency (modal) limit state can be assessed to identify the most critical among them.

- It is emphasized that the analyses of the wind turbine and recommendations presented do not consider the foundation aspect of the design. This issue deserves future investigation.

- As future work, by use of the conditional samples generated by MCMC simulation and SUS, a numerical method can be used to estimate and realize the reliability sensitivity of the conditional failure probabilities and estimate the influence of the random variables in the reliability assessment.

- For a long-term future work, a Multi-fidelity Convolutional Neural Networks framework can be used to predict HF outputs treating the multi-fidelity data as image data and processes them using Convolutional Neural Networks.

REFERENCES

[1] Melchers RE, Beck AT. Structural Reliability Analysis and Prediction: Wiley; 2018.

[2] Chojaczyk AA, Teixeira AP, Neves LC, Cardoso JB, Guedes Soares C. Review and application of Artificial Neural Networks models in reliability analysis of steel structures. Structural Safety. 2015;52:78-89.

[3] Gaspar B, Teixeira AP, Guedes Soares C. Adaptive surrogate model with active refinement combining Kriging and a trust region method. Reliability Engineering & System Safety. 2017;165:277-91.

[4] Matthies HG, Brenner CE, Bucher CG, Guedes Soares C. Uncertainties in probabilistic numerical analysis of structures and solids Stochastic finite elements. Structural Safety. 1997;19:283-336.

[5] Gaspar B, Naess A, Leira BJ, Guedes Soares C. System Reliability Analysis by Monte Carlo Based Method and Finite Element Structural Models. Journal of Offshore Mechanics and Arctic Engineering. 2014;136:031603.

[6] Roy A, Chakraborty S. Support vector regression based metamodel by sequential adaptive sampling for reliability analysis of structures. Reliability Engineering & System Safety. 2020;200:106948.

[7] Narváez NS, Rojas NR, Evangelista Jr F. Reliability analyses of shear strengthened RC beams with externally bonded fiber reinforced polymer. Materials and Structures. 2020;53:31.

[8] Gaspar B, Teixeira AP, Guedes Soares C. Assessment of the efficiency of Kriging surrogate models for structural reliability analysis. Probabilistic Engineering Mechanics. 2014;37:24-34.

[9] Yoo K, Bacarreza O, Aliabadi MHF. A novel multi-fidelity modelling-based framework for reliability-based design optimisation of composite structures. Engineering with Computers. 2020;38:595–608.

[10] Yun W, Lu Z, Wang L, Feng K, He P, Dai Y. Error-based stopping criterion for the combined adaptive Kriging and importance sampling method for reliability analysis. Probabilistic Engineering Mechanics. 2021;65:103131.

[11] Guo Q, Liu Y, Chen B, Zhao Y. An active learning Kriging model combined with directional importance sampling method for efficient reliability analysis. Probabilistic Engineering Mechanics. 2020;60:103054.

[12] Hurtado JE. Filtered importance sampling with support vector margin: A powerful method for structural reliability analysis. Structural Safety. 2007;29:2-15.

[13] Au S-K, Beck JL. Estimation of small failure probabilities in high dimensions by subset simulation. Probabilistic Engineering Mechanics. 2001;16:263-77.

[14] Dhulipala SLN, Shields MD, Chakroborty P, Jiang W, Spencer BW, Hales JD, et al. Reliability estimation of an advanced nuclear fuel using coupled active learning, multifidelity modeling, and subset simulation. Reliability Engineering & System Safety. 2022;226:108693.

[15] Papaioannou I, Betz W, Zwirglmaier K, Straub D. MCMC algorithms for Subset Simulation. Probabilistic Engineering Mechanics. 2015;41:89-103.

[16] Bourinet JM, Deheeger F, Lemaire M. Assessing small failure probabilities by combined subset simulation and Support Vector Machines. Structural Safety. 2011;33:343-53.

[17] Zhang X, Quek ST. Efficient subset simulation with active learning Kriging model for low failure probability prediction. Probabilistic Engineering Mechanics. 2022;68:103256.

[18] Abaei MM, Abbassi R, Garaniya V, Chai S, Khan F. Reliability assessment of marine floating structures using Bayesian network. Applied Ocean Research. 2018;76:51-60.

[19] Zhang K, Collette M. Predicting growth and interaction of multiple cracks in structural systems using Dynamic Bayesian Networks. Marine Structures. 2022;86:103271.

[20] Han Y, Zhen X, Moan T, Huang Y. Real time prediction of operational safety limits for dynamic positioning of an FPSO in a Deepwater Artificial Seabed system. Marine Structures. 2021;80:103093.

[21] Evangelista Jr F, Afanador Garcia N. A chaos expansion polynomial approach to uncertainty analysis in viscoelastic structural elements (in spanish). Dyna. 2016;83:172–82.

[22] Evangelista Junior F, Almeida IF. Machine learning RBF-based surrogate models for uncertainty quantification of age and time-dependent fracture mechanics. Engineering Fracture Mechanics. 2021;258:108037.

[23] Morató A, Sriramula S. Calibration of safety factors for offshore wind turbine support structures using fully coupled simulations. Marine Structures. 2021;75:102880.

[24] Vosooghi N, Sriramula S, Ivanović A. Response surface based reliability analysis of critical lateral buckling force of subsea pipelines. Marine Structures. 2022;84:103246.

[25] Shi X, Teixeira ÂP, Zhang J, Guedes Soares C. Kriging response surface reliability analysis of a ship-stiffened plate with initial imperfections. Structure and Infrastructure Engineering. 2014;11:1450-65.

[26] Bucher C, Most T. A comparison of approximate response functions in structural reliability analysis. Probabilistic Engineering Mechanics. 2008;23:154-63.

[27] Hosni Elhewy A, Mesbahi E, Pu Y. Reliability analysis of structures using neural network method. Probabilistic Engineering Mechanics. 2006;21:44-53.

[28] Forrester AIJ, Sóbester A, Keane AJ. Multi-fidelity optimization via surrogate modelling. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2007;463:3251-69.

[29] Liu X, Zhao W, Wan D. Multi-fidelity Co-Kriging surrogate model for ship hull form optimization. Ocean Engineering. 2022;243:110239.

[30] Yi J, Cheng Y, Liu J. A novel fidelity selection strategy-guided multifidelity kriging algorithm for structural reliability analysis. Reliability Engineering & System Safety. 2022;219:108247.

[31] Song X, Lv L, Sun W, Zhang J. A radial basis function-based multi-fidelity surrogate model: exploring correlation between high-fidelity and low-fidelity models. Structural and Multidisciplinary Optimization. 2019;60:965-81.

[32] Guo Q, Hang J, Wang S, Hui W, Xie Z. Buckling optimization of variable stiffness composite cylinders by using multi-fidelity surrogate models. Thin-Walled Structures. 2020;156:107014.

[33] Zhang C, Song C, Shafieezadeh A. Adaptive reliability analysis for multi-fidelity models using a collective learning strategy. Structural Safety. 2022;94:102141.

[34] Abdallah I, Lataniotis C, Sudret B. Parametric hierarchical kriging for multi-fidelity aero-servo-elastic simulators — Application to extreme loads on wind turbines. Probabilistic Engineering Mechanics. 2019;55:67-77.

[35] Liao X, Carin L. Radial Basis Function Network for Multi-task Learning. Advances in Neural Information Processing Systems. 2005;18:1-8.

[36] Xue Y, Liao X, Carin L, Krishnapuram B. Multi-Task Learning for Classification with Dirichlet Process Priors. Journal of Machine Learning Research. 2007;8:35-63.

[37] Kapil P, Ekbal A. A deep neural network based multi-task learning approach to hate speech detection. Knowledge-Based Systems. 2020;210:106458.

[38] Gaspar B, Teixeira AP, Guedes Soares C, Wang G. Assessment of IACS-CSR implicit safety levels for buckling strength of stiffened panels for double hull tankers. Marine Structures. 2011;24:478-502.

[39] Bucher CG, Bourgund U. A fast and efficient response surface approach for structural reliability problems. Structural Safety. 1990;7:57-66.

[40] Rajashekhar MR, Ellingwood BR. A new look at the response surface approach for reliability analysis. Structural Safety. 1993;12:205-20.

[41] Kmiecik M, Guedes Soares C. Response surface approach to the probability distribution of the strength of compressed plates. Marine Structures. 2002;15:139-56.

[42] Moharrami MJ, Shiri H. Reliability assessment of drag embedment anchors in clay for catenary mooring systems. Marine Structures. 2018;58:342-60.

[43] Dong Y, Garbatov Y, Guedes Soares C. Improved effective notch strain approach for fatigue reliability assessment of load-carrying fillet welded cruciform joints in low and high cycle fatigue. Marine Structures. 2021;75:102849.

[44] Echard B, Gayton N, Lemaire M. AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation. Structural Safety. 2011;33:145-54.

[45] Gaidai O, Fu S, Xing Y. Novel reliability method for multidimensional nonlinear dynamic systems. Marine Structures. 2022;86:103278.

[46] Pellegrini R, Iemma U, Leotardi C, Campana EF, Diez M. Multi-fidelity Adaptive Global Metamodel of Expensive Computer Simulations. IEEE Congress on Evolutionary Computation (CEC). Vancouver2016.

[47] Thenon A, Gervais V, Ravalec ML. Multi-fidelity meta-modeling for reservoir engineering - application to history matching. Computational Geosciences. 2016;20:1231-50.

[48] Liu H, Ong Y-S, Cai J, Wang Y. Cope with diverse data structures in multi-fidelity modeling: A Gaussian process method. Engineering Applications of Artificial Intelligence. 2018;67:211-25.

[49] Kennedy MC, O'Hagan A. Predicting the output from a complex computer code when fast approximation are available. Biometrika. 2000;87:14.

[50] Myers DE. Matrix formulation of co-kriging. Mathematical Geology. 1982;14.

[51] Skandalos K, Chakraborty S, Tesfamariam S. Seismic reliability analysis using a multi-fidelity surrogate model: Example of base-isolated buildings. Structural Safety. 2022;97:102222.

[52] Qian PZG, Wu CFJ. Bayesian Hierarchical Modeling for Integrating Low-Accuracy and High-Accuracy Experiments. Technometrics. 2008;50:192-204.

[53] Chung HS, Alonso J. Using gradients to construct cokriging approximation models for high-dimensional design optimization problems. 40th AIAA Aerospace Sciences Meeting & Exhibit2002. p. 317.

[54] Han Z-H, Görtz S, Zimmermann R. Improving variable-fidelity surrogate modeling via gradient-enhanced kriging and a generalized hybrid bridge function. Aerospace Science and Technology. 2013;25:177-89.

[55] Krishnan KVV, Ganguli R. Multi-fidelity analysis and uncertainty quantification of beam vibration using co-kriging interpolation method. Applied Mathematics and Computation. 2021;398:125987.

[56] Zhang X, Xie F, Ji T, Zhu Z, Zheng Y. Multi-fidelity deep neural network surrogate model for aerodynamic shape optimization. Computer Methods in Applied Mechanics and Engineering. 2021;373:113485.

[57] Aruna A, Ganguli R. Multi-fidelity response surfaces for uncertainty quantification in beams using coarse and fine finite element discretizations. International Journal for Computational Methods in Engineering Science and Mechanics. 2020:1-20.

[58] Guedes Soares C, Dogliani M, Ostergaard C, Parmentier G, Pedersen PT. Reliability based ship structural design. Transactions - Society of Naval Architects and Marine Engineers. 1996;104:357-89.

[59] Chen N-Z, Guedes Soares C. Reliability assessment for ultimate longitudinal strength of ship hulls in composite materials. Probabilistic Engineering Mechanics. 2007;22:330-42.

[60] Teixeira AP, Guedes Soares C. Reliability analysis of a tanker subjected to combined sea states. Probabilistic Engineering Mechanics. 2009;24:493-503.

[61] Xu MC, Teixeira AP, Guedes Soares C. Reliability assessment of a tanker using the model correction factor method based on the IACS-CSR requirement for hull girder ultimate strength. Probabilistic Engineering Mechanics. 2015;42:42-53.

[62] Zayed A, Garbatov Y, Guedes Soares C. Time-variant reliability assessment of ship structures with fast integration techniques. Probabilistic Engineering Mechanics. 2013;32:93-102.

[63] Gaspar B, Guedes Soares C. Hull girder reliability using a Monte Carlo based simulation method. Probabilistic Engineering Mechanics. 2013;31:65-75.

[64] Gaspar B, Bucher C, Guedes Soares C. Reliability analysis of plate elements under uniaxial compression using an adaptive response surface approach. Ships and Offshore Structures. 2014;10:145-61.

[65] Gaspar B, Naess A, Leira BJ, Guedes Soares C. System reliability analysis of a stiffened panel under combined uniaxial compression and lateral pressure loads. Structural Safety. 2012;39:30-43.

[66] Leheta HW, Mansour AE. Reliability-based method for optimal structural design of stiffened panels. Marine Structures. 1997;10:323-52.

[67] Guedes Soares C. Stochastic Models of Load Effects for the Primary Ship Structure. Structural Safety. 1990;8:16.

[68] Guedes Soares C. Combination of primary load effects in ship structures. Probabilistic Engineering Mechanics. 1992;7:9.

[69] IACS. Common structural rules for Bulk Carriers and Oil Tankers. In: Societies IAoC, editor.2018.

[70] Xu MC, Fujikubo M, Guedes Soares C. Influence of Model Geometry and Boundary Conditions on the Ultimate Strength of Stiffened Panels Under Uniaxial Compressive Loading. Journal of Offshore Mechanics and Arctic Engineering. 2013;135.

[71] Paik JK, Kim BJ, Seo JK. Methods for ultimate limit state assessment of ships and ship-shaped offshore structures: Part II stiffened panels. Ocean Engineering. 2008;35:271-80.

[72] Xu MC, Yanagihara D, Fujikubo M, Guedes Soares C. Influence of boundary conditions on the collapse behaviour of stiffened panels under combined loads. Marine Structures. 2013;34:205-25.

[73] Hørte T, Wang G, White N. Calibration of the Hull Girder Ultimate Capacity Criterion for Double Hull Tankers. 10th International Symposium on Practical Design of Ships and Other Floating Structures. Texas2007.

[74] IACS. Background document – Section 9/1, design verification, hull girder ultimate strength. Common structural rules for double hull oil tankers. In: International Association of Classification Societies L, editor.2006.

[75] Amlashi H, Hørte T, Steen E, J. K. Probabilistic assessment of stiffened panel strength - an interactive tool using proban and puls. ASME 2010 29th International Conference on Ocean, Offshore and Arctic Engineering. China: ASME; 2010. p. 1021-30.

[76] Guedes Soares C, Teixeira AP. Structural reliability of two bulk carrier designs. Marine Structures. 2000;13:107-28.

[77] Forrester A, Sobester A, Keane A. Engineering Design via Surrogate Modelling_ A Practical Guide (2008) - libgen.lc: Wiley; 2008.

[78] Xing W, Razi M, Kirby RM, Sun K, Shah AA. Greedy nonlinear autoregression for multifidelity computer models at different scales. Energy and AI. 2020;1:100012.

[79] Toal DJJ. Some considerations regarding the use of multi-fidelity Kriging in the construction of surrogate models. Structural and Multidisciplinary Optimization. 2015;51:1223-45.

[80] Shi M, Lv L, Sun W, Song X. A multi-fidelity surrogate model based on support vector regression. Structural and Multidisciplinary Optimization. 2020;61:2363-75.

[81] Székely GJ, Rizzo ML. Brownian distance covariance. The Annals of Applied Statistics. 2009;3:1236-65.

[82] Székely GJ, Rizzo ML, Bakirov NK. Measuring and testing dependence by correlation of distances. The Annals of Statistics. 2007;35:2769-94.

[83] Székely GJ, Rizzo ML. The distance correlation t-test of independence in high dimension. Journal of Multivariate Analysis. 2013;117:193-213.

[84] Forrester AI, Sóbester AK, A. J. Multi-fidelity optimization via surrogate modelling. Proceedings of the royal society of london A: mathematical physical and engineering sciences. 2007;463:3251-69.

[85] Kroetz HM, Moustapha M, Beck AT, Sudret B. A Two-Level Kriging-Based Approach with Active Learning for Solving Time-Variant Risk Optimization Problems. Reliability Engineering & System Safety. 2020;203:107033.

[86] Pendola M, Mohamed A, Lemaire M, Hornet P. Combination of finite element and reliability methods in nonlinear fracture mechanics. Reliability Engineering & System Safety. 2000;70:15-27.

[87] Kang S-C, Koh H-M, Choo JF. An efficient response surface method using moving least squares approximation for structural reliability analysis. Probabilistic Engineering Mechanics. 2010;25:365-71.

[88] Taflanidis AA, Cheung S-H. Stochastic sampling using moving least squares response surface approximations. Probabilistic Engineering Mechanics. 2012;28:216-24.

[89] Kaymaz I, McMahon CA. A response surface method based on weighted regression for structural reliability analysis. Probabilistic Engineering Mechanics. 2005;20:11-7.

[90] Li X, Zhu H, Chen Z, Ming W, Cao Y, He W, et al. Limit state Kriging modeling for reliability-based design optimization through classification uncertainty quantification. Reliability Engineering & System Safety. 2022;224:108539.

[91] Zhan H, Xiao N-C, Ji Y. An adaptive parallel learning dependent Kriging model for small failure probability problems. Reliability Engineering & System Safety. 2022;222:108403.

[92] Bourinet JM. Rare-event probability estimation with adaptive support vector regression surrogates. Reliability Engineering & System Safety. 2016;150:210-21.

[93] Lee S. Monte Carlo simulation using support vector machine and kernel density for failure probability estimation. Reliability Engineering & System Safety. 2021;209:107481.

[94] Hawchar L, El Soueidy C-P, Schoefs F. Principal component analysis and polynomial chaos expansion for time-variant reliability problems. Reliability Engineering & System Safety. 2017;167:406-16.

[95] Liu Z, Lesselier D, Sudret B, Wiart J. Surrogate modeling based on resampled polynomial chaos expansions. Reliability Engineering & System Safety. 2020;202:107008.

[96] Li M, Sadoughi M, Hu Z, Hu C. A hybrid Gaussian process model for system reliability analysis. Reliability Engineering & System Safety. 2020;197:106816.

[97] Wang Y, Gao Y, Liu Y, Ghosh S, Subber W, Pandita P, et al. Bayesian-entropy gaussian process for constrained metamodeling. Reliability Engineering & System Safety. 2021;214:107762.

[98] Lee J, Mitici M. Multi-objective design of aircraft maintenance using Gaussian process learning and adaptive sampling. Reliability Engineering & System Safety. 2022;218:108123.

[99] Xu Z, Saleh JH. Machine learning for reliability engineering and safety applications: Review of current status and future opportunities. Reliability Engineering & System Safety. 2021;211:107530.

[100] Afshari SS, Enayatollahi F, Xu X, Liang X. Machine learning-based methods in structural reliability analysis: A review. Reliability Engineering & System Safety. 2022;219:108223.

[101] Snoek J, Larochelle H, Adams RP. Practical Bayesian optimization of machine learning algorithms. Advances in Neural Information Processing Systems. 2012;25:9.

[102] Wu J, Chen XY, Zhang H, Xiong LD, Lei H, Deng SH. Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimization. Journal of Electronic Science and Technology. 2019;17:26-40.

[103] Giannakis N, Gorgolis N, Hatzilygeroudis I. Bayesian optimization for the design of deep neural networks. 2021 12th International Conference on Information, Intelligence, Systems & Applications (IISA)2021. p. 1-8.

[104] Aydin RC, Braeu FA, Cyron CJ. General Multi-Fidelity Framework for Training Artificial Neural Networks With Computational Models. Frontiers in Materials. 2019;6.

[105] Motamed MJIJfUQ. A multi-fidelity neural network surrogate sampling method for uncertainty quantification. 2020;10.

[106] Liu D, Wang Y. Multi-Fidelity Physics-Constrained Neural Network and Its Application in Materials Modeling. Journal of Mechanical Design. 2019;141.

[107] Meng X, Karniadakis GE. A composite neural network that learns from multi-fidelity data: Application to function approximation and inverse PDE problems. Journal of Computational Physics. 2020;401:109020.

[108] Chen J, Gao Y, Liu Y. Multi-fidelity Data Aggregation using Convolutional Neural Networks. Computer Methods in Applied Mechanics and Engineering. 2022;391:114490.

[109] Lu L, Dao M, Kumar P, Ramamurty U, Karniadakis GE, Suresh S. Extraction of mechanical properties of materials through deep learning from instrumented indentation. Proc Natl Acad Sci U S A. 2020;117:7052-62.

[110] Islam M, Thakur MSH, Mojumder S, Hasan MN. Extraction of material properties through multi-fidelity deep learning from molecular dynamics simulation. Computational Materials Science. 2021;188:110187.

[111] Zhou K, Tang J. Efficient characterization of dynamic response variation using multi-fidelity data fusion through composite neural network. Engineering Structures. 2021;232:111878.

[112] Le Gratiet L, Garnier J. Recursive co-kriging model for design of computer experiments with multiple levels of fidelity. Int J Uncertain Quantif. 2014;4:22.

[113] Dhulipala SLN, Shields MD, Spencer BW, Bolisetti C, Slaughter AE, Labouré VM, et al. Active learning with multifidelity modeling for efficient rare event simulation. Journal of Computational Physics. 2022;468:111506.

[114] Proppe C, Kaupp J. On information fusion for reliability estimation with multifidelity models. Probabilistic Engineering Mechanics. 2022;69:103291.

[115] Peherstorfer B, Willcox K, Gunzburger M. Survey of Multifidelity Methods in Uncertainty Propagation, Inference, and Optimization. SIAM Review. 2018;60:550-91.

[116] Giselle Fernández-Godino M, C. Park C, Kim NH, Haftka RT. Review of Multi-fidelity Models. American Institute of Aeronautics and Astronautics (AIAA). 2016.

[117] Duchi J, Hazan E, Singer Y. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. Journal of Machine Learning Research. 2011;12:2121-59.

[118] Kingma DP, Ba JL. Adam: A method for stochastic optimization. International Conference on Learning Representations 2015. San Diego2017.

[119] Liu X, Wu J, Chen S. A context-based meta-reinforcement learning approach to efficient hyperparameter optimization. Neurocomputing. 2022;478:89-103.

[120] Brochu E, Hoffman MW, de Freitas N. A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning. CoRR. 2010;abs/1012.2599:1-49.

[121] Brochu E, Hoffman MW, de Freitas N. Portfolio Allocation for Bayesian Optimization. CoRR. 2010;abs/1009.5419:1-20.

[122] Klein A, Falkner S, Bartels S, Hennig P, Hutter F. Fast Bayesian Optimization of Machine Learning Hyperparameters on Large Datasets. CoRR. 2016;abs/1605.07079:1-9.

[123] Evangelista F, Alves GdS, Moreira JFA, Paiva GOFd. A global–local strategy with the generalized finite element framework for continuum damage models. Computer Methods in Applied Mechanics and Engineering. 2020;363:112888.

[124] DNV. Structural reliability analysis of marine structures. Classification notes no 306. Det Norske Veritas Classification, AS1992.

[125] Guo M, Manzoni A, Amendt M, Conti P, Hesthaven JS. Multi-fidelity regression using artificial neural networks: Efficient approximation of parameter-dependent output quantities. Computer Methods in Applied Mechanics and Engineering. 2022;389:114378.

[126] GWEC. Global Wind Report 2022. GLOBAL WIND ENERGY COUNCIL; 2023.

[127] Ivanhoe RO, Wang L, Kolios A. Generic framework for reliability assessment of offshore wind turbine jacket support structures under stochastic and time dependent variables. Ocean Engineering. 2020;216:107691.

[128] Shittu AA, Mehmanparast A, Amirafshari P, Hart P, Kolios A. Sensitivity analysis of design parameters for reliability assessment of offshore wind turbine jacket support structures. International Journal of Naval Architecture and Ocean Engineering. 2022;14:100441.

[129] Liao D, Zhu S-P, Correia JAFO, De Jesus AMP, Veljkovic M, Berto F. Fatigue reliability of wind turbines: historical perspectives, recent developments and future prospects. Renewable Energy. 2022;200:724-42.

[130] Meng D, Yang S, Jesus AMPd, Zhu S-P. A novel Kriging-model-assisted reliability-based multidisciplinary design optimization strategy and its application in the offshore wind turbine tower. Renewable Energy. 2023;203:407-20.

[131] Ren C, Aoues Y, Lemosse D, Souza De Cursi E. Ensemble of surrogates combining Kriging and Artificial Neural Networks for reliability analysis with local goodness measurement. Structural Safety. 2022;96:102186.

[132] Shittu AA, Mehmanparast A, Wang L, Salonitis K, Kolios A. Comparative Study of Structural Reliability Assessment Methods for Offshore Wind Turbine Jacket Support Structures. Applied Sciences2020.

[133] Shi X, Teixeira AP, Zhang J, Guedes Soares C. Structural reliability analysis based on probabilistic response modelling using the Maximum Entropy Method. Engineering Structures. 2014;70:106-16.

[134] Bonfiglio L, Perdikaris P, Brizzolara S, Karniadakis GE. Multi-fidelity optimization of super-cavitating hydrofoils. Computer Methods in Applied Mechanics and Engineering. 2018;332:63-85.

[135] Conti P, Guo M, Manzoni A, Hesthaven JS. Multi-fidelity surrogate modeling using long short-term memory networks. Computer Methods in Applied Mechanics and Engineering. 2023;404:115811.

[136] Nitzler J, Biehler J, Fehn N, Koutsourelakis P-S, Wall WA. A generalized probabilistic learning approach for multi-fidelity uncertainty quantification in complex physical simulations. Computer Methods in Applied Mechanics and Engineering. 2022;400:115600.

[137] Meng X, Wang Z, Fan D, Triantafyllou MS, Karniadakis GE. A fast multi-fidelity method with uncertainty quantification for complex data correlations: Application to vortex-induced vibrations of marine risers. Computer Methods in Applied Mechanics and Engineering. 2021;386:114212.

[138] Thai H-T. Machine learning for structural engineering: A state-of-the-art review. Structures. 2022;38:448-91.

[139] Li J, Yin G, Wang X, Yan W. Automated decision making in highway pavement preventive maintenance based on deep learning. Automation in Construction. 2022;135:104111.

[140] Lisboa RC, Teixeira PRF, Fortes CJ. Numerical evaluation of wave energy potential in the south of Brazil. Energy. 2017;121:176-84.

[141] Chawla A, Spindler DM, Tolman HL. Validation of a thirty year wave hindcast using the Climate Forecast System Reanalysis winds. Ocean Modelling. 2013;70:189-206.

[142] Veritas DN. DNV-RP-C205 Recommended Practice—Environmental Conditions and Environmental Loads. 2014.

[143] Kodaira T, Sasmal K, Miratsu R, Fukui T, Zhu T, Waseda T. Uncertainty in wave hindcasts in the North Atlantic Ocean. Marine Structures. 2023;89:103370.

[144] Zieger S, Babanin AV, Erick Rogers W, Young IR. Observation-based source terms in the third-generation wave model WAVEWATCH. Ocean Modelling. 2015;96:2-25.

[145] Vorpahl F, Popko W, Kaufer D. Description of a Basic Model of the "UpWind Reference Jacket" for Code Comparison in the OC4 Project under IEA Wind Annex 30. 2013.

[146] Jonkman JM, Butterfield S, Musial W, Scott G. Definition of a 5-MW Reference Wind Turbine for Offshore System Development. 2009.

[147] Ren C, Aoues Y, Lemosse D, Souza De Cursi E. Comparative study of load simulation approaches used for the dynamic analysis on an offshore wind turbine jacket with different modeling techniques. Engineering Structures. 2021;249:113308.

[148] Damiani R, Song H. A Jacket Sizing Tool for Offshore Wind Turbines Within the Systems Engineering Initiative. Offshore Technology Conference2013. p. OTC-24140-MS.

[149] Tian X, Sun X, Liu G, Deng W, Wang H, Li Z, et al. Optimization design of the jacket support structure for offshore wind turbine using topology optimization method. Ocean Engineering. 2022;243:110084.

[150] Veritas DN. DNV-OS-J101 Design of Offshore Wind Turbine Structures. 2014.

[151] IEC. IEC 61400-3-1 Wind energy generation systems – Part 3-1: Design requirements for fixed offshore wind turbines. 2019.

[152] Gentils T, Wang L, Kolios A. Integrated structural optimisation of offshore wind turbine support structures based on finite element analysis and genetic algorithm. Applied Energy. 2017;199:187-204.

[153] Morató A, Sriramula S, Krishnan N. Kriging models for aero-elastic simulations and reliability analysis of offshore wind turbine support structures. Ships and Offshore Structures. 2018;14:545-58.

[154] Okpokparoro S, Sriramula S. Uncertainty modeling in reliability analysis of floating wind turbine support structures. Renewable Energy. 2021;165:88-108.

[155] Hong HP, Sheng C. Reliability-based calibration of site-specific design typhoon wind and wave loads for wind turbine. Engineering Structures. 2022;270:114885.

[156] Laboratory NRE. OpenFAST documentation. 2022.

[157] IEC. IEC 61400-1 Wind Turbines — Part 1: Design Requirements. 2005.

[158] Reiss RD, Thomas M. Statistical Analysis of Extreme Values: with Applications to Insurance, Finance, Hydrology and Other Fields. 3 ed: Birkhäuser Basel; 2007.

[159] Coles S. An Introduction to Statistical Modeling of Extreme Values. 1 ed: Springer London; 2001.

[160] Yang Q, Li Y, Li T, Zhou X, Huang G, Lian J. Statistical extrapolation methods and empirical formulae for estimating extreme loads on operating wind turbine towers. Engineering Structures. 2022;267:114667.

[161] IEC. IEC CD TS 61400-9 Wind Energy Generation Systems—Part 9: Probabilistic Design Measures for Wind Turbines. International Electrotechnical Commission. Geneva, Switzerland.2023.

[162] Nielsen JS, Toft HS, Violato GO. Risk-Based Assessment of the Reliability Level for Extreme Limit States in IEC 61400-1. Energies2023.

[163] Abhinav KA, Saha N. Stochastic response of jacket supported offshore wind turbines for varying soil parameters. Renewable Energy. 2017;101:550-64.