Surrogate-assisted unified optimization framework for investigating marine structural design under information uncertainty

by

Yan Liu

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Naval Architecture and Marine Engineering)
in The University of Michigan
2016

Doctoral Committee:

Assistant Professor Matthew D. Collette, Chair
Professor Romesh Saigal
Assistant Professor David J. Singer
Professor Nickolas Vlahopoulos

ACKNOWLEDGEMENTS

I want to express my deep gratitude to my advisor, Professor Matthew Collette. This dissertation work would not have been possible without his guidance and mentorship. Throughout my study at University of Michigan, Professor Collette gave me invaluable help and advice with his intelligence, knowledge, and vision. His unwavering optimism always inspires me to aim high and look for new challenges. I am forever grateful for the belief and trust he placed in me.

I would like to thank the faculty who have served in my doctoral committee, Professor Romesh Saigal, Professor Nickolas Vlahopoulos, and Professor David Singer, for their encouragement and critical comments which have improved this dissertation. I appreciate each one of them for lending their expertise and time in examining my dissertation.

I would like to thank Ms. Kelly Cooper, Office of Naval Research, for sponsoring this dissertation work under grant N00014-11-1-0845. The funding support from ONR on my PhD study is gratefully acknowledged.

I would like to thank all the people I have worked with in the Marine Structure Design Lab, Dr. Jiandao Zhu, Dr. Dylan Temple, Thomas Devine, and Mark Groden, for their help and feedbacks on my research work. Thanks also goes to Dr. Hankoo Jeong; his collaboration on the composite structural design helped to increase the exposure of my research.

Last but not least, I want to thank my parents, Fengming Liu and Defang Xiong. Because of their love and support, I have been able to pursue the intellectual chase

in life.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ii

LIST OF FIGURES vii

LIST OF TABLES ix

LIST OF ABBREVIATIONS xi

ABSTRACT xii

CHAPTER

I. Introduction 1

1.1 Research Background 1
1.2 Research Overview 3

II. Background 8

2.1 Introduction 8
2.2 Overview of design optimization problem 9
2.2.1 Deterministic design optimization 9
2.2.2 Design optimization considering uncertainty 12
2.3 Optimization methods 17
2.3.1 Evolutionary algorithm 18
2.3.2 Elitist non-dominated sorting genetic algorithm . . 23
2.4 Early stage design uncertainty: Epistemic uncertainty 26
2.4.1 Epistemic uncertainty models 28
2.5 Surrogate modeling 32
2.5.1 RBF 33
2.5.2 Kriging 35
2.5.3 Sampling for surrogate modeling 37
2.6 Summary 40
2.6.1 Design optimization considering epistemic uncertainty 41

2.6.2 Surrogate-assisted evolutionary optimization 41

III. Robust and reliable design optimization considering interval uncertainty 43

3.1 Introduction 43
3.2 Surrogate-assisted robust design optimization framework . . . 45
3.2.1 Evolutionary optimization with interval uncertainty 45
3.2.2 Surrogate-assisted robust design optimization 47
3.2.3 Robust compartmentalization case study 50
3.3 Robust and reliable structural design considering interval uncertainty 57
3.3.1 Reliability analysis with interval parameters 58
3.3.2 Multi-objective optimization with surrogate model . 60
3.3.3 Box girder structure case study 62
3.4 Summary 67

IV. Surrogate-assisted variable fidelity multi-objective optimization 69

4.1 Introduction 69
4.2 Variable fidelity optimization 71
4.3 Clustered multiple surrogate modeling for VFO 73
4.4 Numerical benchmark problem tests 79
4.4.1 Results for FON, POL and QV with clustered surrogates VFO 79
4.4.2 Comparison test using ZDT problems 81
4.5 Stiffened panel design study 85
4.5.1 Design problem statements 85
4.5.2 Performance metric for evaluating VFO results . . . 90
4.5.3 Variable fidelity optimization results 95
4.6 Summary 101

V. Efficient reliability-constrained structural design optimization including interval uncertainty 103

5.1 Introduction 103
5.2 Reliability analysis with interval uncertainty 107
5.2.1 Interval uncertainty 107
5.2.2 Worst-case reliability index with interval uncertainty 108
5.2.3 Impacts of interval reliability analysis 110
5.3 Trade-off analysis with adaptive surrogate modeling 111
5.3.1 Overview 111
5.3.2 Kriging modeling 112
5.3.3 Worst-case prediction and adaptive refinement . . . 114

5.3.4 Multi-objective optimization with interval variables 117
5.4 Interval reliability benchmark design problem 119
5.5 Composite top-hat stiffened panel structural design 125
5.6 Summary 132

VI. Summary 134

6.1 Conclusions 134
6.2 Recommendations for future work 138

BIBLIOGRAPHY 140

LIST OF FIGURES

Figure

1.1 Uncertainty-based information theory (*Klir*, 2005) 4

2.1 Example of a Pareto optimal front (*Deb et al.*, 2002). 11

2.2 Definition of safety index (*Hasofer and Lind*, 1974). 16

2.3 Typical evolutionary algorithm procedure. 20

2.4 Example of crowding distance metric calculation 25

2.5 Schematic of NSGA-II procedure (*Deb et al.*, 2002) 26

2.6 Example of Fuzzy logic (*Li et al.*, 1996) 30

2.7 A Latin Hypercube Sample with N=6, K=2 (*Stein*, 1987) 38

2.8 General off-line surrogate construction 39

2.9 General on-line surrogate construction 40

3.1 Evans's general design diagram (*Evans*, 1959). 44

3.2 Flow chart of surrogate assisted robust optimization. 49

3.3 Sketch of two compartments. 51

3.4 Cross section view. 52

3.5 Design performance under interval uncertainty. 56

3.6 Reliability analysis using RBF surrogate model. 61

3.7 Pareto front of direct optimization. 65

3.8 Pareto front of surrogate method. 66

4.1 Flow chart of the clustering embedded VFO method. 76

4.2 Optimization results from FON, POL, and QV test problems 82

4.3 Comparison of Zhu's single surrogate with proposed clustering approach on 30 variable ZDT problems 84

4.4 Sketch of independent variables. 87

4.5 Enumeration results using high-fidelity Faulkner strength function. . 93

4.6 Enumeration results using low-fidelity Paik strength function. 94

5.1 Cantilever tube problem (*Du*, 2007). 119

5.2 Cantilever tube problem: comparison of Pareto fronts for exact and approximate reliability analysis (Case 5). 124

5.3 Cantilever tube problem: effect of interval range on the performance of multi-objective optimization. 124

5.4 Composite FRP top-hat stiffened grillage plate. 126

5.5 Schematic of a multi-stiffener grillage. 127

5.6 A sectional view of the CFRP single skin with top-hat stiffeners. . . 128

5.7 CFRP panel problem: comparison of Pareto fronts for exact and approximate reliability analysis. 131

LIST OF TABLES

Table

2.1 Example of single-point crossover 21

2.2 Common basis functions in Radial Basis Functions (RBF) 34

2.3 Correlation functions for Kriging 36

3.1 Standard T-stiffener dimensions 51

3.2 Design variables 52

3.3 SOGA parameters 54

3.4 Optimization results 55

3.5 Design variables description for interval uncertainty MOGA 63

3.6 Description of variables in limit sate function 64

3.7 Computational cost comparison in two optimizations 67

4.1 Test problems description 80

4.2 VFO parameters for FON, POL and QV tests 81

4.3 Optimization Parameters for Comparison Test 85

4.4 Chromosome for Panel Design 86

4.5 Components of the cost model. 88

4.6 VFO updates schedules 96

4.7 VFO Updating Metrics from standard VFO method 96

4.8 VFO updating metrics from proposed method 97

4.9 Case statistics of distance metric 98

4.10 Average wall clock time for one complete run of algorithm over 25 replicates (sec) 99

4.11 Extended VFO run of case o50s5d60 using multiple surrogates . . . 100

5.1 Random variables of the cantilever tube problem 120

5.2 Interval variables of the cantilever tube problem 120

5.3 Design variables of the cantilever tube problem 121

5.4 Determination of worst-case reliability for the cantilever tube problem 122

5.5 Score metrics from 10 independent optimization runs 123

5.6 Distributions of variables in CFRP problem 129

5.7 Design variables of CFRP problem 130

LIST OF ABBREVIATIONS

RDO Robust Design Optimization

RBDO Reliability Based Design Optimization

VFO Variable Fidelity Optimization

MSE Mean Square Error

COV Coefficient of Variation

PSO Particle Swarm Optimization

SBX Simulated Binary Crossover

EA Evolutionary Algorithm

LHS Latin Hypercube Sampling

RBF Radial Basis Functions

UNDEX Underwater Explosion

FORM First Order Reliability Method

SORM Second Order Reliability Method

SOGA Single Objective Genetic Algorithm

MOGA Multi Objective Genetic Algorithm

CFRP Carbon Fiber-Reinforced Plastic

BPA Basic Probability Assignment

ABSTRACT

Surrogate-assisted unified optimization framework for investigating marine structural design under information uncertainty

by

Yan Liu

Chair: Matthew Collette

Structural decisions made in the early stages of marine systems design can have a large impact on future acquisition, maintenance and life-cycle costs. However, owing to the unique nature of early stage marine system design, these critical structure decisions are often made on the basis of incomplete information or knowledge about the design. When coupled with design optimization analysis, the complex, uncertain early stage design environment makes it very difficult to deliver a quantified trade-off analysis for decision making. This work presents a novel decision support method that integrates design optimization, high-fidelity analysis, and modeling of information uncertainty for early stage design and analysis. To support this method this dissertation improves the design optimization methods for marine structures by proposing several novel surrogate modeling techniques and strategies.

The proposed work treats the uncertainties that are sourced from limited information in a non-statistical interval uncertainty form. This interval uncertainty is treated as an objective function in an optimization framework in order to explore the impact

of information uncertainty on structural design performance. In this examination, the potential structural weight penalty regarding information uncertainty can be quickly identified in early stage, avoiding costly redesign later in the design. This dissertation then continues to explore a balanced computational structure between fidelity and efficiency. A proposed novel variable fidelity approach can be applied to wisely allocate expensive high-fidelity computational simulations. In achieving the proposed capabilities for design optimization, several surrogate modeling methods are developed concerning worst-case estimation, clustered multiple meta-modeling, and mixed variable modeling techniques. These surrogate methods have been demonstrated to significantly improve the efficiency of optimizer in dealing with the challenges of early stage marine structure design.

CHAPTER I

Introduction

1.1 Research Background

Early stage structural design has a disproportionally large impact on the success of a ship acquisition, as major decisions regarding construction and total ownership costs are often made in early design stage. Despite the development and growth of the design theories for marine systems, recent unconventional vessel design programs, especially naval vessels, are still struggling to make a robust and reliable preliminary design decision that support later stage design and production. A large part of this struggle is due to the fact that complete information is not available in early stage design for decision makings. The aim of this dissertation is to propose advanced design tools that address the information uncertainty issues regarding early stage marine structural design optimization.

Structural system design is particularly challenging in early stage design. The structural system consists the majority of a ship's lightship weight (*Keane*, 2012). A correct structural weight estimation has a substantial impact over future design cycles. Additionally, a reliable structure is a key component for the marine system to safely fulfill its missions. Therefore, the initial structural decisions regarding materials, configurations and safety margins are critical for the success of a marine system design project.

While conventional structural designs like cargo ships are often similar and rely heavily on prior designs, unconventional marine designs are often first of their types and thus have no access to successful prototypes for design guidance. Innovative marine designs can not use prior design practice and knowledge to guide the early stage design process. Complete information and understanding about the system is rarely available in the early stage. Reversing incorrect decisions made under incomplete information can easily disrupt the design and planning, result in expensive design modifications and cost growth beyond budget. Moreover, designing under incomplete information could bring risk into the structural design. In the early stage, design loads are likely to be an approximate, structural responses are often from simplified model due to the lack of physical model test data. This can lead to inadequate structural designs that are problematic when put into services. Such design issue have manifested in vessels such as the FFG7, CG47, and DDG51 classes, where structures under-performs in extreme operational conditions and structural problems such as cracking are experienced in-service (*Keane*, 2012). These structural problems are costly to repair, more importantly they could significantly impair the missions of navy ships.

Traditional structural design tools are incapable of dealing with this complex early stage design situation. Much of the design failures are due to the inability to handle uncertainties that are related to lack of information in the early design stage. Early stage design requires an effective method to describe this unique information uncertainty, and then achieve robust baseline design that is least sensitive to this information uncertainty. A desired early stage design needs to be stable and at the same time flexible enough to design changes made in the future.

Designing against uncertainty is never a one-dimensional endeavor. Designers need a well-balanced technical solution that is both robust to information uncertainty, and avoid unnecessary conservatism in the design. Moreover, design tools need to help the

designers identify consequences of potential information uncertainty, thus prioritize how to address the technical uncertainties in later design stage. To achieve these goals, it is essential to couple information uncertainty analysis into multi-objective design optimization.

Meanwhile, a more challenging part of early stage analysis is to complete these analyses within a short time frame. Reducing the time-to-solution cycle time has always been an urgent need in early stage structural design. It requires advanced design tools and strategy to allocate the limited high-fidelity simulations available and deliver satisfactory solutions. Addressing the challenging task of early stage design can only be achieved with a combined information uncertainty-aware optimizer and efficient implementation technique. The aim of this thesis is to deliver both collectively, and present a scalable early stage design optimization tool for decision making purposes.

1.2 Research Overview

As discussed, design uncertainty in early stage design can degrade the performance of marine system if not handled properly. Normally uncertainty is treated in stochastic form and optimization techniques are available to consider it in both objective and constraint functions (*Jin and Branke*, 2005). However, sufficient information is usually not available to build a stochastic definition of all sources of uncertainty in early stage design, especially for unconventional innovative design projects. An example from ship structure design is the determination of key hull dimensions. Normally an educated guess of initial dimensions is proposed in early stage design. However, it is subject to change later on in the design when better information becomes available (*Gale*, 2003). The potential variability for such a parameter is very difficult to quantify through stochastic distributions. Therefore applying existing stochastic based design tools may not be sufficient.

Unlike stochastic uncertainty, the specific type of uncertainty in early stage design is epistemic uncertainty, or uncertainty due to lack of information. Such uncertainty can only be reduced by gaining relevant information. Here the underlying relation between uncertainty and information is illustrated in Figure 1.1. The amount of information gained by taking actions such as conducting experiments, collecting data etc., is measured by the difference between a *prior* uncertainty and a *posteriori* uncertainty (*Klir*, 2005). The direct impact of information-related uncertainty on the design project can be very valuable in terms of early stage planning in the design project.

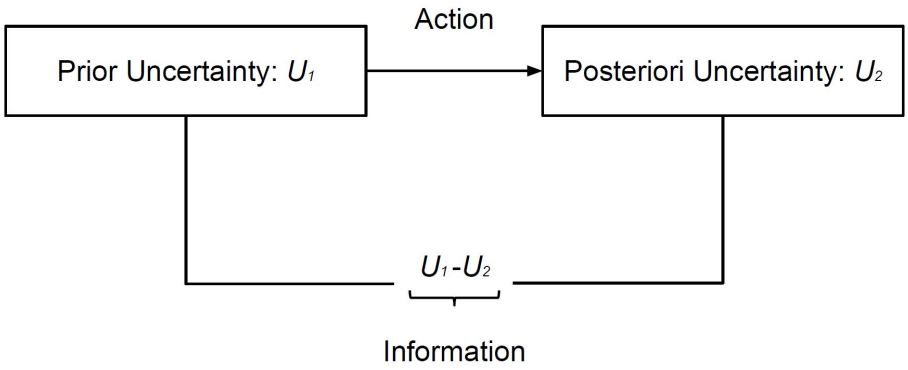


Figure 1.1: Uncertainty-based information theory (*Klir*, 2005)

This dissertation proposed an interval uncertainty formulation to define the uncertainty that arises due to a lack of information. Interval uncertainty provides an appropriate formulation when there is insufficient basis to infer a probabilistic distribution, considering that incorrect assumptions about stochastic information may bring additional risk into the design and result in costly redesign. An interval uncertainty reduction metric is used to represent the information gaining process in Figure 1.1. Feasible structural configurations with different uncertainty intervals need to be defined in order to understand whether gaining information is worthwhile for the design project. To this end, a systematic study of various ranges of uncertainty intervals involved in design optimization is conducted via optimization, aiming to

unveil the overall impact of lack of information in early stage design on the design performance.

Uncertainty analysis and optimization are both computational expensive analyses, where frequent function evaluations are needed. Sometimes these computational expenses may prevent the designers from exploring the trade space in the early stage design. Surrogate modeling is often suggested as an alternative method to improve computational efficiency. Despite rapid growth on surrogate modeling, it still struggles in certain critical modeling problems such as large sampling size models and model updating. This dissertation presents the development of novel surrogate model methods and construction strategies that can be used in early stage design optimization given significant information uncertainty.

Chapter II contains the background techniques that the proposed optimization frameworks are built upon. It reviews the single and multi-objective optimization process, epistemic uncertainty models, surrogate modeling theories and correspondingly the off-line and on-line surrogate construction methods within optimization run. Chapter II concludes with a discussion of the drawbacks of current state-of-art methods in terms of dealing unique early stage design problems.

Chapter III initially explores the robust design optimization against interval uncertainty with a surrogate-assisted optimization framework. The parametric uncertainty in the vessel's early stage compartmentalization layout has been examined through an interval uncertainty metric. The achieved robust baseline structural design is compared with a deterministic design to show its superiority in dealing with a changing design definitions. Interval uncertainty with reliability-constraint is then explored, and a trade-off study on various ranges of interval uncertainties in both structural geometry and reliability simulation is delivered. This trade study is achieved through a surrogate-assisted multi-objective optimization framework.

Chapter IV explores novel surrogate modeling technique to strengthen the capa-

bility for efficient multi-objective optimizations. It considers the existence of various fidelity models in structural simulation and presents a unique solution to better address this issue for early stage multi-objective optimization studies. A variable-fidelity multi-objective optimizer is proposed to efficiently deliver the trade-off analysis. The presented method builds multiple surrogate models on-line as the optimizer runs and successfully addresses large sampling size modeling problem in surrogate-assisted optimization research area.

Chapter V investigates interval uncertainty applied in novel composite structural design via a reliability-based design optimization scheme, to account for incomplete loading and structural response information in early stage structural design. An adaptive surrogate modeling method is developed to accurately estimate reliability analysis in the optimization. The resulting Pareto front indicates how information uncertainty reduces design performance while maintaining a constant structural adequacy.

Chapter VI presents the contributions of dissertation work summarized as follows:

- This work examined the early stage design uncertainty in a new form. It proposed a reducible interval uncertainty form to represent the changeable nature of uncertainty in the design process. The impact of interval uncertainty on design performance is quantified to gain a higher level design information for early stage decision making.

- This work presented a numerically efficient surrogate modeling method to solve the interval analysis problems. It enables the implementation of interval uncertainty modeling within multi-objective evolutionary optimizations with moderate computational resources.

- This work developed clustered multiple surrogate modeling technique to address the variable fidelity problems in structural response models. This method enables multi-objective optimization with limited high-fidelity simulations in early

stage design.

Chapter VI ends with a discussion of the outlook for future applications and extensions of the current research portfolio.

When facing challenging design problems for marine structural systems, this dissertation models the design uncertainty with the awareness of unique nature of early stage design, and merges design optimization with uncertainty analysis through efficient surrogate modeling methods. Overall, this dissertation successfully demonstrated that developed advanced computational analysis tools addressed critical early stage structural design problems.

CHAPTER II

Background

2.1 Introduction

This chapter will cover the fundamental theories that this dissertation work is built upon. The general design optimization problem will be discussed, including proposed design optimization formulations which deal with uncertain environments. The evolutionary algorithm that is used to solve the single and multi-objective optimization problems will be reviewed. Then, the discussion of epistemic uncertainty, the type of uncertainty that is of interest to this research work is presented. Afterwards, surrogate modeling techniques are introduced as a solution to efficient implementation of the type of optimization proposed in this research. This chapter surveys some popular surrogate theories and the common construction methods to implement these surrogate models within optimization. It concludes with a brief discussion on the special needs of surrogate modeling methods for structural design and optimization in early stage design.

2.2 Overview of design optimization problem

2.2.1 Deterministic design optimization

Engineering designs are expected to meet certain demands or performance requirements from customers. One technique for synthesizing designs which will attain these goals is optimization. In design optimization, the mathematical formulation for an optimization problem can be stated as:

$$\begin{aligned} minimize &: f(x) \\ with\ respect\ to &: \mathbf{x} \in \mathbb{R}^n \\ subject\ to &: c_i(x) = 0, i = 1, 2, ..., m \\ &\hat{c}_j(x) \geq 0, j = 1, 2, ..., \hat{m} \end{aligned} \tag{2.1}$$

In Equation 2.1 $f(x)$ is the objective function to be minimized, representing the process to achieve the optimal design performance. A maximization problem can be solved by minimizing the inverse or negative of the function. The design domain is defined with a vector of design variables $\mathbf{x} = (x_1, x_2, ..., x_n)^T$ that are often bounded. The optimization search is normally subjected to several constraint functions including a vector of m equality constraints $c_i(x)$, and a vector of $\hat{m}$ inequality constraints $\hat{c}_j(x)$. These constraint functions ensure that the optimal design $\mathbf{x}^{opt}$ is in the feasible design domain. Though there are variations of the form in the optimization formulation, or there might be implicit objective functions that are not expressed in mathematical form, the general concept of design optimization is broadly applied in engineering design community to improve the quality and performance of engineering products.

In the marine structure design community, design optimization tool is very appealing for designers. Using various design requirements or rules as constraint functions, designers can utilize optimization process to pursue clearly definable objectives such

as minimized structural weight (*Sekulski*, 2009). Other criteria can also be stated as objectives such as cost effectiveness, durability, etc.

However, in the concurrent demanding and competitive environment, a single objective-oriented optimal design is not likely to be adequate for complex projects. Designers are constantly facing multiple criteria in decision making. To put multiple criteria into the single objective design optimization frame in Equation 2.1, one has to subjectively determine weighting factors for various criteria. Apart from weighted sum method, other classical formulations in transforming multi-objective problem into single-objective problem are ϵ-constraint method, and goal programming method, etc. An alternative approach in dealing with multiple criteria is for optimization tools to capture the trade-off between multiple priorities or objectives. Then decision makers are more aware of the design choices at hand. Multi-objective optimization frameworks have been developed to meet this need in multi-criteria decision making. A typical multi-objective optimization problem can be stated as:

$$\begin{aligned} minimize &: \mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x}), ..., f_k(\mathbf{x})] \\ with\ respect\ to &: \mathbf{x} \in \mathbb{R}^n \\ subject\ to &: c_i(x) = 0, i = 1, 2, ..., m \\ &\hat{c}_j(x) \geq 0, j = 1, 2, ..., \hat{m} \end{aligned} \tag{2.2}$$

In Equation 2.2 the objective is to minimize a vector of k functions that are all depend on the design variables $\mathbf{x}$, while satisfying all the equality constraint functions $c_i(x)$, and inequality constraint functions $\hat{c}_j(x)$. The multiple objective functions are often in conflict with each other, therefore, in multi-objective design optimization there is not a single design solution that is optimal in every objective function. The outcome of a multi-objective optimization is defined as a set of Pareto optimal solutions. The notion of Pareto optimality was first introduced by *Stadler* (1984) to the field of engineering and science. A Pareto optimal solution is a solution that cannot

be improved in any objective without degrading the performances of at least one of the other objectives. The entire set of the Pareto optimal solutions comprises the Pareto optimal front. For a more detailed definition of Pareto optimality readers can refer to the multi-objective optimization literature (*Miettinen*, 2012). An example of Pareto front from *Deb et al.* (2002) is shown in Figure 2.1.

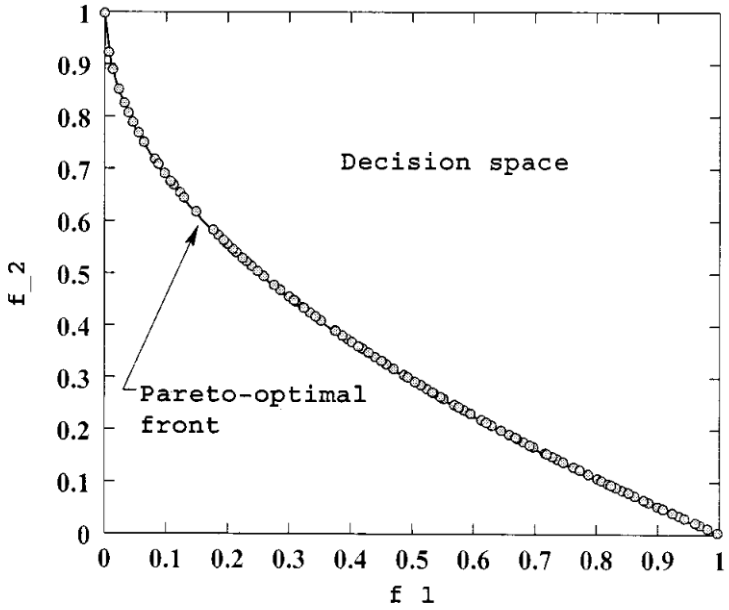


Figure 2.1: Example of a Pareto optimal front (*Deb et al.*, 2002).

In Figure 2.1, two objective functions are being minimized. The Pareto-optimal points in the front are called non-dominated points, in a sense that each point is only better than other points in one objective value. The Pareto-optimal front indicates the trade-off between different objective functions, and it shows the fundamental trade space of the problem to the designer. Therefore, in multi-objective optimization process has to accomplish two goals: search towards the Pareto-optimal front and find a diverse set of Pareto-optimal points.

Multi-objective design optimization allows for simultaneously optimization of all design objectives. The designer can choose a well-balanced decision from the Pareto optimal solutions. Once the designers have the higher level information regarding

the preference among those objectives, or compromises among different objectives have been made, it will be easy to arrive one single optimal design solution using the Pareto-optimal front information. The techniques to solve such optimization problems will be discussed in Section 2.3.

2.2.2 Design optimization considering uncertainty

The deterministic design optimization stated in Equation 2.1 and Equation 2.2 are effective in achieving design targets such as reduced structural weight or cost. However, in the optimization process, deterministic optimization reduces the design margin to the minimum level. Hence, the capability of the design solution to deal with design variability is significantly impaired. In other words, the deterministic optimal design solution can be very sensitive to design uncertainties. A small variation in design variable or environmental conditions may degrade the design performance and make the selected point sub-optimal. Hence, design optimization need to take uncertainty into consideration. This section briefly discuss the general concepts of two main philosophies in dealing with uncertainty in design optimization, namely Robust Design Optimization (RDO) and Reliability Based Design Optimization (RBDO).

The concept of Robust design is first proposed by *Taguchi* (1986) to improve product quality. Robust design optimization is to achieve a optimal design performance that is minimally sensitive to design uncertainties. A typical Robust design formulation can be stated as:

$$\begin{aligned} minimize &: f(x+\delta) \\ with\ respect\ to &: \mathbf{x} \in \mathbb{R}^n \\ subject\ to &: c_i(x) = 0, i = 1, 2, ..., m \\ &\hat{c}_j(x) \geq 0, j = 1, 2, ..., \hat{m} \end{aligned} \tag{2.3}$$

Robust design is focused on the outcome of the design performance when the de-

sign variables are perturbed from nominal condition. In the formulation stated in Equation 2.3, design variables $\mathbf{x}$ are considered as controllable, in a sense that they are chosen by the designer in the design optimization process; δ are non-controllable noise in the design vector, represent the variation in the design system. The noise information may not be exactly known to the designer but is often described by a probabilistic density distribution.

In robust design optimization, the mean and variance of objective function f are both desired design criteria. By establishing weighting parameters for these two criteria, the designer can formulate the robustness objective for design optimization. A comprehensive review of Robust design optimization formulations can be found in *Beyer and Sendhoff* (2007).

Reliability-based design treats uncertainty in a different perspective, it focuses more on the feasibility of design when faced with uncertainty. The optimal solution for traditional deterministic design usually lies against one or more of the constraint boundaries. Any variation may perturb the design into the infeasible domain and lead to design failure. Reliability-based design optimization (RBDO) (*Enevoldsen and Sørensen*, 1994; *Frangopol and Corotis*, 1996; *Tu et al.*, 1999) is aimed at improving the probability constraints will be satisfied with random design variables. RBDO introduces a concept of probabilistic constraints in an optimization framework. A general Reliability-based design optimization can be formulated as:

$$\begin{aligned} &minimize: f(\mathbf{d}) \\ &subject\ to: P(g_i(\mathbf{X}) \leq 0) \leq P_{t,i},\ i = 1, 2, ..., \hat{m} \\ &\mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U, \mathbf{d} \in \mathbb{R}^n \end{aligned} \tag{2.4}$$

In RBDO formulation design variable perturbed by uncertainty is defined as random variable in probabilistic form. $\mathbf{d}$ is the vector of mean values of random variables $\mathbf{X}$; $[\mathbf{d}^L, \mathbf{d}^U]$ is the design domain for the optimization problem. The probabilistic con-

straints are defined in the probability of constraint function violations ($g_i \leq 0$). $P_{t,i}$ is the maximum probability of failure for each of the reliability constraints.

The reliability of a system against failure is define as:

$$R = Pr\{g(X) \geq 0\} \tag{2.5}$$

where $X = [x_1, x_2, ..., x_n]$ represents the random variables such as loading, material strength, and geometry properties, etc. $g(X)$ is the limit state function defined as the boundary separating the two states of safety:

$$\begin{aligned} g(X) > 0 &: \text{the safe state} \\ g(X) = 0 &: \text{the limit state} \\ g(X) < 0 &: \text{the failure state} \end{aligned} \tag{2.6}$$

Once the limit state function has been defined, the next step is to evaluate the probability of failure, hence the reliability. The most direct way to compute P_f is by taking the integral of the joint probability density distribution of the random variables over the failure region:

$$P_f = Pr(g(X) \leq 0) = \int_{g(X) \leq 0} f_X(X) dX \tag{2.7}$$

However, the direct integral method is limited by the fact that it is difficult to obtain the joint probability density function, and the integration is not straightforward to apply. Analytical approximation methods such as the First Order Reliability Method (FORM) (*Hasofer and Lind*, 1974), Second Order Reliability Method (SORM) (*Breitung*, 1994), and numerical simulations such as Monte Carlo simulation (*Rubinstein and Kroese*, 2011) and directional sampling (*Ditlevsen and Bjerager*, 1989) are normally used to solve reliability problems.

The first-order reliability method is briefly introduced here as reliability calculation. More details on other methods can be found in reliability literature (*Melchers*, 1999; *Mahadevan and Haldar*, 2000). The FORM method is developed from the simple safety index concept. Based on the definition of limit state function above, the safety margin can be expresses as:

$$M = g(X) \tag{2.8}$$

The reliability index, (or safety index) is defined as the inverse of the coefficient of variation of the safety margin:

$$\beta = \frac{\mu_M}{\sigma_M} \tag{2.9}$$

This safety index method (*Cornell*, 1969) is a second-moment reliability method and is often used in some reliability calculations as well as code calibrations (*Madsen et al.*, 2006). However, the safety index method is dependent on the formulation of limit state equation, and the safety index may not reflect the true reliability if the random variables are not normally distributed. In addition, the safety index method may not solve the problem well if the limit state function is nonlinear. The FORM method (*Hasofer and Lind*, 1974) has been developed to address these shortcomings.

In the FORM procedure, the random variables X are first mapped into reduced standardized form of variables:

$$\begin{aligned} \phi(U_i) &= F_{X_i}(X_i), i = 1, 2, .., n \\ u_i &= \phi^{-1}(F_{x_i}(x_i)) \\ g(X) &= g(U) \end{aligned} \tag{2.10}$$

where F_{X_i} stands for the cumulative distribution function of X_i, and ϕ is the standard normal distribution function. Using the transform function, the reliability problem is

changed into the following term:

$$P_f = Pr(g(X) \leq 0) = \int\limits_{g(U)\leq 0} f_U(U)dU \tag{2.11}$$

The Hasofer and Lind's reliability index is defined as the shortest distance from the mean to the limit state surface in the reduced space, as illustrated in Figure 2.2. This point on the limit state surface is known as design point or most probable point (MPP), as it has the largest probability in the failure state space among other points at the limit state surface. The Hasofer and Linds reliability index can be solved by the following optimization problem:

$$\beta_{HL} = \min_{g(U)=0} \|U\| \tag{2.12}$$

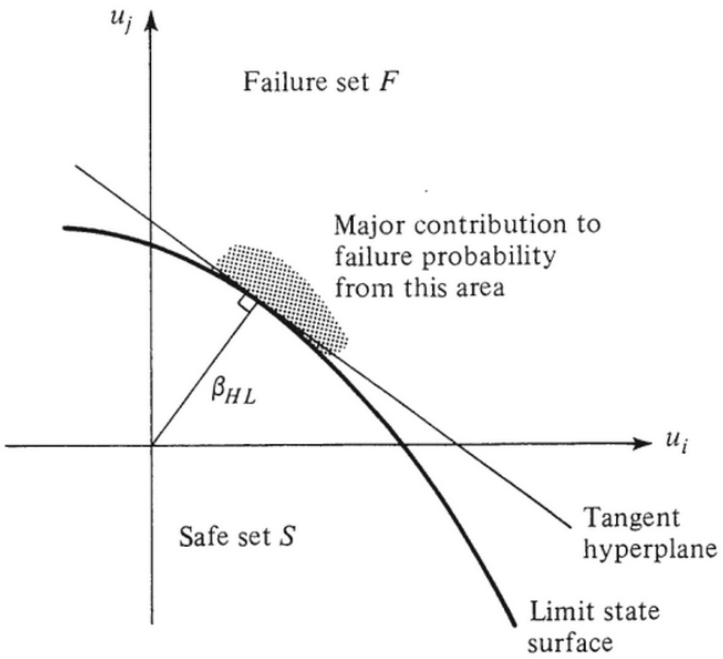


Figure 2.2: Definition of safety index (*Hasofer and Lind*, 1974).

The reliability index computed can be converted to the probability of failure in standard normal distribution function : $P_f = \Phi(-\beta)$. This probability of failure

P_f is then used in the RBDO framework expressed in Equation 2.4 to meet certain target probabilistic constraint: $P_f \leq P_t$. This RBDO formulation in Equation 2.4 is a typical double-loop approach, as the candidate design is evaluated in both reliability level and objective function level. Other formulations of RBDO such as sequential RBDO, single-loop RBDO can be found in *Du and Chen* (2004) and *Liang et al.* (2008). These alternative formulations are generally considered to be more efficient than the double-loop RBDO framework.

2.3 Optimization methods

For most optimization problems stated in Equation 2.1, if the objective functions and constraint functions are sufficiently smooth with regard to variables $\boldsymbol{x}$, gradient based methods are commonly used to solve the optimization problem. Gradient-based approaches can determine the most promising search direction for finding the function minimum, thus are more efficient in locating the optimal design point. The search direction is generally determined with gradient information, the more advanced methods will refine the search direction with second order information by approximating Hessian matrix.

In addressing many practical problems in engineering design, the assumption of gradient-based optimizers are changed. For example, in practical engineering problems the objective function or constraint function can be non-differentiable because of a discrete feasible design space. Besides, when there are multi-modal objective functions, gradient-based optimizer can be easily trapped into local minima and fail to converge to the globally optimal solution.

Gradient-free optimization methods arise to deal with these problems. The strength of gradient-free optimizer is that they do not rely on derivative information of objective function or constraint functions. This makes them suitable to address challenging problems that are difficult for gradient-based optimizer to solve. There are

various gradient-free optimizer developed based on different concepts, most of them are heuristic, stochastic, or population-based approaches. Prominent among them are Nelder-Mead Simplex (*Nelder and Mead*, 1965), simulated annealing (*Kirkpatrick et al.*, 1983), Particle Swarm Optimization (PSO) (*Eberhart et al.*, 1995), and evolutionary algorithm (*Goldberg*, 1989).

These methods each have their own advantage in solving optimization problems. The applicability of optimization algorithm depends on the type of the problem domain. In the marine structural domain, optimization algorithm choice is often adapted to structural problems with different focus (*Klanac et al.*, 2009; *Klanac and Jelovica*, 2009). This dissertation work mainly focuses on the evolutionary algorithms, they are employed as the global optimizer in addressing the marine structure design optimization problem. The rational behind this choice and the detailed description of evolutionary algorithm will be explained in the next section.

2.3.1 Evolutionary algorithm

Evolutionary algorithms (EA) are inspired by natural evolution. EAs are distinguished by the use of natural selection and a population of candidate designs to evolve optimal solutions. Evolutionary algorithms have been widely applied in recent years due to their performance in optimization. Evolutionary algorithms are different from traditional optimization methods in many aspects (*Goldberg*, 1989), EAs are adopted and applied for the research study in this dissertation for the following reasons:

1. As previously introduced, Evolutionary algorithms are gradient-free method, meaning that Evolutionary algorithms do not use gradient information in the search process. Thus, getting accurate derivatives of the fitness function is not necessary in EA optimizations. In the optimization design of marine structures, there can be non-continuous variables and complex function spaces where it is hard to get derivative information (*Temple*, 2015). Evolutionary algorithm

allows the designer to explore these challenging design spaces.

2. Evolutionary algorithms are population-based method. While traditional optimization methods use one solution at each iteration, evolutionary algorithms analyze a group of candidate solutions at one time. This strategy may look redundant in single-objective optimization, but it is advantageous in addressing multi-objective optimization problems. Because multi-objective optimization needs to find a set of Pareto optimal points as shown in the example in Figure 2.1, population-based optimizers are more powerful when searching for multiple optimal solutions simultaneously. When multiple design solutions are equally important for decision makers to evaluate trade-offs between different objectives, evolutionary algorithms provide an intuitive optimizer solution in finding these multiple optimal solutions. This dissertation is focused on discovering the impact of early stage uncertainty in design through multi-objective optimization, Evolutionary algorithms are thus well suited to achieve this task.

3. Evolutionary algorithms use stochastic operators in optimization searches. Traditional optimizers are sensitive to the choice of the starting point. In other words, they are easily trapped into local minimums. The stochastic operators give evolutionary algorithms a global functional perspective in locate the true optimal point. These evolutionary operators increase the probability in moving the population to the optimal point or Pareto optimal front. Evolutionary algorithms have proven to be very robust for engineering designs including marine structural optimizations in recent years.

Every evolutionary algorithm follows a basic procedure described in Figure 2.3. An Evolutionary algorithm is usually initialized with a random population of solutions. Then the algorithm starts the iteration process that evaluates and updates the current population to create a new population of solutions. The creation of new population

relies on three operators: selection, crossover, and mutation. The generation counter represents one iteration of the EA being completed.

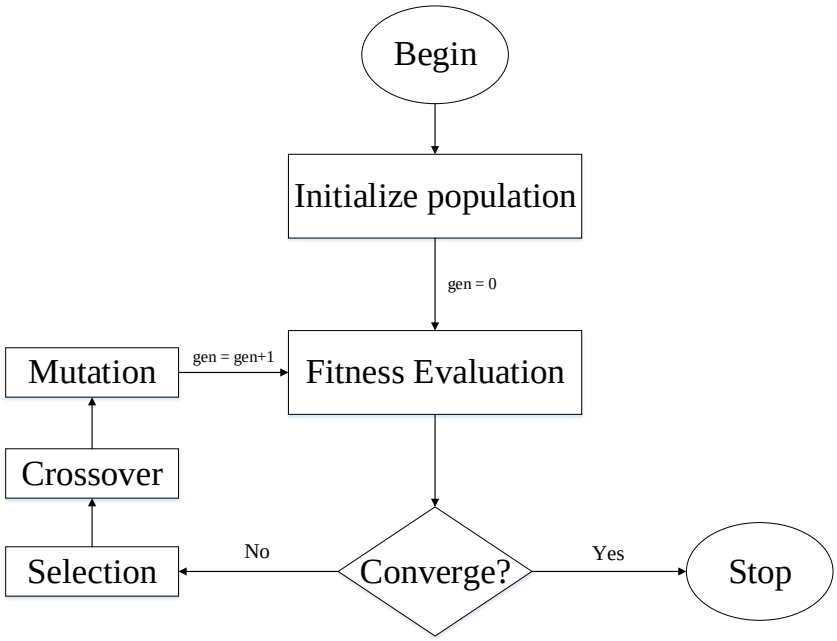


Figure 2.3: Typical evolutionary algorithm procedure.

The initialization randomly creates a population of individuals within the specified lower and upper bound on each design variable. Each individual stands for a possible solution for the optimization problem. It is represented by a chromosome in the EA scheme, where is chromosome is defined with genes for different variables, $\{x_1, x_2, ..., x_n\}$. The chromosome can be coded in either binary or real coding form. Every individual solution in the population is evaluated with objective and constraint function. Then a fitness value is assigned to this individual. For feasible solutions, the fitness values are normally the objective function values, while for infeasible solutions the fitness are objective function plus a penalty term that is proportional to constraint violations (*Deb*, 2001).

Three evolutionary operators are used to mimic the natural selection in the optimization procedure. The purpose of selection is to identify individuals for repro-

Table 2.1: Example of single-point crossover

	Parents		Offspring
Chromosome 1:	1 1 | 1 1 1	→	1 1 | 0 0 0
Chromosome 2:	0 0 | 0 0 0		0 0 | 1 1 1

duction that have good fitness values and eliminate individuals that have less fitness values, and keep the population size constant. There are many selection operators to achieve this goal. In the popularly used *tournament selection*, tournaments are played between two individuals and the better one is selected and placed into the mating pool. Selection emphasis the better solutions in the population and ensures non-decreasing of the overall fitness in the evolutionary algorithm process.

The crossover operator is used next for the mating pool. As the selection process is operating within the existing solutions, the evolutionary algorithm needs to explore other locations in the design domain. Crossover creates new individuals in a population. In a crossover operation, two individuals are randomly picked from the mating pool to create two new individuals. In the binary coded algorithms, commonly used crossover method is single-point crossover.

An example of single-point crossover is illustrated in Table 2.1. Single-point crossover is done by selecting a point in the binary coded variables for exchange. It can be seen that some portion of the Chromosome from two parent individuals are exchanged to formulate two offspring individuals. In this case the exchange is conducted in point 2. This idea can be extended to Multi-point crossover operations in binary coded chromosomes where information is exchanged in multiple positions of the parent chromosomes.

In real-parameter coded Evolutionary algorithm, crossover is directly implemented to real parameter values. An overview of many real-parameter crossover operators can be seen in *Herrera et al.* (1998). Here, the Simulated Binary Crossover (SBX) operator developed by *Deb and Agrawal* (1995) is introduced to illustrate the crossover

process in real-parameter chromosome.

SBX operator is proposed to simulate the power of single-point crossover operation in binary strings. In SBX two offspring $x_i^{(1,t+1)}$ and $x_i^{(2,t+1)}$ are created from parents $x_i^{(1,t)}$ and $x_i^{(2,t)}$ through Equation 2.13 below:

$$\begin{aligned} x_i^{(1,t+1)} &= 0.5[(1+\beta_{q_i})x_i^{(1,t)} + (1-\beta_{q_i})x_i^{(2,t)}] \\ x_i^{(2,t+1)} &= 0.5[(1-\beta_{q_i})x_i^{(1,t)} + (1+\beta_{q_i})x_i^{(2,t)}] \end{aligned} \tag{2.13}$$

β_{q_i} is the spread factor calculated as follows:

$$\beta_{q_i} = \begin{cases} (2u_i)^{\frac{1}{\eta_c+1}} & \text{if } u_i \leq 0.5; \\ \left(\frac{1}{2(1-u_i)}\right)^{\frac{1}{\eta_c+1}} & \text{otherwise.} \end{cases} \tag{2.14}$$

where u_i is a random number between 0 and 1; η_c is a non-negative distribution index given by user choice. A larger value of η_c gives a higher probability of creating 'near parent' offspring and a small value of η_c creates distant solutions in the offspring (*Deb*, 2001). SBX operates in a way that blends parents chromosome in generating offspring chromosome in a real coded Evolutionary Algorithm (EA).

In addition to the crossover operation, the mutation operator adds more search power to the EA. In the binary coded cases, the mutation operator is simple and straightforward. Mutation is normally conducted by selecting a random bit in the chromosome and changing a 1 to 0, and vice versa, with a mutation probability of p_m. The location of next mutation bit can be used with a mutation clock operator proposed in *Goldberg* (1989). In the real-parameter coded EA, the overview of good mutation operator can also be referred to *Herrera et al.* (1998). A polynomial mutation method (*Deb*, 2001) is introduced here for illustration.

In the polynomial mutation operation, the mutated new individual $y_i^{(1,t+1)}$ is obtained from the operation on individual $x_i^{(1,t+1)}$ expressed as follows:

$$y_i^{(1,t+1)} = x_i^{(1,t+1)} + (x_i^{(U)} - x_i^{(L)})\bar{\delta}_i \tag{2.15}$$

where $x_i^{(L)}$ and $x_i^{(U)}$ are lower bound and upper bound of design variable; the parameter $\bar{\delta}_i$ is computed from a polynomial probability distribution:

$$\bar{\delta}_i = \begin{cases} (2r_i)^{1/(\eta_m+1)} - 1 & \text{if } r_i < 0.5, \\ 1 - [2(1-r_i)]^{1/(\eta_m+1)} & \text{if } r_i \geq 0.5. \end{cases} \tag{2.16}$$

where r_i is a random number chosen in $[0, 1)$; η_m is mutation distribution parameter chosen by the user.

Crossover and mutation operations are processed with a probability in the mating pool. Crossover aims to combine two good solutions to form a better chromosome; mutation tries to alter a good solution for a better one. While both operations are not guaranteed to achieve the goal, the selection operator can make sure that if a bad solutions are created they will be eliminated, and if good solutions are created they will be kept. These three evolutionary operators constitute a powerful search for the fittest solution, when the evolutionary algorithm is terminated, the best solution in the current population is considered to be the optimal solution for the optimization problem.

2.3.2 Elitist non-dominated sorting genetic algorithm

The single-objective EA preserves elitism in a simple way: better solution among two individuals is selected for a slot in the next generation. In multi-objective optimization it is more complicated to identify and preserve elite solutions. An Elitist multi-objective evolutionary algorithm - NSGA-II proposed by *Deb et al.* (2002) is

briefly introduced to explain the implementation of an EA in multi-objective optimization problems.

In the multi-objective optimization, the concept of dominance is used in determining the better solutions. A solution $x^{(1)}$ is considered to dominate the solution $x^{(2)}$ is two conditions listed below are both true (*Deb*, 2001):

1. The solution $x^{(1)}$ is no worse than $x^{(2)}$ in all objectives;

2. The solution $x^{(1)}$ is strictly better than $x^{(2)}$ in at least one objective.

Using this criteria, any two solutions in the population can be categorized into one of the following categories:

- $x^{(1)}$ is dominated by $x^{(2)}$
- $x^{(1)}$ dominates $x^{(2)}$
- $x^{(1)}$ is non-dominated by $x^{(2)}$

Dominance provides a way to compare solutions in multi-objective optimization. With the dominance criteria the population of individuals can be classified into various non-dominated sets. There are usually multiple solutions in the same non-dominated set. When there is a need to identify the relative importance within the set of non-dominated solution, the crowding distance metric used in NSGA-II is introduced. Crowding distant sorting is mainly used to avoid achieving a group of solutions that are clustered together in a non-dominated set. The crowding distance metric measures perimeter of the cuboid formed with the nearest neighbors of a solution, as seen in Figure 2.4. The metric $d_{I_j^m}$ is computed by Equation 2.17 as follows:

$$d_{I_j^m} = d_{I_j^m} + \frac{f_m^{(I_{j+1}^m)} - f_m^{(I_{j-1}^m)}}{f_m^{max} - f_m^{min}} \tag{2.17}$$

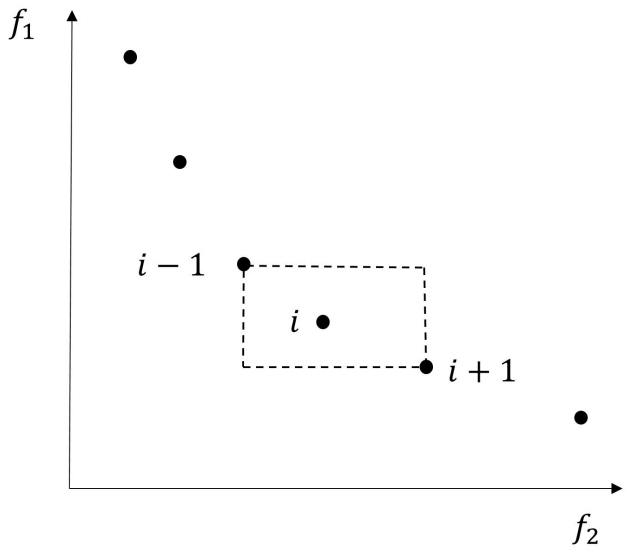


Figure 2.4: Example of crowding distance metric calculation

In Equation 2.17, I_j is the solution index of the j-th member; f_m^{max} and f_m^{min} are the maximum and minimum values of the m-th objective function; $f_m^{(I_{j+1}^m)}$ and $f_m^{(I_{j-1}^m)}$ are the objective function values. The endpoints of the front are assigned with a crowding distance value of ∞. Solutions that have large crowding distance metrics are favored when they are non-dominated with respect to each other. This help to preserve diversity in the population.

The selection process in evolutionary multi-objective optimization is different than that of in single-objective optimization ones. Take NSGA-II as an example, in the selection process illustrated in Figure 2.5, the offspring individuals Q_t is first created from parents P_t. These two population combined to form R_t of size $2N$. R_t is then sorted into various non-dominated sets $(F_1, F_2, ...)$. In formulate the new population for next generation of EA, the best non-dominated front is filled into P_{t+1} first, followed by the second best and so on. When the last allowed front is taken into account, the individuals are sorted with their crowding metrics, the better solutions are chosen to fill the space in P_{t+1}.

This selection process ensures that the fittest solutions are selected to survive in the next iteration of the EA. In addition, the diversity along the non-dominated front

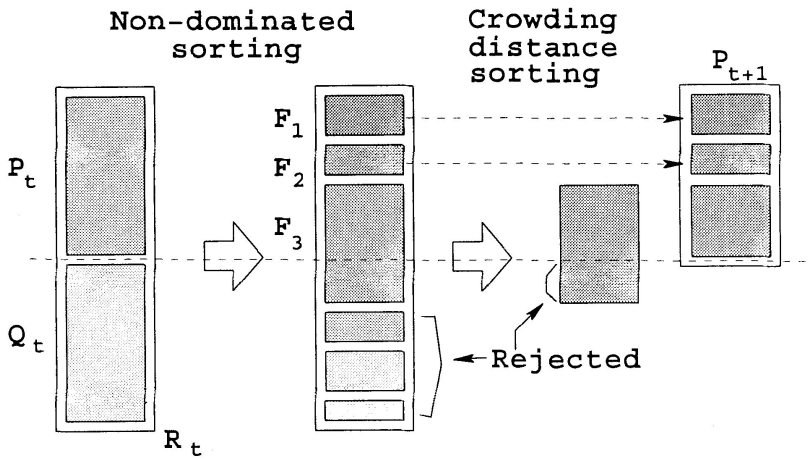


Figure 2.5: Schematic of NSGA-II procedure (*Deb et al.*, 2002)

is also preserved with the crowding metric in the selection process, ensuring that the optimal solutions is a good representation of the entire Pareto optimal front. NSGA-II is used in this dissertation work for multi-objective optimization studies on early stage design uncertainty.

2.4 Early stage design uncertainty: Epistemic uncertainty

Depending on the sources of uncertainty, uncertainty can generally be categorized into two major types: aleatory uncertainty and epistemic uncertainty. Aleatory uncertainty is due to the inherent randomness or variations associated with the system. The often mentioned stochastic uncertainty belongs to this type. Due to randomness, this uncertainty type is irreducible, meaning that aleatory uncertainty can not be removed from the system. In ship designs, the uncertainties in structural material properties and sea loads are considered as aleatory type uncertainties. Aleatory uncertainty is normally modeled with probabilistic distributions. Involving aleatory uncertainty in design optimization has been briefly discussed in Section 2.2.2 with robust design optimization and reliability-based design optimization.

To apply aleatory uncertainty, the probabilistic models and parameters must be specified. However, most of the time the specified models and parameters can not

adequately model the uncertainty without complete understanding of the system simulated. The discussion on the model assumptions give rise to the second type of uncertainty: Epistemic uncertainty. Epistemic uncertainty is the type of uncertainty that is due to a lack of knowledge in the system. Epistemic uncertainty describes a series of situations that are prevalent in engineering analysis where available information is incomplete, ambiguous, or available knowledge is subjective, vague, or deficient in some other way. There are two distinct differences between epistemic and aleatory uncertainty. First, epistemic uncertainty is a reducible type of uncertainty. With more knowledge and information, the epistemic uncertainty level can be reduced or removed completely. Second, epistemic uncertainty can not be described in probabilistic form, rather it must be treated in a non-probabilistic way. For a more detailed distinction between aleatory uncertainty and epistemic uncertainty refer to the works of *Helton and Burmaster* (1996), *Oberkampf et al.* (2004), and *Der Kiureghian and Ditlevsen* (2009).

This dissertation is focused on the epistemic uncertainty type in early stage structural design, optimization, and analysis. Early stage marine design is considered to be subject to lack of adequate information for key decision making. For example, the specifications of design parameters and design margins in the marine design problems has great impact in successful acquisition of the design, but they are made at the outset under incomplete knowledge of the design problem. Epistemic uncertainty suits well in this situation when there is a large design risk at stake for the lack of design knowledge. Thus, critical research need is to improve the capability to quantify the effect of epistemic uncertainty in early stage design. In addition, assuming the amount of epistemic uncertainty can be reduced through gathering relevant information, it is interesting to see how much design performance will improve from reducing epistemic uncertainty.

2.4.1 Epistemic uncertainty models

Three non-probabilistic uncertainty models are briefly introduced in this section as a survey of the overall theories in dealing with epistemic uncertainty. They are interval uncertainty, fuzzy sets theory, and evidence theory. They are all suited for designing robust structures under epistemic uncertainty.

2.4.1.1 Interval uncertainty

Interval uncertainty is a non-probabilistic uncertainty model. In interval uncertainty modeling, the uncertainty model is generally expressed in a closed form as shown below:

$$X = [x_l, x_u] = \{x \in \mathbb{R} | x_l \leq x \leq x_u\} \tag{2.18}$$

As seen in Equation 2.18, interval uncertainty model describes the uncertainty that ranges between crisp lower bound x_l and upper bound x_u, without additional information regarding variations or value frequencies, etc. In the interval uncertainty model only the interval is known, there is no assumption on the probability within this interval. Interval uncertainty is distinctly different from uniform probabilistic distributions in this regard. Interval uncertainty modeling is especially suitable for applications where there is a lack of information, or there is not enough knowledge to specify the parameters within interval.

The mathematical theory dealing with interval valued computation is shown in works of *Moore* (1966) and *Alefeld and Herzberger* (1983). Generally, interval uncertainty modeling analysis studies the mapping of interval input X_i to interval output Z_i:

$$\{X_1, ..., X_n\} \rightarrow \{Z_1, ..., Z_m\} \tag{2.19}$$

where the input values $x_i \in X_i$ and output values $z_j \in Z_j$ is linked through a deter-

ministic analysis model, f.

$$f : x \rightarrow z, x = (x_1,..., x_i, ..., x_n), z = (z_1, ..., z_j, ..., z_m)$$
$$x_i \in X_i, z_j \in Z_j \tag{2.20}$$

The important thing in interval analysis is to search for the bounds z_j^l and z_j^u of the interval output Z_j. In the mapping, z_j^l and z_j^u are associated with the optimal input $x_j^{l,opt}$ and $x_j^{u,opt}$. As the mapping model f may not be monotonic, locating $x_j^{l,opt}$ and $x_j^{u,opt}$ within X_i often involves an optimization process.

Interval uncertainty models are applicable in engineering design and analysis, and also provides additional insight into the design problems. For example, evaluating the widths of interval input and interval output with respect to each other, one can gain valuable insight into sensitivities and robustness of design (*Moens and Vandepitte*, 2007).

2.4.1.2 Fuzzy set

Fuzzy set theory provides an alternative in the boundary treatment for uncertainty models. Fuzzy logic has been used to model ambiguity and subjective knowledge to make decisions. Fuzzy logic is introduced in *Zadeh* (1965) to handle inherently imprecise phenomena. The detailed description of the theory can be referred to *Kosko* (1992) and *Mendel* (2001). Compared to the crisp bound in interval uncertainty model, boundaries in fuzzy theory are imprecise and linguistic. Fuzzy set defines a fuzzy membership values $\mu(x)$. A normalized fuzzy set is described as:

$$X = \{(x, \mu(x))|x \in \mathbb{R}, 0 \leq \mu(x) \leq 1\} \tag{2.21}$$

where $\mu(x)$ is the membership function, it describes the degree that a member belongs to a set between 0 and 1, this allows multiple sets for memberships. An example of

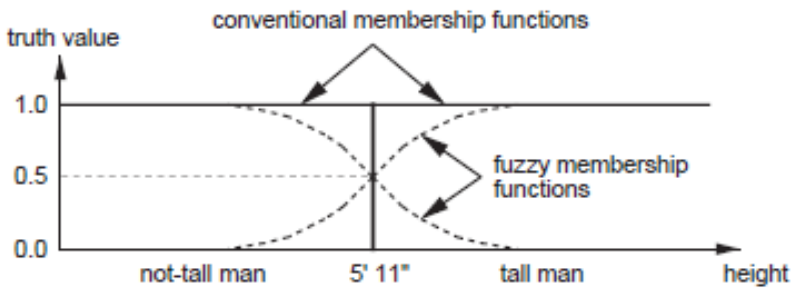


Figure 2.6: Example of Fuzzy logic (*Li et al.*, 1996)

using fuzzy logic from *Li et al.* (1996) is shown in Figure 2.6. The traditional sets theory and fuzzy sets are compared in this figure to describe male height. Traditional sets use crisp boundary of $5'11''$ in defining tall man, so any man above this height is designated as 'tall', and those below this height is 'not tall'. However, in fuzzy sets a man can be categorized into both 'not tall' and 'tall' sets at around $5'11''$.

It can be seen that this membership function is utilized to define uncertainty resulting from vagueness and ambiguity. In addition, fuzzy logic provides a way model uncertainty that can integrate experts judgments and subjective opinions through membership function arrangement. Fuzzy sets provide a good complementary tool for interval model when reference points or boundaries can not defined precisely.

Fuzzy logic has found wide applications in engineering design and analysis. Applications in the marine design domain can be seen in *Li et al.* (1996), *Parsons and Singer* (2000), *Gray et al.* (2010), and *Cuneo et al.* (2011).

2.4.1.3 Evidence theory

The Dempster-Shafer theory of evidence is first described by *Dempster* (1967) and further extended by *Shafer* (1976) into more general framework to model epistemic uncertainty. Evidence theory is a well-suited framework that can represent both epistemic uncertainty and aleatory uncertainty. It relaxes the assumption of probability when there is limited information, or when the information is ambiguous

and conflicting. Here, a very brief introduction of evidence theory is presented.

Differing from classical probability calculus, the probability mass function does not map $\mathbb{R} \to [0, 1]$. Instead the probability masses are assigned into sets therefore it is a mapping from $2^{\mathbb{R}} \to [0, 1]$.

In evidence theory, the user must assign a Basic Probability Assignment (BPA) to each interval, indicating the possibility that the value falls into that interval. The BPA is defined as:

$$\begin{aligned} m(B) \geq 0 \text{ for } B \subset \mathbb{R} \\ \sum_{B \subseteq \mathbb{R}} m(B) = 1 \end{aligned} \tag{2.22}$$

Then a focal element A is defined as:

$$\{A : A \subset \mathbb{R}, m(A) \geq 0\} \tag{2.23}$$

Most of the time, the focal elements are defined as intervals rather than complicated sets.

Evidence theory is based on two dual measures: Bel and Pl, which are belief and plausibility of an event. The belief and plausibility of an event is given by:

$$\begin{aligned} Bel(B) &= \sum_{A \subseteq B} m(A) \\ Pl(B) &= \sum_{A \cap B \neq \emptyset} m(A) \end{aligned} \tag{2.24}$$

The belief and plausibility have the following properties:

$$\begin{cases} Bel(B) \leq Pl(B) \\ Pl(B) = 1 - Bel(\bar{B}) \end{cases} \tag{2.25}$$

where $\bar{B}$ is the complementary hypothesis of B: $B \cap \hat{B} = \emptyset$. $Bel(B)$ and $Pl(B)$ can

also be seen as the bounds of probability $p(X \in B)$, where belief is considered as lower bound probability value, and plausibility, which is the "un-belief" value of the complementary hypothesis $\hat{B}$, can be seen as the upper bound probability value that are consistent with given evidence. In this way, the epistemic uncertain inputs can be represented by the interval values $[Bel(B), Pl(B)]$. This belief interval and the length of the belief interval gives a way to measure the epistemic uncertainty value.

Evidence theory provides a mixture of probability and non-probability in the treatment of uncertainty, and is considered reasonable when incorporating evidence into uncertainty analysis. Evidence theory is in its early stage of development in complex engineering designs. *Oberkampf and Helton* (2005) gave an overall discussion about the engineering applications of evidence theory. An example of applying evidence theory in design optimization can be seen in *Mourelatos and Zhou* (2006).

2.5 Surrogate modeling

An important research focus of this dissertation is on the aspect of numerical efficiency while conducting the early stage structural design and analysis. The need for efficient implementation of optimization is critical in handling the epistemic uncertainty in early stage design. Optimization and uncertainty analysis both require many structural function analysis calls. Thus, combing optimization with uncertainty analysis can easily become a computationally expensive problem. Considering the need for reduced design cycle time in early stage ship design, a computationally expensive framework will not be practical to use.

In this dissertation work, the solution to alleviate the computational burden problem is to use approximation models, known as surrogate models. The concept of surrogate model was originally proposed in "design and analysis of computational experiments" from *Sacks et al.* (1989), where a statistical approach was proposed to build a surrogate model to approximate an unknown, deterministic computational

model. Building a surrogate model of an expensive computational simulation is very promising in today's engineering designs. The development of the surrogate modeling in simulations can be seen in a review papers (*Simpson et al.*, 2001; *Kleijnen et al.*, 2005). Surrogate modeling shows special promise in the area of optimization (*Forrester and Keane*, 2009). The discussion of applying surrogate models in approximating fitness functions in evolutionary optimization algorithms can be seen in *Jin* (2005).

In this section, description of two main surrogate methods that this dissertation relies upon will be first presented. The Radial Basis Function (RBF) approximation and Kriging method are introduced in sequence. Afterwards, the construction method of a surrogate model within optimization is discussed.

2.5.1 RBF

The RBF method (*Buhmann*, 2000) is a very popular technique for use in interpolating function response. This technique is typically used to approximate functions or data only known at finite number of points (*Powell*, 1981).

With a set of training points $\{\mathbf{x_i}|i = 1, \cdots, N\}$ and corresponding functional responses $\{f(\mathbf{x_i})|i = 1, \cdots, N\}$, RBF method is aimed to train a function $\hat{y}$ that pass through all training points. RBF has the following form:

$$\hat{y}(\mathbf{x}) = \sum_{i=1}^{N} w_i \phi(\|\mathbf{x} - \mathbf{x_i}\|) \tag{2.26}$$

where ϕ_i are *radial basis function*, w_i represents the weights, and $\mathbf{x_i}$ are training points. From this formulation it can be seen that RBF approximation relies on the radial distance between input point $\mathbf{x}$ and training points $\mathbf{x_i}$. $\|.\|$ represents a norm to calculate the distances. Usually, a Euclidean norm is used. Each basis function mainly represents the response to a small region of the input space where the respective

training point is centered. Commonly used basis functions include inverse quadratic, linear splines, thin plate splines, gaussian, and multiquadrics (*Bishop*, 1995). They are listed in Table 2.2.

Table 2.2: Common basis functions in RBF

Basis function type	$\phi(r)$ $(r \geq 0)$
Gaussian	$e^{-(\epsilon r)^2}$
Inverse quadratic	$\frac{1}{1+(\epsilon r)^2}$
Inverse multi-quadric	$\frac{1}{\sqrt{1+(\epsilon r)^2}}$
Multi-quadric	$\sqrt{1+(\epsilon r)^2}$
Thin plate spline	$r^2 \ln r$
Linear	r
Cubic	r^3

In this table the notation r is defined as:

$$r = \|\mathbf{x_i} - \mathbf{x_j}\| \tag{2.27}$$

The scale parameter ϵ controls the influence domain of the basis function.

The RBF model is trained by solving the linear system of equations:

$$\mathbf{f} = \mathbf{\Phi w} \tag{2.28}$$

where $\mathbf{f} = [f(\mathbf{x_1}), ..., f(\mathbf{x_n})]$. The $N \times N$ matrix $\mathbf{\Phi}$ is denoted as the interpolation matrix:

$$\mathbf{\Phi} = \begin{pmatrix} \phi(\mathbf{x_1}, \mathbf{x_1}) & \phi(\mathbf{x_1}, \mathbf{x_2}) & \cdots & \phi(\mathbf{x_1}, \mathbf{x_n}) \\ \phi(\mathbf{x_2}, \mathbf{x_1}) & \phi(\mathbf{x_2}, \mathbf{x_2}) & \cdots & \phi(\mathbf{x_2}, \mathbf{x_n}) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(\mathbf{x_n}, \mathbf{x_1}) & \phi(\mathbf{x_n}, \mathbf{x_2}) & \cdots & \phi(\mathbf{x_n}, \mathbf{x_n}) \end{pmatrix} \tag{2.29}$$

The weights $\mathbf{w}$ can be calculated in Equation 2.30 or using least square estimation.

$$\mathbf{w} = \mathbf{\Phi}^{-1}\mathbf{f} \tag{2.30}$$

After calculating the weight $\mathbf{w}$, the prediction of function response at any given input vector is straightforward using Equation 2.26.

RBF can be regarded as a simple type of artificial neural networks, where the radial basis functions act as the activation function in the network.

2.5.2 Kriging

Kriging is a powerful surrogate model that has been widely used to approximate computationally expensive simulations. An in-depth Kriging theory can be found in works of *Sacks et al.* (1989) and *Simpson et al.* (2001), a more recent review can be seen in *Kleijnen* (2009). As a brief explanation, Kriging predicts the function value $\hat{y}$ at an unobserved point based on a set of sampling points through a realization of a regression model and a stochastic process:

$$\hat{y}(\boldsymbol{x}) = \boldsymbol{f}^T\boldsymbol{\beta} + z(\boldsymbol{x}) \tag{2.31}$$

where $\boldsymbol{f}$ is regression basis function selected by the user and $\boldsymbol{\beta}$ are regressional coefficients. In the ordinary Kriging, $\boldsymbol{f}$ is a vector of all 1.0 with length n_s, while in universal Kriging regression models like linear and quadratic models are used. However, ordinary Kriging is sufficient in most practices (*Sacks et al.*, 1989).

The stochastic process z is assumed to have zero means and a covariance of:

$$E[z(x_i)z(x_j)] = \sigma^2 R(\theta, x_i, x_j) \tag{2.32}$$

where σ is the process variance and $R(\theta, x_i, x_j)$ is the correlation model. R normally has the form:

$$R(\boldsymbol{\theta}, \boldsymbol{x_i}, \boldsymbol{x_j}) = \prod_{k=1}^{n} R_k(\theta_k, x_k^i, x_k^j) \tag{2.33}$$

for stationary, one-dimensional correlations. Commonly used correlation functions

are listed in Table 2.3, where d_k denotes $|x_k^i - x_k^j|$.

Table 2.3: Correlation functions for Kriging

Correlation form	$R_k(\theta, d_k)$
Linear	$max\{0, 1 - \theta_k d_k\}$
Exponent	$exp(-\theta_k d_k)$
Gaussian	$exp(-\theta_k d_k^2)$
Cubic	$1 - 3\xi_k^2 + 2\xi_k^3,\ \xi_k = min\{1, \xi_k d_k\}$
Spherical	$1 - 1.5\xi_k + 0.5\xi_k^3,\ \xi_k = min\{1, \xi_k d_k\}$

Assuming a Gaussian correlation model is adopted, the correlation function is expressed as:

$$R(\boldsymbol{\theta}, \boldsymbol{x_i}, \boldsymbol{x_j}) = exp(-\sum_{k=1}^{n_v} \theta_k(|x_k^i - x_k^j|^2) \tag{2.34}$$

where $\boldsymbol{\theta}$ is a correlation parameter vector that is found by optimizing a maximum likelihood function. After that, the Kriging predictor for an unobserved point $\boldsymbol{x}$ can be expressed as the following:

$$\hat{y}(\boldsymbol{x}) = \boldsymbol{f}^T\boldsymbol{\beta} + \boldsymbol{r}^T(\boldsymbol{x})\boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{F}\boldsymbol{\beta}) \tag{2.35}$$

where $\boldsymbol{\beta}$ is computed by least square regression, the vector $\boldsymbol{r}$ measures the correlation between the prediction point $\boldsymbol{x}$ and the sampled points $[\boldsymbol{x_1}...\boldsymbol{x_m}]$.

$$\boldsymbol{r}(\boldsymbol{x}) = [R(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{x_1}), R(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{x_2}), ..., R(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{x_m})]^T \tag{2.36}$$

$\boldsymbol{F}$ is function responses at sampled points:

$$\boldsymbol{F} = [f(x_1), \cdots, f(x_m)]^T \tag{2.37}$$

The least squared estimated $\boldsymbol{\beta}$ and estimated variance σ^2 are computed as:

$$\boldsymbol{\beta} = \left(\boldsymbol{F}^T\boldsymbol{R}^{-1}\boldsymbol{F}\right)^{-1}\boldsymbol{F}^T\boldsymbol{R}^{-1}\boldsymbol{Y} \tag{2.38}$$

$$\sigma^2 = \frac{1}{m}(\boldsymbol{Y} - \boldsymbol{F}\boldsymbol{\beta})^T \boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{F}\boldsymbol{\beta}) \tag{2.39}$$

The mean square error s^2 of the predictor can also be provided:

$$s^2(\boldsymbol{x}) = \sigma^2 \left(1 + \boldsymbol{u}^T \left(\boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{F}\right)^{-1} \boldsymbol{u} - \boldsymbol{r}^T \boldsymbol{R}^{-1} \boldsymbol{r}\right) \tag{2.40}$$

where $\boldsymbol{u} = \boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{r} - \boldsymbol{f}$. One advantage of Kriging is that the estimate of prediction variance s^2 can be derived without much additional computational effort. This term s^2 give a confidence interval of prediction along with the predicted function response.

2.5.3 Sampling for surrogate modeling

Fitting a surrogate model requires sampled data points from the simulation models. Thus sampling plans are needed to generate input training data for surrogate modeling. Classical sampling methods include fractional factorial design and central composite design (*Myers et al.*, 2016) that spread sample points around boundaries and center of a design space. An improved type of sampling methodology is space filling sampling. Typical methods that have space filling properties are orthogonal arrays (*Hedayat et al.*, 1999), uniform designs (*Fang et al.*, 2000), and the Latin Hypercube Sampling (LHS) (*McKay et al.*, 1979) method. This section introduces the Latin Hypercube Sampling as an illustration. For further discussion on selecting input sampling points, readers can be referred to *Müller* (2007).

2.5.3.1 Latin hypercube sampling

In practice, LHS methods are often used to generate sampling points for fitting surrogate models. The main steps of latin hypercube sampling strategy is summarized in the following:

1. Dividing the interval of each dimension of the design domain into N non-overlapping intervals based on an assigned distribution;

2. Sample randomly a point in each dimension;

3. Pair randomly the point from each dimension.

Assuming that the variables of $\mathbf{X}$ are independent, the cumulative distribution function of X_k is denoted as F_k. Let X_{jk} stands for the kth component of $\mathbf{X}_j$, the jth sampled value. Supposed a sample size of N points are required for simulation, a $N \times K$ matrix $P = \{p_{jk}\}$ is built where each column of P is an random permutation of $\{1, 2, ..., N\}$. Afterwards use ξ_{jk} $(j = 1, ..., N; k = 1, .., K)$ as uniformly distributed random variables on [0,1], then the sampling plan X_{jk} is given as:

$$X_{jk} = F_k^{-1}((p_{jk} - 1 + \xi_{jk})/N) \tag{2.41}$$

An example of a latin hypercube sample of uniformly distributed $\boldsymbol{X}$ is shown in Figure 2.7. It can be seen that the metric p_{jk} decides which cell X_j is located, and ξ_{jk} determine the location in cell of X_j.

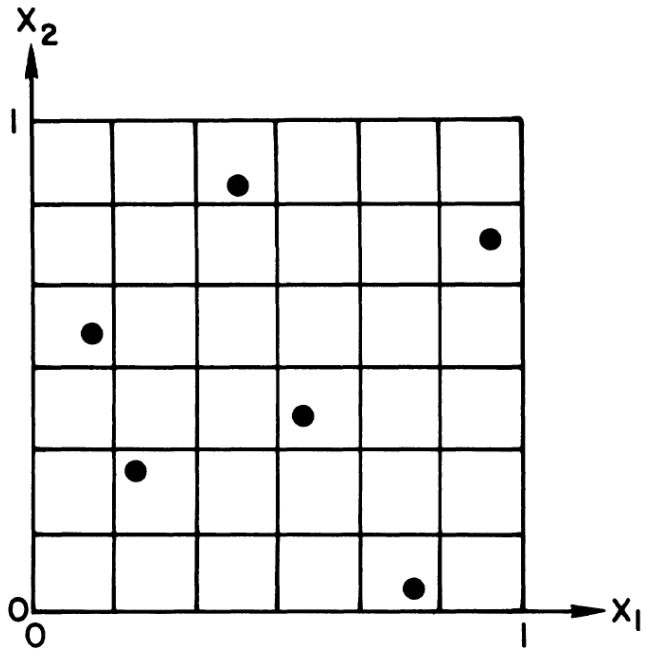


Figure 2.7: A Latin Hypercube Sample with N=6, K=2 (*Stein*, 1987)

Latin hypercube sampling method ensures that all portions of the design space are sampled. The good space-filling properties of LHS generates a diverse set of design points in fitting a surrogate model. Given a fixed sample size, LHS selects sampling points that are representative to help explore the function relationship between input and output.

2.5.3.2 Off-line sampling method

If surrogate modeling is used in optimization, the model fitting process can be classified into two types: the off-line method and the on-line method (*Jin*, 2005). The off-line construction of surrogate model can be generalized in Figure 2.8. Using the introduced sampling methods, a number of sampled points are generated and simulated with objective functions. Then, the surrogate model is trained with the sample points to approximate the input-output functional relation. After the optimizer is initialized, the trained surrogate model is used to estimate the fitness values for the optimization run.

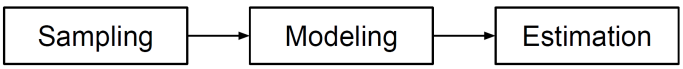


Figure 2.8: General off-line surrogate construction

The off-line sampling method fixes the number of sample points, and hence the computational cost of building the surrogate model. Off-line methods are easy to implement, but the surrogate model's quality is fixed in the sampling stage. Thus, in order to ensure the accuracy of surrogate predictions, generally large number of sampling points are needed to fill all the regions of the design space.

2.5.3.3 On-line sampling method

In contrast with the off-line method, the on-line method samples and refines the surrogate model during the optimization process. The on-line sampling strategy in-

volves selecting data as more is known about the problem for a more efficient sampling in the design space. A general process of on-line sampling is illustrated in Figure 2.9. In the on-line strategy, the correctness of the surrogate model is utilized to refine the surrogate, thus the prediction ability and quality of surrogate can be improved.

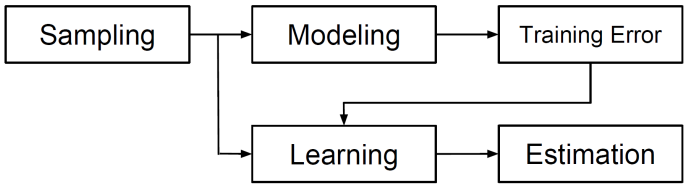


Figure 2.9: General on-line surrogate construction

On-line strategy normally involves model validation of the surrogate models. The accuracy of surrogate model can be validated via comparing the true fitness function values in the validation dataset. In Gaussian process based surrogate models, the estimated prediction errors derived in Equation 2.40 can also be used to indicate the correctness of the surrogate model. The learning and refinement are important as the initial surrogate model can be coarse and may not provide accurate prediction in every domain in the design space. Therefore, on-line strategy must be formulated to increase the accuracy of the surrogate model as the optimization runs. However, such refinement can cause a tractability problem. For example, repeatedly building Kriging construction can be very costly as discussed in *Gano et al.* (2006)

2.6 Summary

This chapter reviews some basic concepts and theories that are applied in the dissertation. As discussed in Chapter I, the main focus of the research is to explore the impact of the early stage design uncertainty on structural designs. Though design optimization in uncertain environments has undergone enormous development, there is little research study in systematically investigating marine structural design

optimization in with early stage design uncertainty. This section briefly discuss the research needs in achieving this goal.

2.6.1 Design optimization considering epistemic uncertainty

Compared to design optimization with aleatory uncertainty, less has been studied on design optimization involving epistemic uncertainty. Thus it is desirable to discover the impact of epistemic uncertainty on design decision making and performance. To this end, this dissertation aims to build an optimization framework that can show different feasible configuration designs associated with different epistemic uncertainty levels. In order to quantify the various uncertainty levels, an interval uncertainty model is chosen for treatment of epistemic uncertainty. The initial study of structural design with interval uncertainty is presented in Chapter III. After investigation of advanced surrogate modeling techniques, the overall optimization framework in dealing with various ranges of interval uncertainty is presented in Chapter V, where the trade-off of various range of uncertain intervals and corresponding feasible structural configurations is shown.

2.6.2 Surrogate-assisted evolutionary optimization

The desired study on design optimization with interval uncertainty is very demanding in terms of computational resources. The problem becomes more critical given that evolutionary algorithms are used for trade-off analysis. This desired research goal puts special challenges on computational efficiency, requiring further surrogate modeling development.

Surrogate modeling techniques have to be adapted to structural design optimization in several areas. First, structural simulations often have high-fidelity and low-fidelity numerical models. High-fidelity models are generally more accurate, but more demanding in computational effort, whereas low-fidelity models are simplified meth-

ods that are less accurate but are easier to evaluate. In order to utilize limited high-fidelity simulations in early stage structural design, a surrogate modeling management strategy that can interact with high-fidelity and low-fidelity models are needed.

Second, surrogate modeling for large problems size is also currently a challenge. A single surrogate model is not likely to be capable of handling a large design domain and complex functional space required for structural analysis. The problem is further complicated by the presence of integer or discrete decision variables in structural designs. A multiple surrogate modeling strategy is needed in this situation to ensure the prediction accuracy of surrogate. Multiple surrogate modeling is a research direction that is promising in reducing the risk of over fitting surrogate model and is explored in this dissertation work.

Last, dealing with interval uncertainty in design also needs an advanced surrogate modeling technique. Introducing interval analysis in evolutionary computation brings additional computational costs. This motivates the traditional surrogate modeling method to adapt to more efficient prediction of interval analysis results. Constructing surrogate models for interval analysis in the context of on-line sampling method is an open research area. This dissertation work presents a unique solution in addressing this issue.

In Chapter III, the off-line surrogate modeling is presented to initially study interval uncertainty in marine structural design. This initial study is followed with the research development addressing the mentioned challenges with on-line surrogate modeling. Chapter IV presents a multiple surrogate model management strategy for variable fidelity structural design optimization. Surrogate methods for worst-case performance estimation is proposed in Chapter V to efficiently investigate interval uncertainty in design optimization.

CHAPTER III

Robust and reliable design optimization considering interval uncertainty

3.1 Introduction

Engineering designs are invariably confronted with design uncertainties. The variability of uncertainty in design variables and parameters can affect engineering design performance. In the worst cases, the physical prototype of the design can be infeasible and failed due to the effect of uncertainty. The ship design process is particularly sensitive to the design uncertainties. Ship design faces various sources of uncertainty that are inherent from the complex nature of ship design process. In ship structural design, some key parameters are determined based on subjective experience in preliminary design stage, while not guaranteeing that the design is immune to rework later in the design stage. The general ship design process can be described in the spiral design diagram (*Evans*, 1959) illustrated in Figure 3.1. The structural decisions directly influence the weight estimation, and then the arrangement of compartment such as engine and auxiliary configurations. There is a need to examine the structural designs responding to this type of uncertainties in the early stage. This chapter presents design optimization studies that are intended to remove the epistemic uncertainty early in marine structural design.

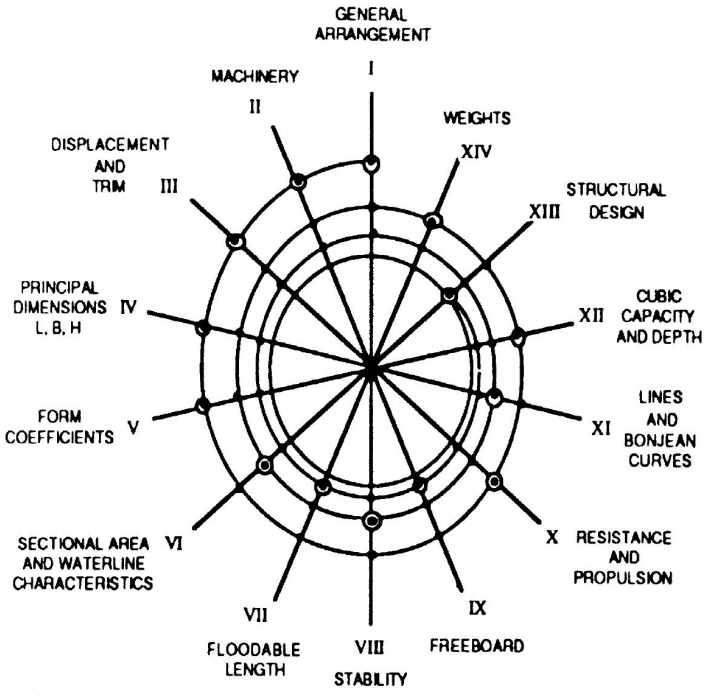


Figure 3.1: Evans's general design diagram (*Evans*, 1959).

As discussed in Chapter II, design optimization under uncertainty has various forms with regards to different types of uncertainties. While stochastic uncertainty is the common form in dealing with uncertainty, this dissertation mainly focus on another form of uncertainty. The uncertainty that is of interest here is defined as epistemic uncertainty, or uncertainty due to lack of knowledge (*Oberkampf et al.*, 2002). Aleatory uncertainty, or stochastic uncertainty, is appropriate to describe uncertainty in material properties or sea environment in ship design. Meanwhile, other uncertainties are more epistemic in nature, such as model uncertainties and configuration uncertainties, and thus require epistemic uncertainty models to study them. Here, the interval uncertainty measure is examined, representing the epistemic uncertainty in structural design. In the context of the ship design process, what is needed in the early stage is a robust baseline structural design that is least affected by design changes arising from other disciplines while the design evolves. Additionally, the designer would be interested to understand the decision making under various interval

uncertainties in different disciplines in the ship design. Thus an overall perspective concerning the epistemic uncertainty can be gained in early stage.

The section will initially demonstrate the surrogate-assisted optimization formulation in addressing the interval uncertainty in structural design. First, a surrogate-assisted robust design optimization is presented to show the difference between robust optimal design compared to deterministic optimal design. It follows with a further study on interval uncertainty in both configuration uncertainty and model uncertainty. A multi-objective optimization is presented to show the trade-off between interval uncertainty width and design performance. A surrogate method is implemented for efficient optimization in both interval uncertainty studies. The content presented in this chapter was first published in *Liu and Collette* (2014b) and *Liu and Collette* (2015). The text that follows is an extended version of that manuscript.

3.2 Surrogate-assisted robust design optimization framework

3.2.1 Evolutionary optimization with interval uncertainty

Evolutionary optimization has long been proposed to address design uncertainty problems, a detailed overview can be referred to *Jin and Branke* (2005). This section discusses the general optimization problem considering interval uncertainty, and the algorithmic scheme in an evolutionary optimization frame.

Considering interval uncertainty in the design, the optimization problem can be formulated as follows:

$$\begin{aligned} &\text{minimize} && f(x+\varphi, c+\delta) \\ &\text{w.r.t.} && x \in R \\ &\text{subject to} && g(x+\varphi, c+\delta) \geq 0 \end{aligned} \tag{3.1}$$

where x and c are vectors of design variables and design parameters respectively. φ

and δ are vector of deviations in the design variable and design parameters respectively, they are both treated with interval uncertainty terms:

$$\begin{aligned} \varphi_{Li} \leq \varphi_i \leq \varphi_{Ui} \\ \delta_{Li} \leq \delta_i \leq \delta_{Ui} \end{aligned} \tag{3.2}$$

Achieving the robust optimal solution with consideration of interval uncertainty normally requires additional effort than traditional optimization. The robust solution needs to perform well even if the design variables or parameters are perturbed by interval uncertainties. Therefore, every design solution needs to be examined by its robustness in presence of uncertainty as part of the fitness evaluation in the optimizer. Depending on the different definitions of robustness, the evaluation in face of uncertainty varies in different problem formulations. Here the worst-case performance is used for design against interval uncertainty. As the interval uncertainty model involves no probabilistic distribution or value density, it is necessary to account for the full range of possible values within the interval.

The basic principle for the optimization problem in Equation 3.1 is to search for the design solution that has the best worst-case performance. Worst-case performance stands for the minimum of a maximization optimization problem, or the maximum in a minimization problem. Thus the design optimization with interval uncertainty becomes a double-loop search process, where the inner loop searches for the worst-case performance in the interval uncertainty domain. The outer loop searches for the solution that has the best worst-case performance. This approach is often referred to as Maxmin optimization (*Ong et al.*, 2006).

Applying the robust scheme in an evolutionary optimization algorithm has been discussed in many precious works such as *Tsutsui and Ghosh* (1997); *Tsutsui* (1999); *Arnold and Beyer* (2002). In this design optimization with interval uncertainty, the robust scheme is defined with the worst-case performance. The general search scheme

is outlined in Algorithm 1.

For each individual x_i evaluated in the evolutionary optimizer, the worse-case design x_l needs to be found within the interval uncertainty domain. Then the fitness value of this individual is set with the worst-case function performance $f(x_l)$ rather then the function value $f(x_i)$ at the nominal design point. In this setting of evolutionary optimization scheme, the individuals that has better worst-case performance in the presence of uncertainty are more likely to survive in the EA operations.

Algorithm 1 Pseudo code for double loop search within evolutionary optimizer.

```
BEGIN
Create a population of candidate designs.
while Termination criteria is not met do
    for Each individual x_i in population do
        Locate the point x_l with the worst case fitness value within interval bounds;
        Set f(x_i) = f(x_l), i.e., the fitness of individual is set to the worst case value;
    end for
    Apply EA operators to create the next generation.
end while
END
```

3.2.2 Surrogate-assisted robust design optimization

Evolutionary algorithms such as genetic algorithms are very robust in solving optimization problems, particularly they are suited to deal with marine structure design problems when there is a mixed continuous and discrete variables. However, EA have been criticized for its large demand of fitness function evaluations. With the worst-case performance search scheme added in the evolutionary optimization framework, the number of the objective functions calls will grow even larger. This section presents a computationally efficient surrogate model to reduce the computational burden incurred by the double-loop approach.

Coupling surrogate models into evolutionary algorithms has been addressed in many previous works, a discussion of this overall methodology can be referred to some review literatures (*Jin*, 2005, 2011). Surrogate models are also widely used for

robust design optimization in evolutionary algorithms. An example can be seen in *Ray and Smith* (2006), where Neural networks are implemented to reduce computational efforts in robust design. *Ong et al.* (2006) have proposed Radial Basis Function surrogate method to address the previously mentioned double-loop maxmin optimization problem. The presented surrogate method is adapted from that in *Ong et al.* (2006), with a different surrogate construction and management strategy in place of the trust region framework used by *Ong et al.* (2006).

The basic procedure of the surrogate-assisted optimization is outlined in the flow chart in Figure 3.2. A Single Objective Genetic Algorithm (SOGA) is used as the global optimizer in searching for the robust solution. RBF (*Buhmann*, 2000) is employed for constructing surrogate models. The theoretical introduction of RBF can be found in the previous chapter.

RBF has the advantage of fast training and only requires small sample size for a realistic model. However, RBF is based on the continuous assumption, and it may not work well with discrete variables. The strategy applied here is to separate discrete variables when building the RBF surrogate model. The sampling points observed for surrogate modeling are first divided into separate clusters based on different combination of discrete variable value set.

The training and use of RBF surrogates in Figure 3.2 is on the basis of individual evaluation. The individual to be evaluated consists of a continuous part x^C and a discrete part x^D. The cluster that is labeled with the same discrete part x^D is identified and chosen for surrogate training. Afterwards, the Euclidean distances of the continuous part x^C to each member of the chosen cluster are computed, and a certain number of the closest points are sorted out to build a local RBF surrogate model. In the presented work, 8,000 points are sampled, while local RBF surrogate model has a size of 100 data points.

When a local RBF surrogate model $\hat{f}$ is constructed, it can then be used for

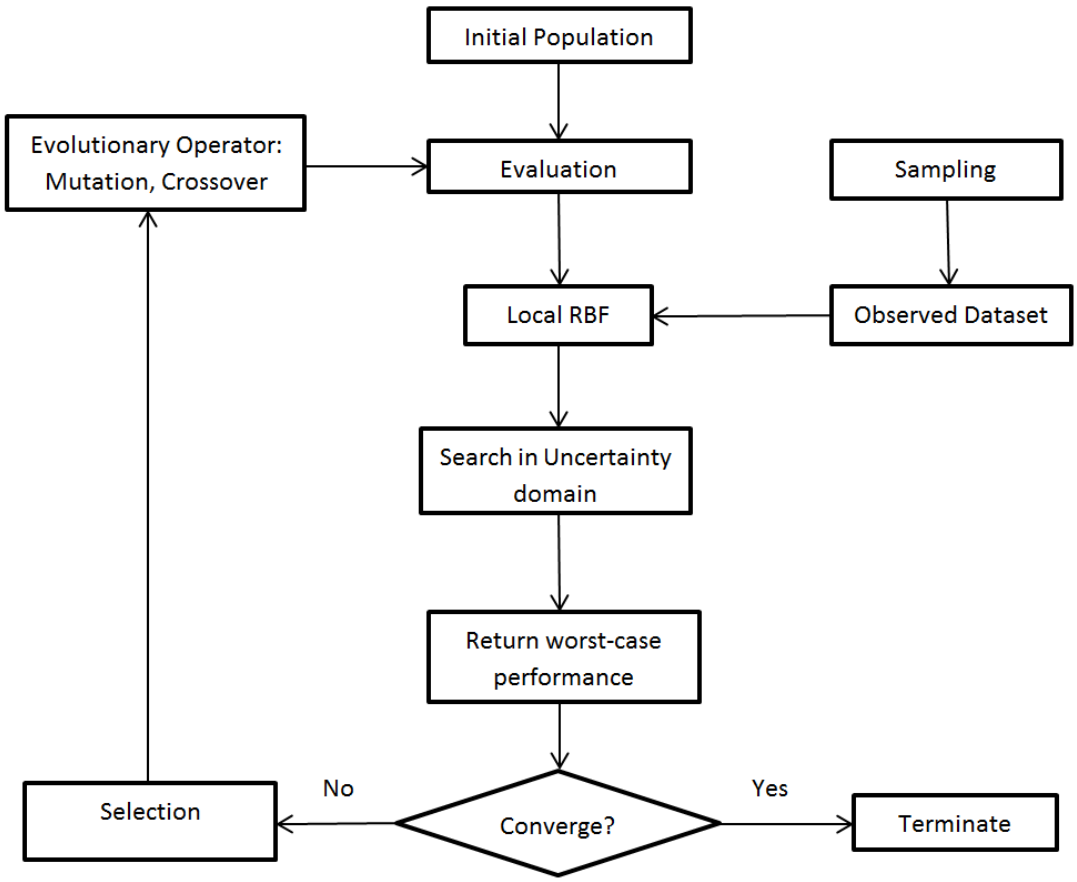


Figure 3.2: Flow chart of surrogate assisted robust optimization.

computing the worst-case performance for this individual. In the inner loop of the double-loop optimization approach, the following subproblem, expressed in Equation 3.3, is solved to get the worst-case performance from the RBF surrogate model, assuming the outer loop is a minimization problem.

$$\begin{aligned} &\text{Maximize} && \hat{f}(x+\varphi, e+\delta) \\ &\text{w.r.t.} && \varphi, \delta \\ &\text{Subject to} && \|\varphi - \varphi_0\| \leq \Delta_1 \\ & && \|\delta - \delta_0\| \leq \Delta_2 \end{aligned} \tag{3.3}$$

where $\hat{f}$ is the local RBF constructed, and Δ_1, Δ_2 are the bounded interval uncertainty domains. φ_0, δ_0 are nominal values of the uncertainty, and are set as zero in this case. In the inner loop optimization search, the Nelder-Mead optimization method is employed to search for the worst-case uncertainty perturbations $[\varphi^*, \delta^*]$. They are then added to the individual candidate design, and the worst-case functional performance is set as the fitness value for this individual. This is sent to the evolutionary algorithm.

3.2.3 Robust compartmentalization case study

As the early stage structure decisions have a large impact on later stage design, this section attempts to establish a robust baseline structural weight estimation while the ship subdivision and hence the structural support arrangements within compartments are evolving. Thus the desired robust structural design can withstand certain design changes without disrupting the entire design scheme. This motivates a robust design optimization to find a compartment's structural configuration that keeps the design feasible over a range of potential compartment lengths.

In the illustration of the structural design example, a simple box girder structure is adopted as it is broadly representative of a single-hull ship hull girder. This box

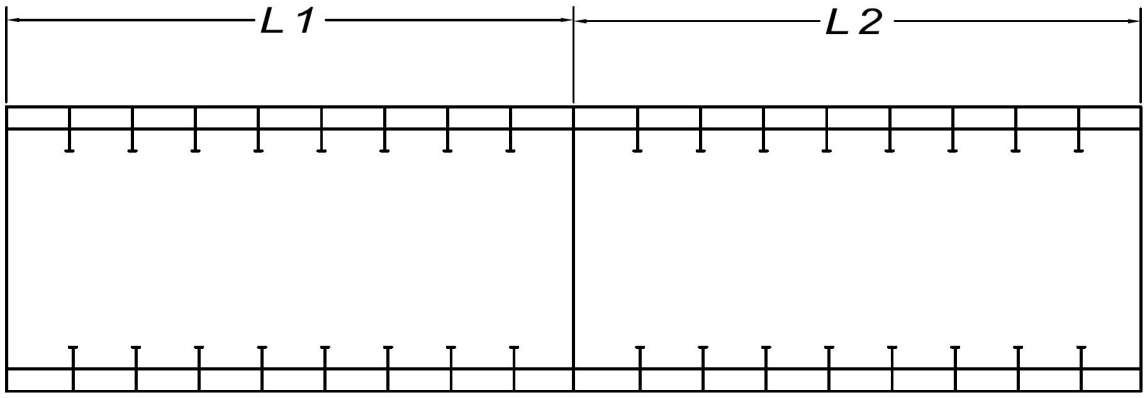


Figure 3.3: Sketch of two compartments.

Table 3.1: Standard T-stiffener dimensions

Description	Value
t_w: Web thickness	6.6 mm
h_w: Web height	313.0 mm
t_f: Flange thickness	11.0 mm
b_f: Flange breadth	166.0 mm

structure is composed of two compartments with a fixed total length L of 16 m. The depth, D, is fixed at 4 m; the breadth, B, at 4 m for simplicity. All four sides of the box girder have the identical structures. The sketch of the structure is shown in Figure 3.3. The length of the first compartment, L_1, has a nominal value of 8 m. This design parameter is treated with interval uncertainty form bounded by $[-2, 2]$. As the total length of the two compartments is fixed at 16 m, the length of the second compartment L_2 also varies within a 4 m interval uncertainty band. The schematic cross section of the structure is shown in Figure 3.4.

In the box structure, the longitudinal stiffener configuration is chosen from the library of Navy steel t-stiffeners, its dimensions are listed in Table 3.1. The number of longitudinal stiffeners is fixed at eight per side of this structural design problem. This leaves the number of design variables to six defining transverse structures. The design variables description is shown in Table 3.2.

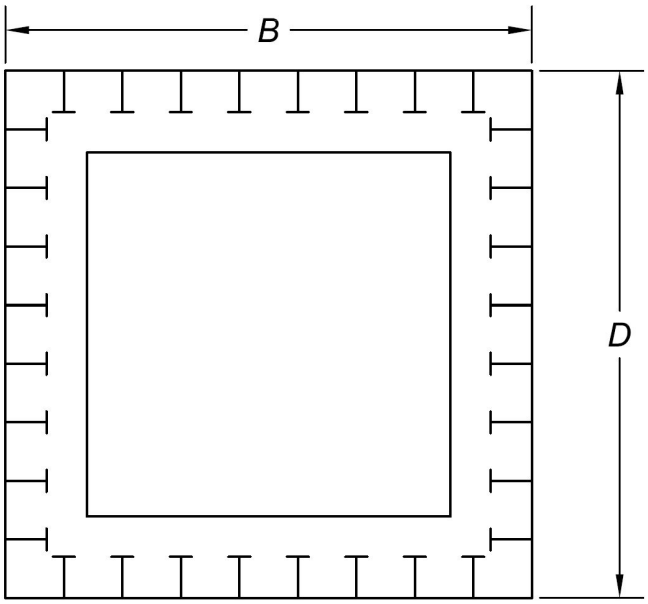


Figure 3.4: Cross section view.

Table 3.2: Design variables

Design variables	Description	Type	Min value	Max value
t_p	Thickness of plate	Continuous	3 mm	15 mm
h_{wt}	Transverse web height	Continuous	100 mm	800 mm
t_{wt}	Transverse web thickness	Continuous	3 mm	30 mm
b_{ft}	Web flange breadth	Continuous	5 mm	50 mm
t_{ft}	Web flange thickness	Continuous	4 mm	20 mm
n_{trans}	Number of transverse webs	Discrete	1	8

In this structural design, the weight of structure is the single objective to minimize. The structural weight is calculated by determining the volume of the material and then multiplying the density of steel, which is taken as 7.85 g/cm^3. The weight of structure is computed following the formula in Equation 3.4.

$$
\begin{aligned}
W &= 7.85\frac{g}{cm^3}(A_s * Length + A_w * n_{trans}) \\
A_s &= n_{stiff} * (h_w * t_w + t_f * b_f) + t_p * (2B + 2D) \\
A_w &= (h_{wt} * t_{wt} + b_{ft} * t_{ft}) * (2B + 2D)
\end{aligned}
\tag{3.4}
$$

In order to keep the design realistic, several requirements adopted from American Bureau of Shipping (ABS) High Speed Navy Craft rules are set as design constraints for this early stage structural design example. The six constraints are listed in Equation 3.5. Constraint g_1 puts a requirement on section modulus at amidship; g_2 is minimum plate thickness requirement; g_3 sets the buckling criteria for the plate; g_4 gives te minimum section modulus requirement for transverse web; g_5 is the buckling criteria for longitudinal stiffener; and g_6 makes sure that the web thickness meets the rules.

$$
\begin{aligned}
g_1 &: SM \geq SM_{required} \\
g_2 &: t_p \geq s\sqrt{\frac{pk}{1000\sigma_a}} \\
g_3 &: \sigma_E \geq 0.9m_1E(\frac{t_b}{s})^2 \\
g_4 &: SM_{web} \geq \frac{83.3 \times psl^2}{\sigma_a} \\
g_5 &: \sigma_e \geq \frac{EI_a}{C_1Al^2} \\
g_6 &: t_w \geq d_w/1.54(\frac{E}{\tau_y})^2
\end{aligned}
\tag{3.5}
$$

Table 3.3: SOGA parameters

GA parameter	Value
Population	400
Real crossover operator	Simulated binary
Binary crossover operator	Single point
Crossover percent	0.99
Mutation percent	0.01

These constraint functions are evaluated for every candidate design. Any violation of constraint function will be treated as penalty added to objective function. Therefore, any violation of the constraints would make the candidate design less desirable in evolutionary optimization. The robust design optimization problem is formulated in Equation 3.6.

$$\begin{aligned} \min \quad & \phi = W(\mathbf{x}, e) + \beta_k \sum G_i(\mathbf{x}, e) \quad (i = 1, 2, ..., 6) \\ \text{w.r.t.} \quad & \mathbf{x} = [t_p, h_{wt}, t_{wt}, b_{ft}, t_{ft}, n_{trans}]^T \\ \text{where} \quad & G_i = [min(0, g_i)]^2 \\ & U_l \leq e \leq U_u \end{aligned} \tag{3.6}$$

β_k is penalty coefficient set as 1000; e is design parameter representing compartment length, which is treated with bounded interval uncertainty. The optimal design outcome from this optimization is a design solution that has the best worst-case performance, and also satisfies all the design constraints while the design parameter varies.

The parameters used for the SOGA to solve the optimization problem in Equation 3.6 are listed in Table 3.3. In order to have an overall view of the performance of surrogate methodology, two additional optimization runs are conducted for references. First, the deterministic approach is used for the structural optimization, where the compartment length parameter is fixed without consideration of interval uncertainty. Second, the robust optimization problem formulated in Equation 3.6 is optimiza-

Table 3.4: Optimization results

Design variable	Deterministic optimum	Robust optimum	Surrogate optimum
$t_p\ (mm)$	8.26	8.26	8.26
$h_{wt}\ (mm)$	537.78	620.05	619.42
$t_{wt}\ (mm)$	9.29	10.68	10.73
$b_{ft}\ (mm)$	50.00	50.00	50.00
$t_{ft}\ (mm)$	20.00	20.00	20.00
n_{trans}	7	8	8
ϕ	35.53	40.31	40.44

tion with no surrogate mechanism turned on. Last, the introduced surrogate-assisted robust design optimization framework is applied here to solve the same robust structural design problem in Equation 3.6. The results from these three optimization runs are listed in Table 3.4 in terms of optimal design variables and penalized objective function values.

It can be observed from the results comparison in Table 3.4 that the robust design is more conservative design with approximately 15% more weight than the deterministic design. Consider that the length uncertainty band of 4 m is 50 % of the nominal compartment length for the case problem, this weight penalty for robust design is relatively small. Using the robust weight estimate would assure that the early stage structural weight estimates and configurations would remain feasible, if not optimal, throughout the subsequent design process as long as the compartment length did not vary outside the uncertainty band. Also this result is assured in a sense that with the interval uncertainty treatment, there is no corresponding probability that a constraint may be violated.

In comparison, the deterministic design is very sensitive to uncertainty in compartment length. In Figure 3.5 it is found that the deterministic design can easily violate constraints as the design parameter varies, causing the penalty function change dramatically. Clearly the deterministic design is not desirable, because any change in the design parameter L will make the design unfeasible. The robust design solution

provides a better solution considering the changing compartment length. It can be seen in Figure 3.5 that the robust solution is very stable in the interval range, this means future changes in compartment arrangement will not disrupt the performance of the robust design. In this regard, the robust solution trade a small amount of weight performance to achieve flexibility for later design stages.

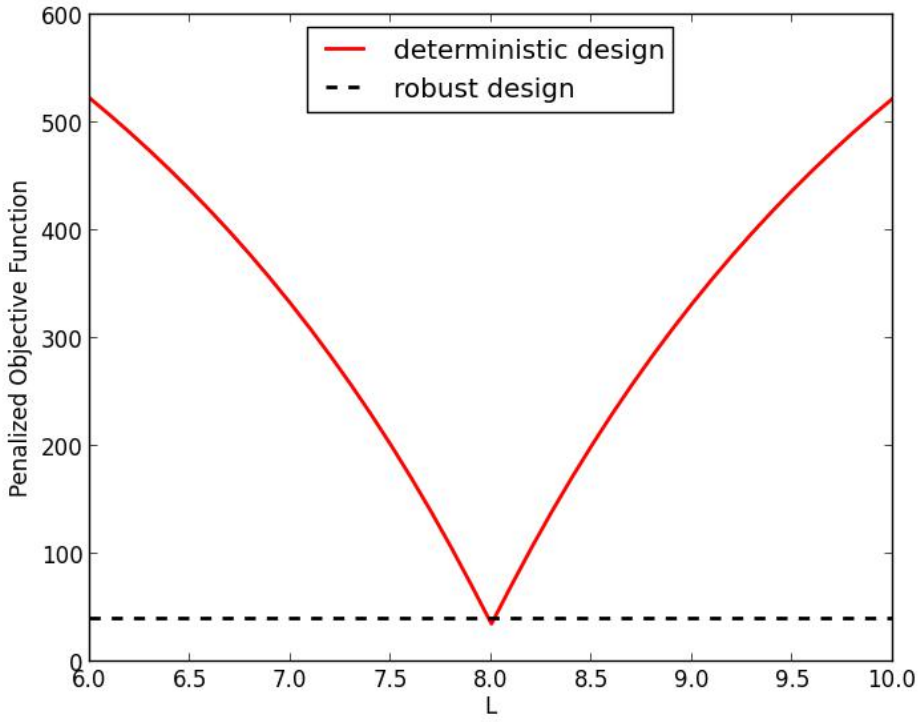


Figure 3.5: Design performance under interval uncertainty.

The surrogate solution is very satisfactory as it closely captures the robust solution, while greatly reducing the computational cost. Direct robust optimization requires 1,798,419 objective function evaluations, which means approximately 38 function calls are used per individual evaluation. This suggests that the inner loop search consumed a large amount of function evaluations to locate the worst-case performance. Surrogate method only used 8,000 function evaluations - the size of sampling dataset. The closely converged solutions in Table 3.4 indicate the surrogate model

performed well in approximating the original function with only a fraction of the function evaluations. This proves that proposed surrogate framework is a viable alternative to conduct robust optimization study in early stage design.

3.3 Robust and reliable structural design considering interval uncertainty

The interval uncertainty is assumed to be present in the parameters of a reliability-based design constraint as well as in the geometric configuration of the structure in various uncertain levels. This chapter continues to extend the robust design optimization framework in two ways.

First, in addition to compartment geometry uncertainty, interval uncertainty is also considered inside the formulation of a reliability-based structural constraint in the optimization process. This better represents early-stage marine design, where probabilistic models often employed, but the true value of their parameters remains uncertain. The probabilistic characteristic of the uncertainty encountered in early stage design may not be known precisely due to lack of adequate knowledge. A motivating example is the model uncertainty for loads computation models. Model uncertainty arises due to the scarce test results of the loads model, especially in cases of novel vessel design. *Ben-Haim and Elishakoff* (2013) have indicated that any error in the subjective assumption of distribution of the uncertainty in the reliability analysis can be harmful later on in the design. Interval uncertainty is used in this work to quantify the model uncertainties in the limit state function, as in most cases the non-deterministic parameters are only known within intervals (*Ferson et al.*, 2004). The worst case reliability (*Jiang et al.*, 2011) from interval analysis is examined to meet the target reliability requirement. Hence, the design is guaranteed to be also reliable against the uncertainty due to lack of knowledge.

Second, the optimization is now multi-objective with the width of the uncertainty interval treated as both a design variable and a component of one of the objective functions. This allows the trade space between structural performance and the amount of uncertainty to be quantified. The interval width reflects the confidence in the parameter choice and computation models. By performing additional measurements and collecting more information about the design, it is possible to improve the accuracy of the model and hence reduce the interval width. This design space reducing concept is similar to the set-based design concept (*Singer et al.*, 2009; *Hannapel and Vlahopoulos*, 2014) in naval design. However, such actions require time and the investment or resources. In order to show the payoff of reducing interval uncertainty in terms of structural performance improvement, a Multi Objective Genetic Algorithm (MOGA) optimizer - NSGA-II (*Deb et al.*, 2002) is employed to resolve the trade-off information between structural design performance and interval uncertainty measures.

3.3.1 Reliability analysis with interval parameters

In traditional reliability methods where only random variables are involved, the reliability can be calculated using the following equation:

$$Pr\{g(x) \geq 0\} = \int_{g(X)\geq 0} f_X(X)dx \tag{3.7}$$

To avoid the complex direct integration in Equation 3.7, an approximation method FORM (*Hasofer and Lind*, 1974) is normally used for computation. As introduced in Chapter II, the FORM method can be formulated as a minimization of the distance problem expressed as:

$$\beta = \min_{g(U)=0} \|U\| \tag{3.8}$$

This distance is termed as reliability index to assess the safety of structures. In

a typical reliability based design problem, the calculated reliability index is required to be larger than a threshold value β_t, termed as target reliability index. Considers interval uncertainty parameter $Y \in [Y_l, Y_h]$ in the limit state function, the transformed limit state function $g(U, Y)$ is bounded due to interval uncertainty Y, resulting in a bounded reliability index $\beta \in [\beta_L, \beta_H]$. Reliability analysis with interval parameter (*Jiang et al.*, 2011) focuses on finding the worst-case reliability index β_L using the following equations:

$$\begin{aligned} \beta_L = & \min_U \|U\| \\ \text{s.t.} \ & \min_Y \ g(U, Y) = 0 \end{aligned} \tag{3.9}$$

Locating of the worst case reliability index requires a nested optimization process. The inner loop finds the lower bound limit surface corresponding to a Y value in the interval. Then the outer loop follows the traditional FORM method as a distance minimization to find β_L.

The feasibility robustness introduced in Equation 3.6 is adopted here to consider interval uncertainty in structural geometry. To satisfy the feasibility robustness requirement, a candidate design must meet the following criteria:

$$\begin{aligned} & g_j(x, c) \leq 0, \ j = 1, 2, ..., J \\ & \forall c \in [c_l, c_h] \end{aligned} \tag{3.10}$$

x and c are vectors of design variables and design parameters respectively and J is total number of rule-based constraints. In the present work, only the design parameters are considered to be impacted by interval uncertainty.

With both feasibility robustness and worst-case reliability in the optimization process to account for interval uncertainties, the optimization formula incorporating

two schemes is presented as:

$$
\begin{aligned}
\min \quad & \mathrm{f}(x) \\
\text{s.t.} \quad & g_j(x,c) \leq 0, j = 1,2,\ldots,J \\
& \forall c \in [c_0 - \Delta_1, c_0 + \Delta_1] \\
& \beta_L = \min_U \|U\| \geq \beta_t \\
& g(U, Y^*) = \min_Y g(U,Y) \\
& \forall Y \in [Y_0 - \Delta_2, Y_0 + \Delta_2]
\end{aligned}
\tag{3.11}
$$

The optimal solution from the above procedure will be both robust and reliable against the interval uncertainties involved. The objective of this section is to disclose the impact of various interval ranges on the optimal design. Thus, interval width is treated as one of the design variables in the optimization framework. Doing so will modify Equation 3.11 into a multi-objective optimization approach.

3.3.2 Multi-objective optimization with surrogate model

The research focus is to determine the importance of interval uncertainty parameter on the design. Therefore, it is desirable to show how much reducing the interval uncertainty would improve design performance. This trade-off study is developed through a surrogate-assisted multi-objective optimization framework.

Some interval reduction measures can be found in *Li et al.* (2009). In this paper, a simple inverse of interval width is used as an indication of the investment needed to reduce the interval uncertainty. The investments to reduce both interval uncertainty in Equation 3.11 and a performance function are set as objective functions to minimize in the multi-objective optimizer.

The off-line RBF surrogate modeling is employed here to approximate the reliability simulations. The basic procedure of the surrogate-assisted reliability analysis

is outlined in the flow chart in Figure 3.6. To begin, a sampling plan using Latin Hypercube Sampling is made, and then the sampled dataset is collected using FORM analysis. In the optimization process, when the reliability analysis is required to evaluate a candidate design, a sorting algorithm is used to collect a number of sample points in the database. These points are used to train a local RBF surrogate model to replace the FORM method for reliability analysis. Afterwards, the worst-case reliability index will be located using Nelder-Mead method to search RBF surrogate in the interval uncertainty space.

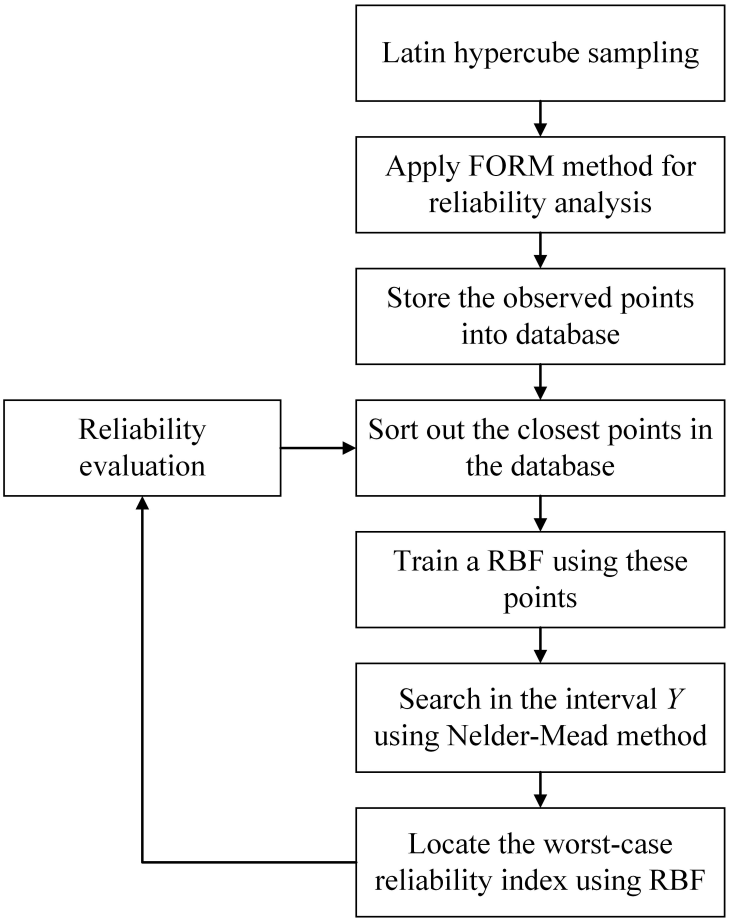


Figure 3.6: Reliability analysis using RBF surrogate model.

After introducing the multi-objective optimization problem and surrogate model-

ing strategy, the overall optimization framework is presented in the follows:

$$\begin{aligned} \min \quad & [W, I_1, I_2] \\ \text{where} \quad & I_1 = 1/\Delta_1,\ I_2 = 1/\Delta_2 \\ \text{s.t.} \quad & g_j(x, c) \leq 0, j = 1, 2, \ldots, J \\ & \forall c \in [c_0 - \Delta_1, c_0 + \Delta_1] \\ & \beta_L = \min \hat{f}(x, Y) \geq \beta_t \\ & \forall Y \in [Y_0 - \Delta_2, Y_0 + \Delta_2] \end{aligned} \tag{3.12}$$

In this formulation, W stands for a performance metric of the design, I_1 and I_2 represent the investment to reduce interval uncertainty in the uncertain design parameter and in the uncertain limit state function parameter respectively. $\hat{f}$ means that the RBF surrogate model is used as replacement of FORM method in reliability analysis.

3.3.3 Box girder structure case study

The simple box girder problem presented in the compartmentalization study is used here for the expanded investigation on interval uncertainty. The sketch of the structure is shown in Figure 3.3, the schematic cross section of the structure is shown in Figure 3.4. Independent design variables for this problem are listed in Table 3.5.

Two interval uncertainty terms are included. First, the length between bulkheads is assumed to have interval uncertainty. This represents a common situation where a structural configuration and weight estimate must be made before the final subdivision of the hull is established. The second interval uncertainty term is related to a first-yield reliability criteria for the girder under combined stillwater and wave bending moments. The second uncertainty term modifies the wave model uncertainty in the limit state. This case study attempts to minimize three objective functions.

Table 3.5: Design variables description for interval uncertainty MOGA

Design variables	Description	Min value	Max value
t_p	Thickness of plate	3 mm	15 mm
h_{wt}	Transverse web height	100 mm	800 mm
t_{wt}	Transverse web thickness	3 mm	30 mm
b_{ft}	Web flange breadth	5 mm	50 mm
t_{ft}	Web flange thickness	4 mm	20 mm
Δ_1	Geometry uncertainty interval range	0.1	1.0
Δ_2	Model uncertainty interval range	0.01	0.1

One is a performance metric defined as the structural weight, the other two are the investment in reduction of interval width in each domain respectively.

The six deterministic design constraint functions stated in Equation 3.5 are also used in this case study for feasibility robustness analysis against the geometric interval uncertainty bounded by $[-\Delta_1, \Delta_1]$.

In the reliability constraint function, a first-yield limit state function is used for the box girder design as a simple formula that is broadly representative of more complex ultimate limit states. It is expressed as follows:

$$\begin{aligned} G(x) =& SM * \sigma_y - M_{sw} - x_w * M_w \\ \text{where } SM &= \text{section modulus} \\ \sigma_y &= \text{yielding stress} \\ M_{sw} &= \text{still-water bending moment} \\ M_w &= \text{wave-induced bending moment} \\ x_w &= \text{model uncertainty with interval parameter } \Delta_2 \end{aligned} \tag{3.13}$$

The description of each of the component in the limit state function is shown in Table 3.6. SM, σ_y, M_{sw}, and M_w are assumed to be independent random variables. Their distribution type and parameters are chosen from *Mansour* (1993).

A hybrid uncertainty model x_w is used to represent the limited understanding of loads and load combinations on the structure. x_w takes the form of a normal

Table 3.6: Description of variables in limit sate function

Variable Term	Distribution	Mean	C.O.V.
SM	Lognormal	SM($\mathbf{x}$)	0.04
σ_y	Lognormal	235Mpa	0.07
M_{sw}	Normal	μ_{sw}	0.40
M_w	Gumbel	μ_w	0.09
x_w	Hybrid model	0.9	$[0.15 - \Delta_2, 0.15 + \Delta_2]$

distribution but its key parameter - coefficient of variance - is treated with interval uncertainty. The mean value of x_w is taken as 0.9, the coefficient of variance has a nominal value of 0.15, but bounded by $[-\Delta_2, \Delta_2]$.

The reliability analysis with interval uncertain parameter described previously can be applied to compute a worst case reliability index β_L. A target reliability index of 3.0 is set as the reliability constraint:

$$g_7 : \beta_L - 3.0 \geq 0 \tag{3.14}$$

The multi-objective optimization problem is stated as:

$$\begin{aligned}
&\min && [W(x), I_1(x), I_2(x)] \\
&\text{w.r.t.} && x = [t_p, h_{wt}, t_{wt}, b_{ft}, t_{ft}, \Delta_1, \Delta_2]^T \\
&\text{where} && I_1 = 1/\Delta_1,\ I_2 = 1/\Delta_2 \\
&\text{s.t.} && g_j(x, c) \leq 0, j = 1, 2, \ldots, 6 \\
& && \forall c \in [c_0 - \Delta_1, c_0 + \Delta_1] \\
& && g_7 :\ \beta_L - 3.0 \geq 0
\end{aligned} \tag{3.15}$$

W is the structural weight function to be minimized along with two interval reduction objective functions. The optimization problem is solved in NSGA-II optimizer. The simulated binary crossover is adopted here for crossover with an exponent of 4.0. Random mutation rate is set at a low probability of occurrence of 0.1%. The pop-

ulation size is set as 600 with generation number of 100 as the NSGA-II parameter choice.

3.3.3.1 Direct optimization results

First, the optimization problem was run without the surrogate modeling method to provide a reference solution. In this case, FORM method was used in computing the worst-case reliability. The direct optimization results are shown in the Figure 3.7. In order to better interpret the optimal solutions achieved, the objectives of interval width function I_1, I_2, are inverted to become the interval width Δ_1 and Δ_2.

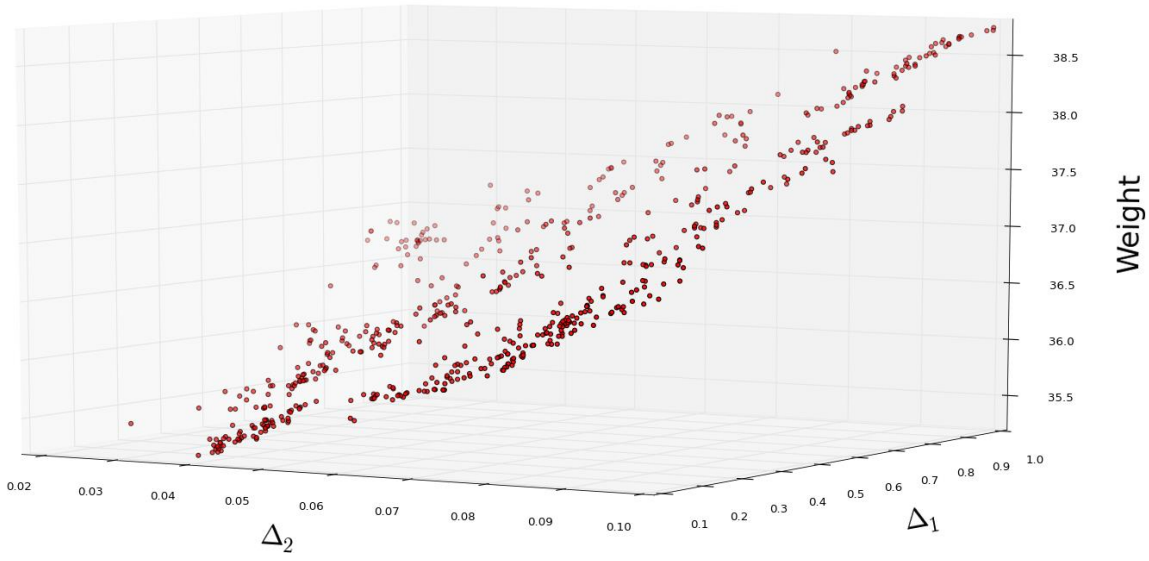


Figure 3.7: Pareto front of direct optimization.

Clearly the trade information is shown in the Pareto front. As seen, when interval uncertainty range becomes large (less investment effort was made to reduce the interval), an increase in weight is observed in order to keep the structure feasibility robust and reliable against larger interval uncertainties. It is also noticed that the Pareto optimal front followed a near hyperplane shape, indicating that the increase of interval range in either Δ_1 or Δ_2 can cause linearly penalty on design performance measured by the structure weight, though a more in depth analysis is needed to disclose the

relative importance between the two uncertainty parameters involved. The trade-off information from the Pareto front demonstrates that the information availability in the early stage can cause a possible monetary burden on the design in the worst case scenario.

3.3.3.2 Surrogate optimization results

The optimization process was repeated with the surrogate-assisted method introduced in Figure 3.6. A sampling size of 9,000 points and a local RBF modeling size of 150 points are used. The Pareto-optimal front generated from the multi-objective optimizer is shown in Figure 3.8 below.

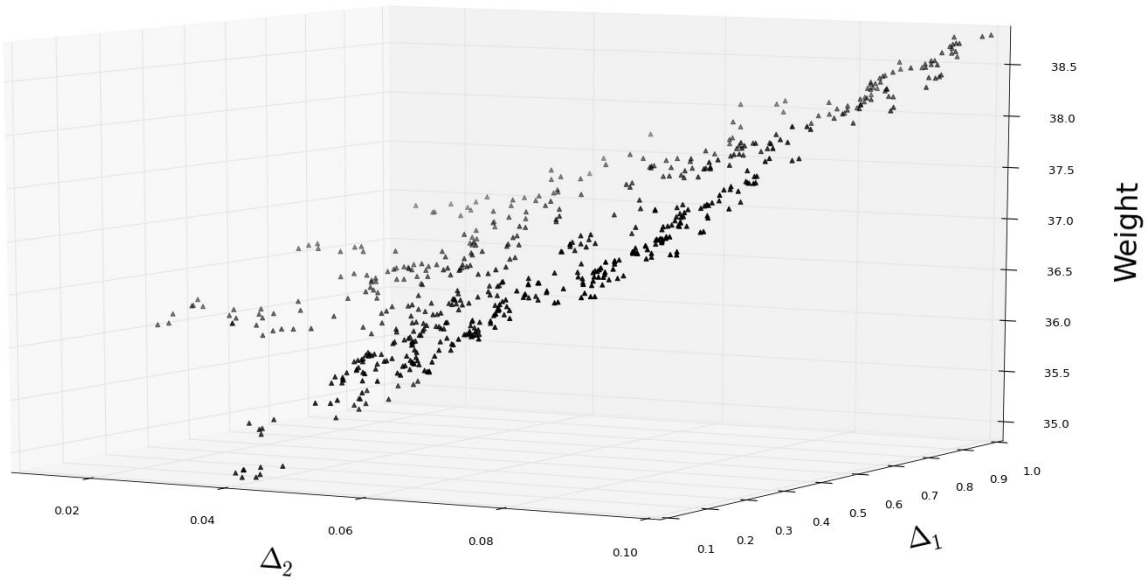


Figure 3.8: Pareto front of surrogate method.

On visual inspection, the Pareto front with the surrogate method is broadly similar to the direct optimal solutions. However, it is not exactly converged to the overall high-fidelity front. The number of reliability simulations spent for these two optimization run is presented in Table 3.7 below. It is obvious that the surrogate method significantly improved the efficiency of the optimization, as it used only 1.2% of the

Table 3.7: Computational cost comparison in two optimizations

Method	Number of reliability analysis used	Average number of reliability calls to locate β_L per individual
Direct FORM method	754,690	12.57
RBF approximation	9,000	0.15

reliability analysis calls of the direct optimization with the full FORM method.

3.4 Summary

This chapter has presented efficient design optimization frameworks to address epistemic uncertainties in early stage design. The presented frameworks are novel in the naval field to first examine interval uncertainty in marine structural designs. A robust design optimization framework is presented to consider interval uncertainty in compartment configuration. Afterwards, a multi-objective optimization framework further investigates interval uncertainty in both configuration and reliability constraint function. The overall impact of interval uncertainty is shown in trade study result. A RBF surrogate modeling method is implemented within the evolutionary algorithm for the efficient implementation of both uncertainty studies.

The presented robust design optimization framework is shown to successfully applied in investigating the compartmentalization problem in marine structural design. The case study indicated a 15% weight penalty to cover a 4 m uncertainty band in the compartment length. Using a fixed databased of 8,000 functional evaluations, the surrogate method can resolve a virtually identical robust design with less than 1% of the function evaluations used by a direct double-loop method. This demonstrated that such efficient robust design optimization strategy can be applicable to convert epistemic uncertainties in early stage into static weight penalties that can allow robust design convergence.

In the multi-objective optimization study, the method was shown to successfully

resolve the trade-off information between epistemic availability represented by various ranges of interval uncertainty and structural design performance. This provides a better perspective for decision makers to explore structural design options in early stage design. Also, the RBF surrogate modeling strategy proved to be capable in addressing the computational tractability problem associated with interval analysis. The surrogate method was demonstrated to deliver a closely converged Pareto front when compared to the front found by direct optimization, while used only a fraction of the exact solutions.

Nevertheless, improvements in the surrogate construction can still be expected. The presented surrogate modeling method is an off-line method, meaning the accuracy of surrogate model is fixed after the sampling procedure. In the multi-objective optimization study, the surrogate result does not converge exactly to high-fidelity FORM results. In the following chapters, the dissertation will discuss more on efficient surrogate-assisted multi-objective optimization methods, in particularly investigating advanced on-line surrogate modeling techniques that are capable in flexibly distribute computational resources, while delivering more converged Pareto optimal solutions.

CHAPTER IV

Surrogate-assisted variable fidelity multi-objective optimization

4.1 Introduction

In the last chapter, surrogate-assisted evolutionary optimization initially demonstrated its ability to address interval uncertainty design problems efficiently through an off-line construction methodology. However the off-line construction method of surrogate modeling is limited to applications in problem scope and complexity. In dealing with more challenging problems such as multi-objective optimization, or high dimensional function modeling, the off-line methodology may not be capable enough in providing accurate predictions for all the function space to be explored. The on-line surrogate construction method provides solution to continuously refine surrogate models within optimization process. This chapter will discuss the on-line surrogate modeling method, and a proposed novel surrogate method that can improve on-line surrogate modeling. The contents in this chapter was published by the author in *Liu and Collette* (2014a).

On-line surrogate method is aimed to adaptively improve the accuracy of the surrogate model throughout optimization runs. Various on-line surrogate modeling methods have been proposed to accomplish this task. Here, a Variable Fidelity Op-

timization (VFO) framework is used to illustrate the concept of on-line surrogate modeling strategy (*Zhu et al.*, 2014). In a typical VFO method, a Kriging surrogate model is constructed on-line to scale a low fidelity version of objective function to a high fidelity version of the same function. The VFO approach schedules high fidelity simulation in given generations within evolutionary optimization, so that the computational budget of VFO is fixed in achieving the true Pareto front. This method proved to work in simple multi-objective optimization problems. However, the previous VFO method became inefficient and struggled to converge when the number of objective functions increased beyond two. The key reason for this struggle is that Kriging surrogate model cannot deal with increasingly larger modeling problems.

In the VFO approach, the Kriging surrogate model is constructed around the evolving Pareto front in the optimization run. As the optimization problem moves to a higher number of objectives, the location of non-dominated solutions in the independent design variable space becomes more diverse. To maintain accuracy, more sampling points are needed, consequently causing the surrogate model size to increase. During the Kriging modeling process, an $N \times N$ matrix will be inverted. Thus Kriging surrogate model construction can be very time-consuming for large sample sizes (*Jin et al.*, 2001). Solving an extremely large Kriging model can be numerically unstable as the matrix becomes nearly singular, in which case, the Kriging predictions are unlikely reliable.

Multiple surrogates modeling (*Viana*, 2011) has been suggested to improve the prediction quality in surrogate-assisted evolutionary algorithms. *Jin and Sendhoff* (2004) have proposed using neural network ensembles for surrogate modeling, the center points in each ensemble was chosen for original fitness simulations. *Hamza and Saitou* (2005) have used polynomial surrogate ensembles in genetic algorithm for vehicle crash-worthiness design, where the ensemble of surrogate models effectively compensated the errors associated with individual surrogate model. The benefits of

using multiple surrogates have also been reported in *Goel et al.* (2007) and *Sanchez et al.* (2008). *Isaacs et al.* (2007) constructed spatially distributed multiple radius basis functions for multi-objective optimization.

This chapter presents a novel multiple surrogates management strategy to improve the performance of VFO method. A k-means clustering algorithm is employed to partition model data into local surrogate models. The comparison between single surrogate model strategy and multiple surrogates strategy through VFO is presented through benchmark multi-objective optimization test problems, and stiffener panel structural design problem. Performance metrics show that the proposed multiple surrogate handling strategy clearly outperforms the single surrogate strategy as surrogate size increase.

4.2 Variable fidelity optimization

In the real-world design optimization problems, usually there are various simulation functions for fitness evaluation with different levels of fidelity. As high-fidelity evaluations are generally more time-consuming, there is a need to balance fidelity with computational cost in optimization design. *Zhu et al.* (2014) proposed the VFO framework based on a variant of global-local approach (*Haftka*, 1991). In *Zhu et al.* (2014), the VFO was shown to work well in approximating the Pareto optimal front with fewer high-fidelity fitness function calls. In VFO approach, the high-fidelity function $f_h(x)$ can be approximated by a global simplified low-fidelity function $f_g(x)$ and a Kriging correction surrogate model $f_f(x)$, as shown in Equation 4.1.

$$f_h(x) = f_g(x) \times f_f(x) \tag{4.1}$$

In this formulation, the global approximation mathematical model f_g is a simplified function that runs rapidly with a relatively high coefficient of variation (COV)

in its prediction. The bias of the prediction is defined as:

$$bias = \frac{predicted}{actual} \tag{4.2}$$

This leads to a bias of 1.0 for perfect approximation, and a bias below or above 1.0 for less accurate prediction methods. The VFO formulation take advantage of the fact that the genetic algorithm starts off with random population, so lower fidelity can be used to evolve a rough Pareto front. During this process, an interpolation surrogate model can be constructed on-line. Afterwards, the surrogate model can be used through VFO formulation to transition the rough Pareto front to an approximated true Pareto front.

The central component in the variable fidelity scheme is the Kriging surrogate model f_f, which acts as a bridge function between the low-fidelity function and the high-fidelity function. The detailed description of Kriging derivation can be seen in Chapter 2.5.2. Here the ordinary Kriging method is used to approximate f_f through Equation 4.3.

$$f_f(x) \approx \hat{f}_f(x) = F + Z(x) \tag{4.3}$$

where F is the constant global error between the simplified and high-fidelity method, and Z is the local error term. A Gaussian correlation model is used here for the random process. Kriging prediction will provide both a predictor value for an unknown point to be evaluated as well as the Mean Square Error (MSE) s^2 for the prediction at this point to be returned. The prediction error provides key information in developing updating criteria to refine the Kriging model within variable fidelity framework (*Gano et al.*, 2006).

Training a Kriging surrogate model requires finding of the correlation parameter θ. The approach used here is to solve the following maximum likelihood optimization

problem in Equation 4.4.

$$Max(-\frac{n_s In(\sigma^2) + In(|R|)}{2}),\ 0 < \theta_k \leq \infty \tag{4.4}$$

The computational cost for solving the Kriging model becomes non-trivial when the sampling size is large. The determination of θ vector in Equation 4.4 via optimization becomes much more difficult as the size of the model grows, owing to both numerical difficulties in matrix inversion and the fact that this inversion must be repeated within the VFO framework. Though VFO provides good computational strategy in building Kriging model on-line, a single Kriging model is not capable enough for VFO to tackle larger problems in higher dimensional space.

4.3 Clustered multiple surrogate modeling for VFO

Multiple surrogate modeling strategy is developed to address the challenge in VFO framework. The means to determine how to split a single large Kriging model into multiple Kriging models is needed. Clustering provides an attractive technique for this purpose. In clustering, a large dataset is separated into subsets that are more closely related to each other than the other members of the overall data set. The advantage of clustering is that the algorithm can carry out this separation without external guidance, making it ideal for inclusion in an automated optimization procedure. A clustering algorithm is employed to partition the Kriging sample dataset into several local subsets, where a single, smaller Kriging model is trained in each one of these local subsets.

The rational of proposing clustering embedded Kriging construction strategy is twofold:

First, from the viewpoint of computational cost, solving several small Kriging models is much faster than solving a Kriging model of large sample size. The com-

putational complexity of Kriging is $O(N^3)$, where N is the sample size. If the sample size is divided into k small Kriging samples, then the computational complexity of all those Kriging constructions becomes $k \times O((N/k)^3)$, the computational cost of k-means algorithm is roughly $O(N)$ (*Wu et al.*, 2008). Therefore, a large amount of computational cost reduction is possible if N is large.

Second, using multiple Kriging models can improve the prediction quality when the number of sample points is large. The optimization of the Kriging parameter indicated in Equation 4.4 requires solving a linear system. When the number of Kriging sample points becomes large, it can cause numerical difficulties as the matrix inversion becomes ill-conditioned. Solving several small Kriging models rather than one large model can substantially reduce the ill-conditioning matrix encountered in calculation. As less numerical round-off occurs, the overall predictor accuracy is improved.

The k-means clustering algorithm (*MacQueen et al.*, 1967) is employed here to improve the efficiency of surrogate modeling in VFO. k-means clustering is a common method in partition a data set into k clusters. The k-means clustering procedure follows the steps described in Algorithm 2.

Algorithm 2 General procedure for k-means clustering algorithm.

BEGIN
1: Initialize the clustering algorithm by randomly generating k cluster centers;
2: Assign each point in the dataset to the nearest cluster center;
3: Re-compute the cluster centers based on the points assigned to each cluster;
4: Step 2 and Step 3 are repeated until convergence.
END

With the proposed new Kriging construction strategy, the variable fidelity updating scheme proposed in *Zhu et al.* (2014) is reformulated to incorporate the clustering embedded Kriging surrogate models. The new variable fidelity scheme is implemented in the standard NSGA-II multi-objective genetic algorithm (*Deb et al.*, 2002). The major difference compared to previous standard VFO method is in the on-line Kriging

model construction, where multiple Kriging surrogates are built and refined with the help of k-means clustering algorithm. Figure 4.1 shows the flowchart of the method.

The detailed process of each step in the proposed VFO updating strategy using clustered multiple Kriging models is described as follows:

1. To initialize the optimization, a random population is selected. The first few generations only use the low-fidelity model for fitness evaluation. A rough Pareto front is evolved based on the low fidelity model.

2. After a given number of generations, termed the offset, a number of individuals are chosen from the current non-dominated Pareto front and sent for high-fidelity analysis. The individuals are selected based on a simple inter-individual distance metric in objective function space. The number of individuals is termed the density in the update scheme. Also worth noting is that the number selected could be set equal to an integer that is a multiple of the available processing cores of the computer being used. This ensures that full capacity of available processing power is always used when updating the surrogate model. Based on the number of high- fidelity evaluations processed, the clustering algorithm partitions all sample points into several cluster point sets. Afterwards, multiple Kriging models are formed in each of these clusters. The centers of all these clusters are stored in the database. The number of clusters is set to dynamically increase with increasing data set size.

3. Subsequently in the optimization process, when a candidate point is to be evaluated, the distances of this point to all the cluster centers are calculated and compared. The closest Kriging model is then selected. After, the fitness evaluation of this point uses either the low-fidelity solution scaled by the Kriging model, or just the low-fidelity solution, if the prediction error of the Kriging model does not pass the accuracy criteria discussed below.

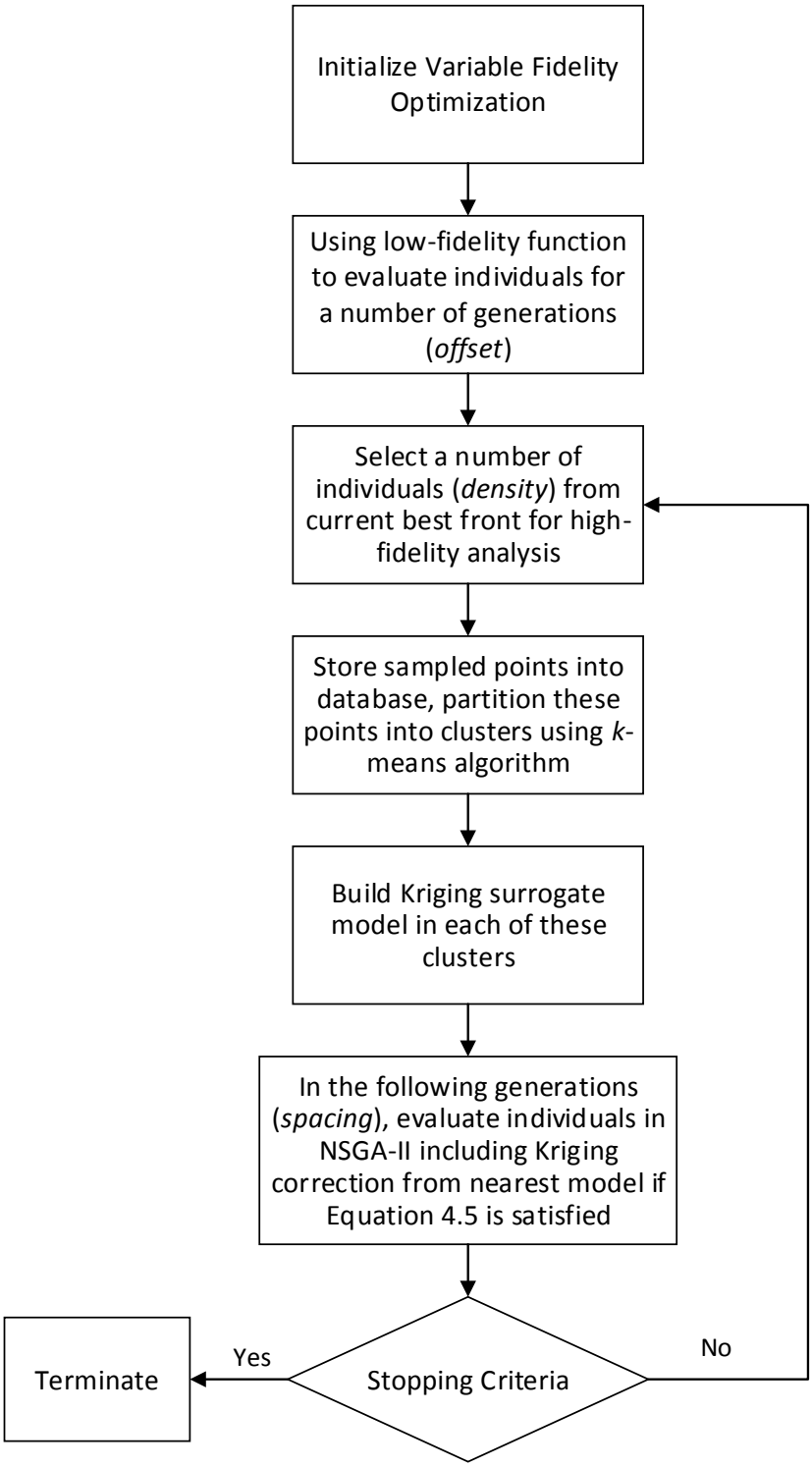


Figure 4.1: Flow chart of the clustering embedded VFO method.

4. After a fixed number of generations have passed, where the number is termed spacing, another set of individuals equal to the density is selected from the current non-dominated front and sent for high-fidelity analysis. These individuals are selected by computing the prediction error from their respective Kriging models for all the points in the current best non-dominated front. The individuals with the highest error are sent for high- fidelity analysis. In this updating, the clustering algorithm is applied again to help form a new set of Kriging models. The number of clusters dynamically increases as more high-fidelity analyzed points are added. Generally, the number of Kriging models increases in every update generation.

By updating multiple Kriging models in this way, a variable fidelity optimization scheme is built to interleave the high-fidelity fitness prediction model with a rapid low-fidelity model corrected by a computationally cheap surrogate model. As the best non-dominated front converges to the true global Pareto front for the given problem, each of the constructed Kriging models performs better in its region of the Pareto front. The Kriging models work together to provide a very accurate approximation in the complex multi-dimensional design space occupied by the Pareto front.

In the variable fidelity optimization process, the Kriging models are refined in every high-fidelity analysis update. It is expected that the Kriging models may not be accurate enough for all the objective function calls made by the optimizer, especially in the early part of the optimization process. To address this concern, the prediction error of the Kriging model is used to examine the surrogate model error, and then the optimizer will determine whether to use the uncorrected or the corrected form of the low-fidelity model, based on the following criteria (*Zhu et al.*, 2014):

$$\sqrt{\epsilon_K^2 + \epsilon_H^2} < \epsilon_L \tag{4.5}$$

where ϵ_K is the estimated non-dimensional error of the chosen local Kriging model at the design point being evaluated, the value of ϵ_K can be expressed as:

$$\epsilon_K = \frac{\sqrt{S^2(x_s)}}{\hat{f}_f(x_s)} \tag{4.6}$$

where $\hat{f}_f(x_s)$ is the standard Kriging predictor from Equation 4.3 and $S^2(x_s)$ is the mean square error of the predictor. In Equation 4.5, ϵ_H and ϵ_L are prior estimates of the Coefficient of Variation (COV) of the bias of the high and low-fidelity models, respectively.

The variable fidelity criterion in Equation 4.5 compares the relative error of the Kriging model used plus the high-fidelity model to the error of only using the low-fidelity model. If the criterion in Equation 4.5 is not satisfied, $f_f(x)$ is taken as 1.0, which means that the optimizer uses the uncorrected low-fidelity model for fitness objective function evaluations without Kriging correction. This ensures that Kriging prediction, if accepted, will not degrade the accuracy of the fitness evaluation.

For stability of the method, two additional conditions are added. If the update density is greater than the number of points in the current best non-dominated front, the process continues on down through the current fronts in domination order. Additionally, as high-fidelity analyzed points are added to the Kriging models, they are required to fall a certain distance away from the existing points in the Kriging models. This prevents a group of very similar design points from being added during an update.

The NSGA-II multi-objective genetic algorithm optimizer proposed by *Deb et al.* (2002) was used in this work to solve the optimization problem. The detailed description of NSGA-II can be seen in Chapter 2.3.2. The simulated binary crossover algorithm (*Deb*, 2001) is adopted here for crossover with an exponent of 4. Random mutation rate is set at a low probability of occurrence of 0.1%.

4.4 Numerical benchmark problem tests

To verify the applicability of the proposed variable fidelity optimization scheme using clustering implemented Kriging models, a series of benchmarking test problems from past significant multi-objective optimization studies were used to conduct the examination. First,three test problems from *Fonseca and Fleming* (1998)'s study (QV),*Poloni* (1995)'s study (POL) and *Quagliarella* (1998)'s study (QV) were chosen for initial tests. Afterwards, the same Zitzler-Deb-Thiele (ZDT) tests from *Zitzler et al.* (2000) that was conducted in the standard VFO method (*Zhu et al.*, 2014) is used again to present a comparison between the proposed clustered multiple Kriging models approach and the original single Kriging surrogate model approach.

The test problems are described in Table 4.1. All of them are minimization problems with two objective functions, covering different level of difficulties. None of the test problems were structured with a low and high fidelity version of the objective functions, therefore, a low-fidelity version of the second objective function was developed for each test problem. These are listed in the rightmost column of Table 4.1.

4.4.1 Results for FON, POL and QV with clustered surrogates VFO

The first three problems were run using the VFO parameters given in Table 4.2. The optimization solutions are visualized in Figure 4.2. The true Pareto fronts of these three problems are visualized in the figures in the left hand column, plotted as a heavy solid line. As seen, FON has a nonconvex front, POL has a nonconvex and disconnected front, and QV has an extreme concave front. The low-fidelity front solutions are shown in a dashed line on the same plots, as can be seen the low-fidelity solutions reflect the overall trend of the true Pareto front, but with some significant differences in detail and magnitude. The NSGA-II algorithm is able to robustly converge to each front as shown by the circular and downward-facing triangle solutions.

Table 4.1: Test problems description

Problem	n	Variable bounds	Objective functions	Low fidelity function for test
FON	3	$[-4,4]^n$	$f_1(x) = 1 - exp\left(-\sum_{i=1}^{3}(x_i - \frac{1}{\sqrt{3}})^2\right)$ $f_2(x) = 1 - exp\left(-\sum_{i=1}^{3}(x_i + \frac{1}{\sqrt{3}})^2\right)$	$f_{2\_low} = f_2(a) + f_2'(a)(x-a)$ a=[0,0,0]
POL	2	$[-\pi,\pi]^n$	$f_1(x) = 1 + (A_1 - B_1)^2 + (A_2 - B_2)^2$ $f_2(x) = (x_1+3)^2 + (x_2+1)^2$ $A_1 = 0.5\sin 1 - 2\cos 1 + \sin 2 - 1.5\cos 2$ $A_2 = 1.5\sin 1 - \cos 1 + 2\sin 2 - 0.5\cos 2$ $B_1 = 0.5\sin x_1 - 2\cos x_1 + \sin x_2 - 1.5\cos x_2$ $B_2 = 1.5\sin x_1 - \cos x_1 + 2\sin x_2 - 0.5\cos x_2$	$f_{2\_low} = x_1^2 + x_2^2$
QV	5	$[-5,5]^n$	$f_1(x) = (1/n\sum_{i=1}^{n} x_i^2 - 10(2\pi x_i) + 10)^{1/4}$ $f_2(x) = (1/n\sum_{i=1}^{n}(x_i - 1.5)^2 - 10(2\pi(x_i - 1.5)) + 10)^{1/4}$	$f_{2\_low} = f_2(a) + f_2'(a)(x-a)$ a=[0,0,0,0,0]
ZDT1	30	$[0,1]^n$	$f_1(x) = x_1$ $f_2(x) = g(x)\left(1 - \sqrt{\frac{x_1}{g(x)}}\right)$ $g(x) = 1 + \frac{9(\sum_{i=2}^{n} x_i)}{n-1}$	$f_{2\_low} = f_2(a) + f_2'(a)(x-a)$ a=[0,...0]
ZDT2	30	$[0,1]^n$	$f_1(x) = x_1$ $f_2(x) = g(x)\left(1 - \frac{x_1}{g(x)^2}\right)$ $g(x) = 1 + \frac{9(\sum_{i=2}^{n} x_i)}{n-1}$	$f_{2\_low} = f_2(a) + f_2'(a)(x-a)$ a=[0,...0]
ZDT3	30	$[0,1]^n$	$f_1(x) = x_1$ $f_2(x) = g(x)\left[1 - \sqrt{\frac{x_1}{g(x)}} - \frac{x_1}{g(x)}\sin(10\pi x_1)\right]$ $g(x) = 1 + \frac{9(\sum_{i=2}^{n} x_i)}{n-1}$	$f_{2\_low} = f_2(a) + f_2'(a)(x-a)$ a=[0,...0]

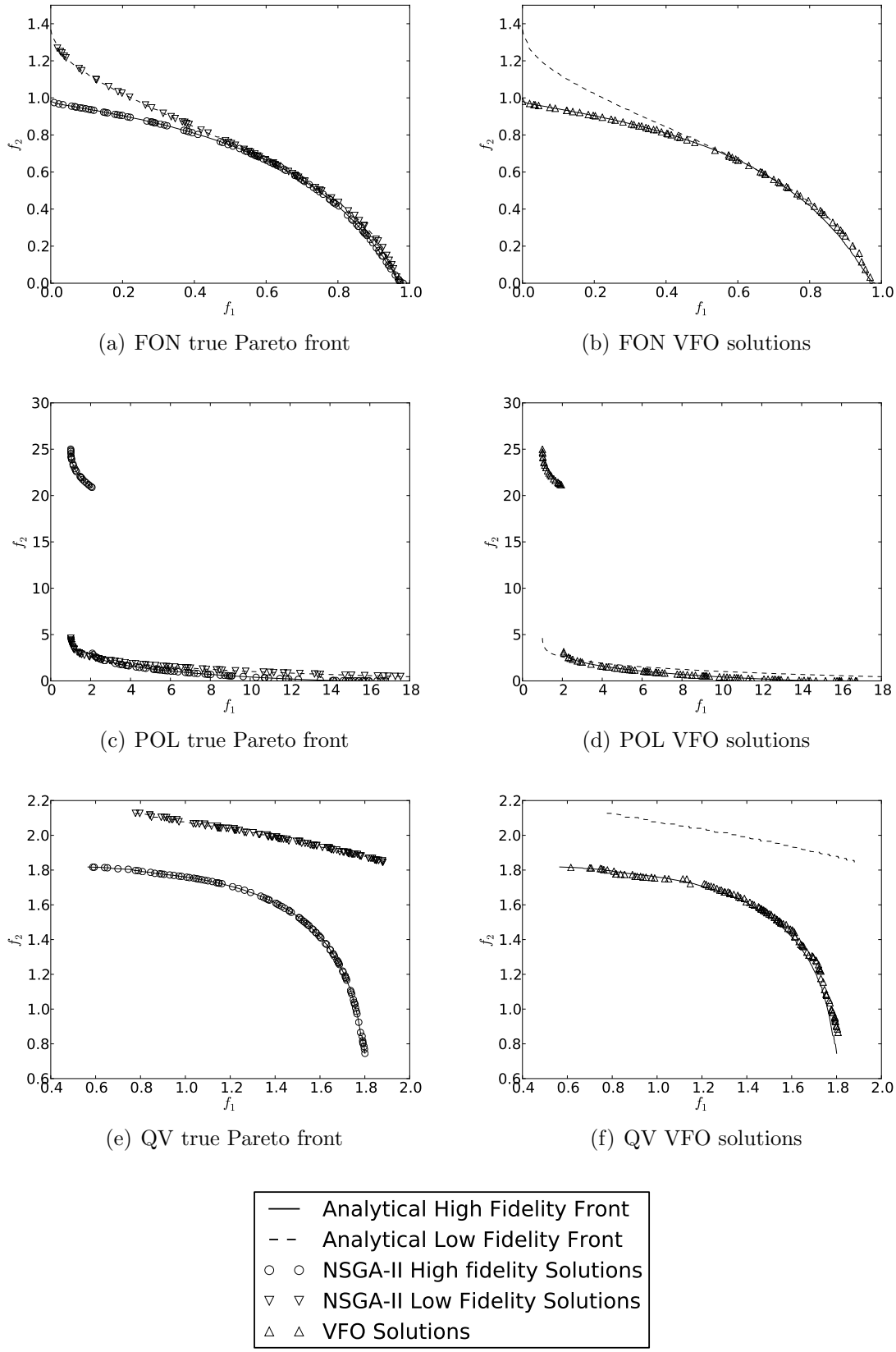
Table 4.2: VFO parameters for FON, POL and QV tests

Parameter	FON	POL	QV
Population Size		200	
Generation number		300	
Offset		30	
Spacing		20	
Density		10	

For each problem, the proposed VFO method from Chapter 4.3 was applied, and the result are shown in the right-hand column of Figure 4.2. All the three VFO solutions have managed to converge to the true or high-fidelity Pareto fronts, showing that the mechanism of VFO is working effectively. In general, the solutions along the resolved Pareto fronts are well-distributed, though the results for the QV test case are not quite as smooth as the other two test cases. According to the VFO schedule defined in Table 4.2, each case used 140 high-fidelity function calls in total. That means for each found point along the Pareto front in the proposed VFO method, only 0.7 high-fidelity analysis was needed. These 140 sampled points were partitioned and trained in 4 clustered Kriging models at the end of the optimization run for each case.

4.4.2 Comparison test using ZDT problems

The popular ZDT test problems ZDT1, ZDT2, and ZDT3 were previously explored with the early version of the VFO algorithm, with strong results presented for 5-variable version of the problem (*Zhu et al.*, 2014). Here, we expand the size of the problem to 30 variables, and compare the clustered Kriging model algorithm with the original single Kriging model algorithm proposed by *Zhu et al.* (2014). The problem sets contain features that are very representative of real world optimization problems. ZDT1 has a convex, continuous Pareto-optimal front, while ZDT2 has a non-convex, continuous Pareto-optimal front. In ZDT3 problem, a sine function is introduced to



(a) FON true Pareto front

(b) FON VFO solutions

(c) POL true Pareto front

(d) POL VFO solutions

(e) QV true Pareto front

(f) QV VFO solutions

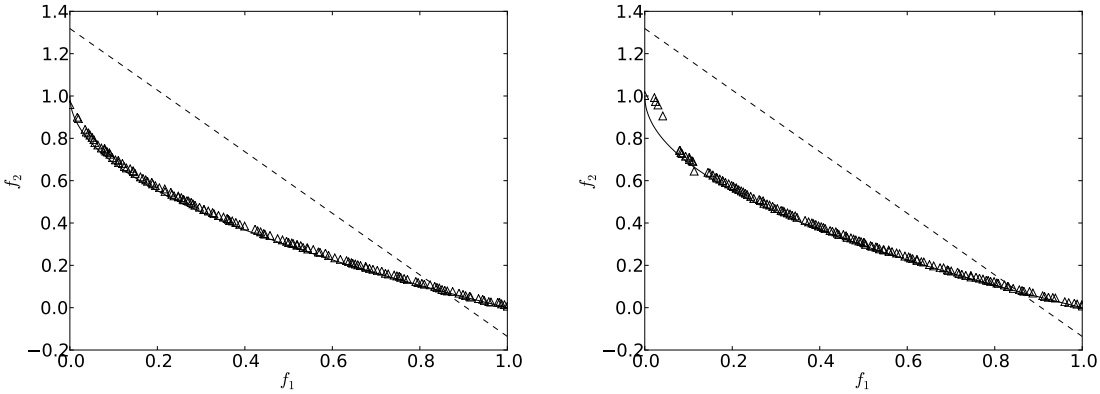
Figure 4.2: Optimization results from FON, POL, and QV test problems

the objective function, which leads to several discrete non-continuous convex parts in the Pareto front. This discrete feature causes additional difficulty in converging to the true Pareto front.

These ZDT test problems are simple mathematical functions with no correspondingly low-fidelity functions. Here, the same approach as that which was presented in *Zhu et al.* (2014)'s work is applied. Within this approach, the original ZDT function was used as the high-fidelity function, and a Taylor series expansion with the function value and first derivative terms was used as the low-fidelity function. The VFO approach was applied in the second objective function for each ZDT problem. For problem ZDT1 and ZDT2, the Taylor expansion was taken at the midpoint of the design variable space. As for ZDT3, the expansion has to be shifted to the point $x_n = 0.6$, otherwise the two objectives do not compete in the low-fidelity model. The high-fidelity and proposed low-fidelity functions for each ZDT problem are shown in Table 4.1.

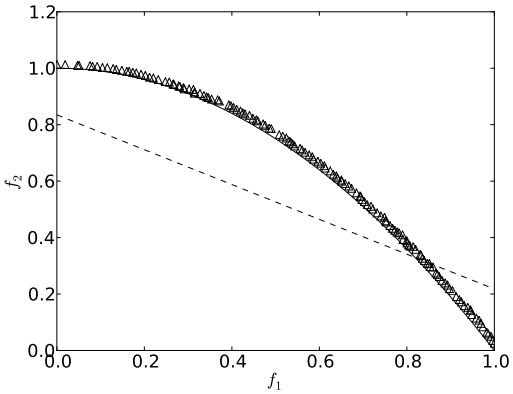
As previously stated, building an approximation model in higher number of independent variable space is difficult and requires a large number of sampling points. It is often referred to as the curse of dimensionality in literature (*Forrester et al.*, 2008). Here the proposed VFO method with multiple Kriging models and *Zhu et al.* (2014)'s VFO method with a single Kriging model were both tested in the same ZDT problem set with an increased 30 variables. The optimization parameters are defined in Table 4.3, and were kept the same for both the single Kriging model and clustered multiple Kriging model versions of the problem. A visualization of the optimization results are shown in Figure 4.3.

In Figure 4.3 the analytical Pareto fronts are plotted in solid and dashed lines for the high and low fidelity versions of each ZDT problem. Found solutions by the VFO method are plotted on top as markers, with the revised cluster Kriging model approach shown in the left-hand column and the original single Kriging model approach

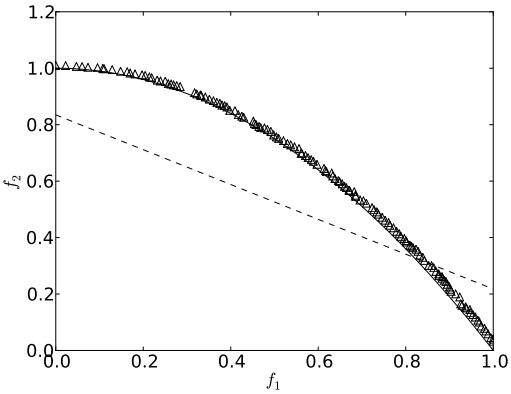


(a) ZDT1 solutions using Multiple Surrogates

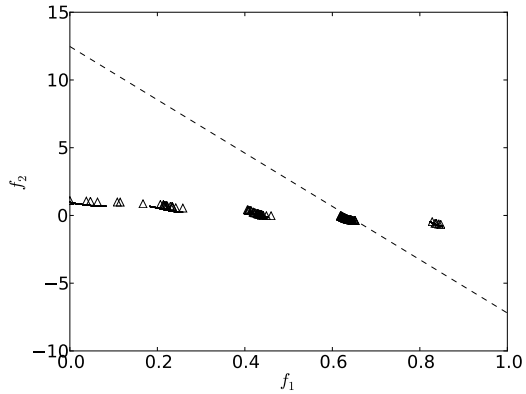
(b) ZDT1 solutions using Single Surrogate



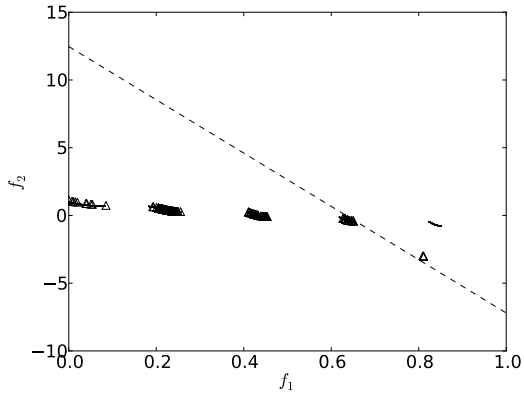
(c) ZDT2 solutions using Multiple Surrogates



(d) ZDT2 solutions using Single Surrogate



(e) ZDT3 solutions using Multiple Surrogates



(f) ZDT3 solutions using Single Surrogate

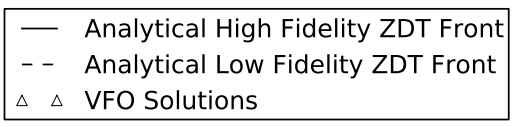


Figure 4.3: Comparison of Zhu's single surrogate with proposed clustering approach on 30 variable ZDT problems

Table 4.3: Optimization Parameters for Comparison Test

Parameter	ZDT1	ZDT2	ZDT3
Population Size		140	
Number of Variables		30	
Generation Number		160	
Offset		30	
Spacing		10	
Density		40	

shown in the right-hand column. It is clear that the VFO results all converged to the true Pareto fronts, indicating that the proposed method works well in solving each of the ZDT benchmarking problems with different levels of difficulties.

Additionally, it is evident that the proposed clustering implemented VFO works better than the original VFO method as the number of variables increases to 30. As can be seen in the subfigure (b) in Figure 4.3, errant solutions emerge at the left side of Pareto front, and in the subfigure (f) in Figure 4.3, the VFO solutions failed to converge to right part of Pareto front. Meanwhile, the proposed multiple Kriging models successfully assisted the VFO converging to the high-fidelity Pareto fronts in these ZDT test problems. Additionally, the proposed method took less wall-clock time to complete the optimization run. This potential speed advantage was more rigorously quantified in the example of Section 4.5. Based on these results, the clustering VFO approach appears promising for further study on structure design problems.

4.5 Stiffened panel design study

4.5.1 Design problem statements

The proposed method was applied to a structure optimization of a tee-stiffened panel, which is a typical deck structure. *Zhu et al.* (2014) conducted a two-objective optimization, regarding the weight and strength of this structure. Here, the design

Table 4.4: Chromosome for Panel Design

Variable	Minimum Value	Maximum Value
n_{stiff}	1	16
t_p	3 mm	12 mm
t_w	3 mm	12 mm
h_w	30 mm	180 mm
t_f	3 mm	20 mm
b_f	30 mm	90 mm

problem was further explored in the weight, strength, and cost three-dimensional trade space. The added cost objective function increases the complexity of trade space as the distribution of non-dominated solutions extends to include cheap but heavy panels. A complete comparison of the proposed method and original VFO method can be shown by studying this multi-objective optimization problem. The objective functions are defined by six independent variables. Five of the variables relate to the plate thickness and stiffener dimensions, and one concerning the number of stiffeners used. The independent variables are defined in Table 4.4. The panel breadth was fixed at 3000 mm, the panel length was fixed at 1500 mm, and the remaining variables are depicted in Figure 4.4. In this work, only the strength objective function was investigated by the variable fidelity approach. Material was assumed to be mild steel with a 250 MPa yield stress and an elastic modulus of 207,000 MPa.

4.5.1.1 Panel weight

The panel weight was calculated by multiplying the volume of the material and the density of standard carbon structural steel, which was taken as 7.85 g/cm3. The panel weight objective function was expressed as follow:

$$W = 7.85\frac{g}{cm^3}a(t_pB + n_{stiff}A_s)A_s = h_wt_w + b_ft_f \tag{4.7}$$

4.5.1.2 Panel cost

The cost of the panel was computed using standard cost estimating approaches which account for cost of materials and labor cost for panel fabrication. In this work, the formula proposed by *Rahman and Caldwell* (1995) was simplified for a panel stiffened in a single direction and used to predict the total cost, C.

$$C = C_{plate} + C_{stiff} + C_{weld} + C_{intersect} + C_{electric} \tag{4.8}$$

The components in this formula are presented in Table 4.5. A detailed description of the cost formulas in this table can be found in the paper presented by *Rigterink et al.* (2013).

4.5.1.3 Panel strength

In *Zhu et al.* (2014) two different strength calculation methods were selected to illustrate the variable fidelity optimization scheme. Here, the same strength regression models were used to verify the proposed revised VFO scheme. The two compressive strength models both idealized a single tee-stiffener and attached plate as a beam-column, ignoring supports from the longitudinal girders at the panel edge.

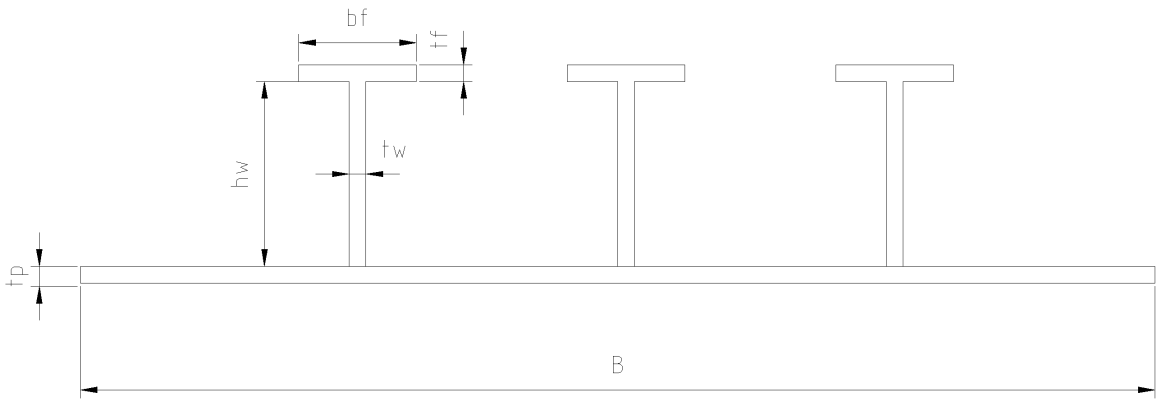


Figure 4.4: Sketch of independent variables.

Table 4.5: Components of the cost model.

Coefficient	Description	Formula
C_{plate}	Cost of materials for plating	$W_p P_a$
C_{stiff}	Cost of materials for longitudinal stiffeners	$W_{ls} C_{lm} P_a$
C_{weld}	Cost of welding for longitudinal stiffeners	$n_{stiff} a C_{ls} P_s$
$C_{intersect}$	Cost of intersections between longitudinal stiffeners and transverse frames and preparation of brackets and joints	$n_{stiff} N_w (C_{bj} + C_{is}) P_s$
$C_{electric}$	Cost of electricity, electrodes and fabrication cost of longitudinal stiffeners	$n_{stiff} a (C_{ee} + C_{fb}) P_s$

Low fidelity function

The regression model from *Paik and Thayamballi* (2003) was used here as the low-fidelity function to compute strength function. Paik's method is usually employed as a rapid strength formula in the early design stage. The ultimate compressive strength σ_u is based on the yield stress, σ_y , and the plate and column elastic slenderness parameters:

$$\begin{aligned} \sigma_u &= \frac{\sigma_y}{\sqrt{q}} \leq \frac{\sigma_y}{\lambda^2} \\ q &= 0.995 + 0.936\lambda^2 + 0.17\beta^2 + 0.188\lambda^2\beta^2 - 0.067\lambda^4 \end{aligned} \tag{4.9}$$

The elastic slenderness parameters are related to the plate thickness,t_p, span of the plate between stiffeners, b, the transverse frame spacing, a, the material properties σ_y and E. The column slenderness parameter, λ, depends on the area, A, and the moment of inertia of one tee-stiffener and attached plate, I, as shown in Equation 4.10.

$$\begin{aligned} \beta &= \frac{b}{t_p}\sqrt{\frac{\sigma_y}{E}} \\ \lambda &= \frac{a}{\pi r}\sqrt{\frac{\sigma_y}{E}} \\ r &= \sqrt{\frac{I}{A}} \end{aligned} \tag{4.10}$$

High fidelity function

The high-fidelity function is the *Faulkner et al.* (1973) panel formulation, extended to include the revised tripping formulation proposed by *Faulkner* (1987). This approximation function has been extensively compared to experimental data and has been found to be very accurate. A description of Faulkner's method in *Zhu et al.* (2014) is adapted here. The elastic column slenderness parameter is modified to account the impact of plate buckling on both the effective width of the plate and the overall column stiffness. Thus, it becomes a function of edge stress, σ_e:

$$\lambda(\sigma_e) = \frac{a}{\pi}\sqrt{\frac{\sigma_y\left(A_S + b_e t_p\right)}{EI'}} \tag{4.11}$$

$$\beta_e(\sigma_e) = \frac{b}{t_p}\sqrt{\frac{\sigma_E}{E}} \tag{4.12}$$

$$\frac{b_e}{b} = \begin{cases} \frac{2.0}{\beta_e} - \frac{1.0}{\beta_e^2} & \beta_e \geq 1 \\ 1.0 & \beta_e < 1 \end{cases} \tag{4.13}$$

$$\frac{b'_e}{b} = \begin{cases} \frac{1.0}{\beta_e} & \beta_e \geq 1 \\ 1.0 & \beta_e < 1 \end{cases} \tag{4.14}$$

where b_e is estimate of effective breadth accounting for plate buckling, and b'_e is the tangent effective breadth, I' is the tangent moment of inertia computed by replacing b with b'_e. The average stress at failure is determined with consideration of reduced effective plate area in cases where the plate buckles:

$$\frac{\sigma_C}{\sigma_y} = \begin{cases} 1.0 - 0.25\lambda_e^2 & \lambda_e \leq 1.41 \\ \frac{1.0}{\lambda_e^2} & \lambda_e > 1.41 \end{cases} \tag{4.15}$$

$$\frac{\sigma_u}{\sigma_y} = \frac{\sigma_e}{\sigma_y}\sqrt{\frac{A_S + b_e t_p}{A_S + b t_p}} \tag{4.16}$$

An iterative process is required to find the edge stress that cause column failure, $\sigma_e = \sigma_C$. Here, the stiffener tripping stress is also calculated by the method proposed in *Faulkner* (1987), the lower of the panel buckling and tripping stresses is taken as the governing stress.

4.5.1.4 Enumeration

In order to examine the performance of the variable fidelity optimization result, an independent approximation of the true Pareto front was needed. The enumeration simulation conducted by Yi-Jen Wang was used here for reference of Pareto fronts. In Yi-Jen Wang's simulation procedure, each variable range was divided into 21 evenly spaced values for the variable ranges showed in Table 4.4, except for the integer variable of stiffener number, which varies from 1 to 16 at integers only. There are about 4 million design points to be evaluated per stiffener number, and 65 million design points overall. The NyX cluster computer at the University of Michigan was used to conduct the enumeration run. The non-dominated points that consists the true Pareto front of this three-dimensional optimization problem were saved in a database for use in the performance metrics introduced next.

4.5.2 Performance metric for evaluating VFO results

To evaluate the performance of the revised VFO strategy, statistical metrics that can determine the Kriging prediction accuracy and the quality of the Pareto front are needed. Here four measures of the found Pareto front originally proposed in *Zhu et al.* (2014) are employed. The following description summarizes the description in *Zhu et al.* (2014).

4.5.2.1 Distance metric

A metric that measures the normalized Euclidean distances between the Pareto front produced by the optimizer and the true Pareto front generated from enumeration is presented. The metric is the expected value of the normalized distance between each point, f_i , in the optimizer-produced Pareto front and the closest point in the enumeration Pareto front, d_k , calculated as follows:

$$\begin{aligned} Score &= E(d_{min_i}) \\ d_{min_i} &= min_k \sqrt{\sum_{j=1}^{J} a_j^2} \\ a_j &= \frac{d_{k,j} - f_{i,j}}{d_{k,j}} \end{aligned} \tag{4.17}$$

where J is the number of objectives, there are i points in the surrogate produced Pareto front and k points in the exact Pareto front.

4.5.2.2 Error metric

The second metric is to evaluate the quality of the Kriging models in terms of prediction accuracy. The error metric calculates the mean-square error between the VFO-scaled low-fidelity objective function value and the high-fidelity function value. The metric is computed as:

$$MSE_{VFO} = \frac{1}{N} \sum \sqrt{f_h(x)^2 - (f_g(x) f_f(x)^2} \tag{4.18}$$

All the N points in the final Pareto front are calculated regardless of the relative error returned by the Kriging models.

4.5.2.3 Crowding distance

The crowding distance metric is a density estimation described in *Deb et al.* (2002). The average crowding distance of all the points in the Pareto front is used to track whether the Pareto-optimal solutions are evenly distributed. The metric would indicate whether the new algorithm causes undesirable clumping Pareto points.

4.5.2.4 Span of Pareto front

The last metric shows the span of the Pareto front. It originates from a comparison metric proposed by *Lee et al.* (2005), and is adapted for the present work where a reference Pareto front is available. The metric is expressed as follow:

$$span = \prod_{j=1}^{J} \left| \frac{f_{j_{MAX}} - f_{j_{MIN}}}{d_{j_{MAX}} - d_{j_{MIN}}} \right| \tag{4.19}$$

Within this expression, f is the maximum or minimum function evaluation for j^{th} objective function in the Pareto front, and d is the corresponding maximum or minimum function evaluation from the enumeration solutions.

4.5.2.5 Enumeration results

Following the enumeration procedures described before, the optimizer generates two Pareto fronts with respect to the Paik strength method and Faulkner strength method. The total number of individuals that compose the approximated Pareto front with Paik strength method is 18,459, while about half as many points remain for the Faulkner front, totaling 8,852. The three-objective enumeration results are generally similar for both objective functions, and are shown in Figure 4.5 and 4.6.

The results do not form a true surface; instead, a single line through the weight, cost, and strength space is formed for each integral number of stiffeners. The space between these lines appears unreachable. The shapes between these lines are similar

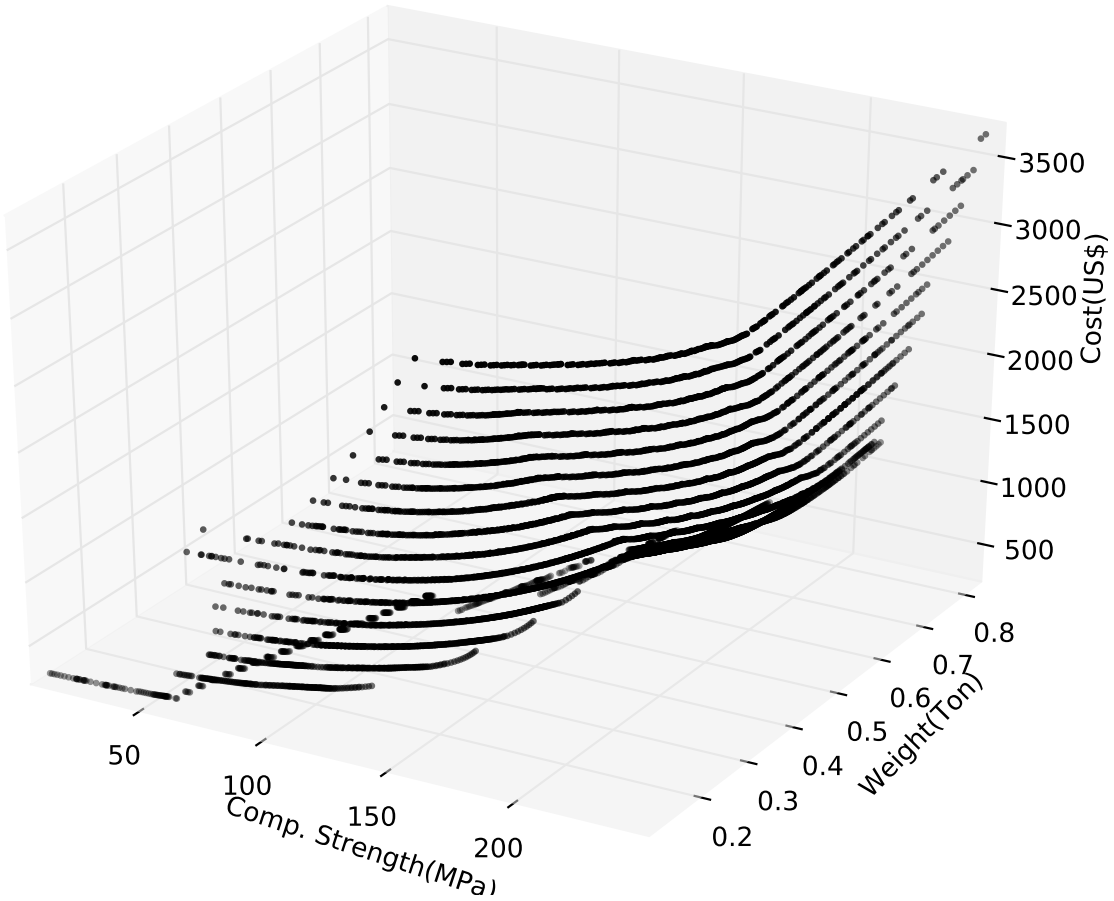


Figure 4.5: Enumeration results using high-fidelity Faulkner strength function.

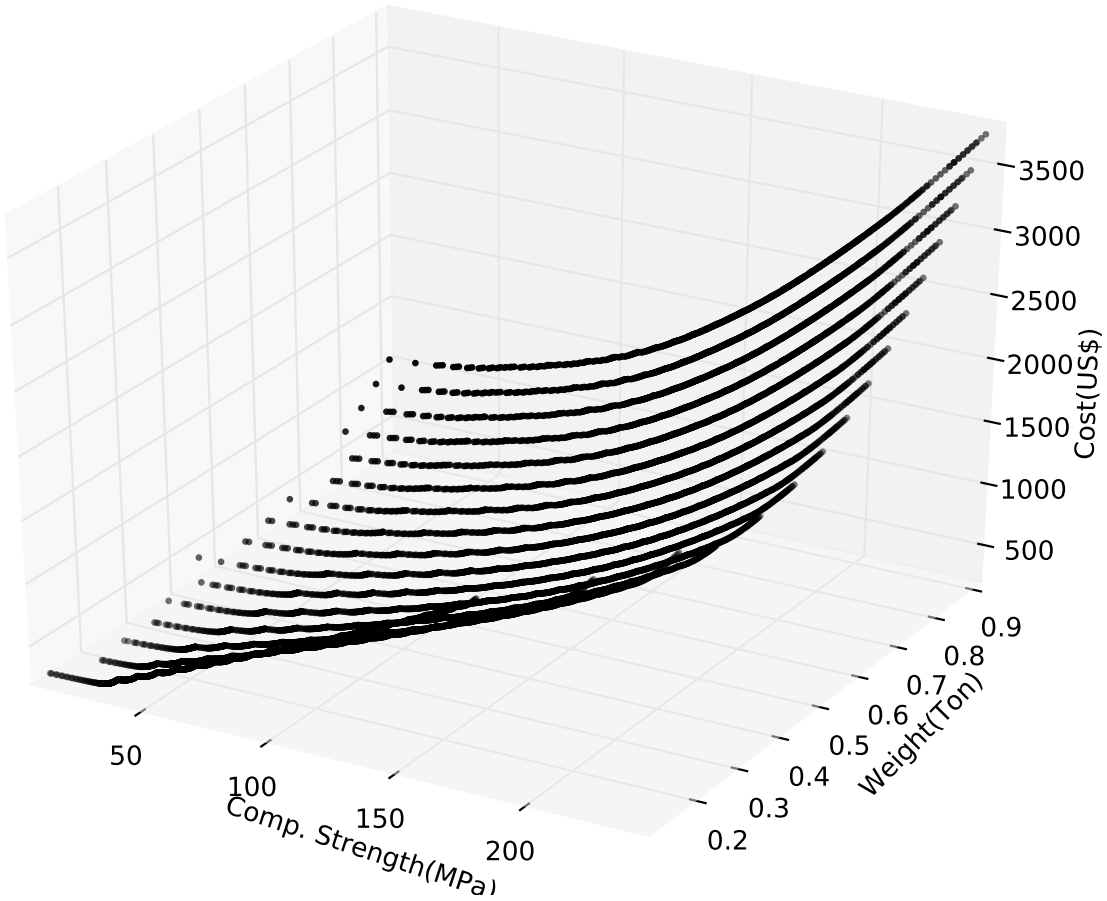


Figure 4.6: Enumeration results using low-fidelity Paik strength function.

between the methods, with slightly a little more complexity and roughness in the Faulkner front, probably owing to the more complex objective function formulation with more discrete failure modes than the Paik formulation. The enumeration results were saved and used as a reference for the distance metric described in the previous section.

4.5.3 Variable fidelity optimization results

In this section, the stiffener panel design problem was optimized using both *Zhu et al.* (2014)'s VFO approach with single Kriging model, and the proposed VFO using clustered multiple Kriging models respectively. To establish a more reasonable reference, the optimizer parameters were set at 130 individuals and 140 generations. The three different updating strategies which had been proved to be the most successful strategies in the standard VFO work (*Zhu et al.*, 2014) were used to examine the three-objective optimization problem.

Due to the stochastic nature of the genetic algorithm, every case was run for 25 replicates using different pseudo-random number generator seeds. It was expected that the influence of the probabilistic selection, crossover, and mutation in the genetic algorithm could be smoothed out using this approach. The four comparison metrics presented in previous section were calculated for each replicate. Then, the average and standard deviation of these metrics were computed for the 25 replicates. The average value was used to estimate the performance of the VFO approach in each case.

4.5.3.1 Single Kriging based VFO method

Zhu et al. (2014)'s VFO approach using one single Kriging surrogate model was studied first as a baseline. Three cases were chosen to represent each of the updating strategies. In addition to these three cases, the optimization results using only high-

Table 4.6: VFO updates schedules

Case	Offset	Spacing	Density	No.HFC
1. All high fidelity	-	-	-	18,200
2. Standard update	30	10	60	660
3. Dense update	30	20	120	660
4. Late update	50	5	60	1080
5. All low fidelity	-	-	-	0

Table 4.7: VFO Updating Metrics from standard VFO method

Case	Distance		Obj. error		Span		Crowding dist.	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
1	0.033	0.007	1.00	0	0.988	0.031	0.047	0
2	0.079	0.017	1.03	0.023	0.577	0.047	0.047	0
3	0.087	0.018	1.15	0.107	0.571	0.055	0.046	0
4	0.077	0.012	1.02	0.109	0.547	0.046	0.047	0
5	0.112	0.010	1.36	0.065	1.097	0.002	0.047	0

fidelity functions and only low-fidelity functions are listed for reference. All the cases are shown in Table 4.6, along with the associate number of high-fidelity function calls required.

The results are shown in Table 4.7. It can be seen that the distance metrics have moved only slightly from low-fidelity results. This shows in a three-objective optimization problem, the high-fidelity and low-fidelity strength objective functions differ strongly in the Pareto front. This could be seen in the objective function error for the all low-fidelity method, which has a relatively larger bias value of 1.36 compared to the two-objective low-fidelity result in *Zhu et al.* (2014). Based on the crowding distance metric results, the VFO results maintain a diverse solutions set in the Pareto front; however, the span metrics were all reduced in VFO solutions. This reduction in the VFO approach is due to the VFO fronts excluding the high-weight, high-strength panels.

Table 4.8: VFO updating metrics from proposed method

Case	Distance		Obj. error		Span		Crowding dist.	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
1	0.033	0.007	1.00	0	0.988	0.031	0.047	0
2	0.077	0.014	1.02	0.016	0.585	0.065	0.047	0
3	0.086	0.018	1.15	0.107	0.571	0.056	0.047	0
4	**0.066**	0.011	1.01	0.006	0.574	0.051	0.047	0
5	0.112	0.010	1.36	0.065	1.097	0.002	0.047	0

4.5.3.2 Clustered multiple Kriging based VFO method

All of the cases in Table 4.6 were also optimized in the proposed VFO scheme using clustered multiple Kriging models. The results are shown in Table 4.8 below.

There are several noticeable differences compared to results in the previous section. With the application of the multiple Kriging surrogates strategy, some improvements in the prediction error metrics are seen, particularly for Case 4. This means that the proposed VFO scheme performed better in approximating the high-fidelity function, using low-fidelity function combined with multiple Kriging models. The span and crowding distance metrics show similar results to the original method.

Even though it has been shown that the amount of VFO updating effort used was insufficient in converging closely to the high-fidelity Pareto front, the new method has improved the distance metric and pushed the VFO results towards the high-fidelity Pareto front. The most obvious improvement occurred in Case 4, which has the largest number of high-fidelity function calls, where the distance metric improves 15% with the same amount of high fidelity simulations. This is supportive of the idea that the multiple surrogates are advantageous for large surrogate data set sizes.

4.5.3.3 Statistical test on distance metric

By comparing the four metrics in Table 4.7 and Table 4.8, the most obvious improvement happens in the distance metric, the other three metrics do not show

Table 4.9: Case statistics of distance metric

Case	Single Kriging		Multiple Kriging		N	t
	Mean	St.Dev.	Mean	St.Dev.		
2	0.079	0.017	0.077	0.014	25	0.45
3	0.087	0.018	0.086	0.018	25	0.19
4	0.077	0.012	0.066	0.011	25	**3.38**

significant changes. In order to examine the impact of proposed multiple kriging method on the improvement of convergence, a one-tail statistical test was conducted to compare the distance metric value obtained in Single Kriging based VFO method and Clustered multiple Kriging based VFO method. In this case, the null hypothesis is the difference between the two algorithms on the convergence of Pareto front is 0, while the alternative hypothesis is that the proposed method did have a statistically significant improvement effect on the distance metric. The formula for T-test is listed as follows:

$$t = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{var_1/n + var_2/n}} \tag{4.20}$$

By substituting the values from Table 4.7 and Table 4.8, the calculated t values are shown in the Table 4.9.

As we conducted 25 replicates for each VFO updating case, the total group size is 50, so the degree of freedom is sample size minus 2, which is 48. By looking up a t table with 48 degrees of freedom, the threshold value for alpha = 0.005 is 2.682, and for alpha = 0.01 the table value is 2.407.

Therefore it is obvious that for Case 4 the t-statistic reached beyond the threshold of statistical significance, and we can reject the null hypothesis and conclude that the new algorithm improved the distance metric and helped converging more closely to the true Pareto front than previous method. As for Case 2 and Case 3, it is likely that the high fidelity function calls in these two VFO updating schedule were insufficient to reveal the difference between the two algorithms. This indicates that the new

Table 4.10: Average wall clock time for one complete run of algorithm over 25 replicates (sec)

Case	Offset spacing density	VFO using single Kriging	VFO using multiple Kriging
2	30/10/60	2677	**826**
3	30/20/120	3483	**852**
4	50/5/60	5138	**1106**

clustering Kriging model method performs at least as well, if not better, than the previous single Kriging model method in terms of convergence to the Pareto front. In the following, the VFO updating effort will be increased to further study its effect on convergence.

4.5.3.4 Computational efficiency study

In achieving more converged Pareto optimal solutions, the multiple surrogate management strategy performed significantly better than single surrogate method. A quantitative comparison of the wall-clock time taken by each method was completed. In comparison of these two different VFO strategies applied in NSGA-II, it is clear that the proposed method outperformed the original method regarding optimization efficiency.

In this work, a profiler code was employed to record the run time of the entire optimization program. The project was sent to the University of Michigans Flux cluster which provided a high performance computing environment. The VFO cases listed in Table 4.6 were analyzed in computational wall time consumption using both methods. The results are shown in Table 4.10. The wall time for one seed listed in the Table 4.10 was based on the average time for 25 replicates in each case.

Due to the new Kriging surrogate model construction approach, the proposed VFO method showed a significant amount of computational time saving, as compared to the original method. It is because of the newly introduced clustering strategy, the new VFO method is capable to manage multiple surrogates in the design space and avoid

Table 4.11: Extended VFO run of case o50s5d60 using multiple surrogates

Case	Generations	High fidelity runs	Distance	Obj. error	Wall time used (s)
50/5/60	140	1080	**0.066**	1.009	1140
50/5/60	280	2760	**0.050**	1.007	4905
50/5/60	350	3600	**0.048**	1.006	8112

solving large Kriging models. These merits helped to result in a large computational efficiency improvement for multiple objective design optimizations. The tractability of the large sampling size surrogate modeling problem is successfully addressed with the clustered VFO method. Now the VFO methodology is equipped with better surrogate capability to converting the Pareto front to the high-fidelity front.

4.5.3.5 Increase the Kriging sample size

As noticed in Table 4.7 and Table 4.8, the VFO results stayed in the middle of high-fidelity result and low-fidelity result. More effort is required to help the VFO converge to the high-fidelity solutions; however, in doing so, the original single Kriging model strategy becomes a limiting factor. As the number of Kriging sample points increases, the single Kriging model becomes time-consuming and difficult to reliably solve. In this investigation, the number of generations was increased in the optimizer runs. The specific Case 4 (o50s5d60) was chosen with generations extended from 140 to 280 and 350. The test runs are summarized in Table 4.11.

The test results verified that the single Kriging struggles in large surrogate size situations. For the single Kriging model, the VFO takes too much computational time and is deemed not practical. The single Kriging model takes more time to complete 140 generations than the multiple does to complete 280 generations. In contrast, the proposed multiple Kriging strategy functioned well in these large problems, including cases with up to 350 generations as shown in Table 4.11.

Table 4.11 demonstrates that the clustering Kriging model can significantly further

reduce the distance metric. This indicates that the problem can be addressed by the VFO method using the new Kriging construction method. In the 280 generations case, the number of surrogate models used increases to 28 in the end, and about 37 surrogate models are used in the 350 generation case. The multiple surrogate strategy works well as better convergence is observed. The prediction error metrics using the new method also show promising improvement, indicating that the Kriging predictions are accurate enough to correct the low-fidelity model to high-fidelity model in the Pareto front region. Most importantly, the distance metric shows significant improvement. As can be seen in Table 11, the VFO is converging from the all low-fidelity Pareto front with a distance metric of 0.112 to 0.048, which is close to the all high-fidelity Pareto front with a distance metric of 0.033. These results are obtained using only 20% of total HFC effort of a non-VFO multi-objective genetic algorithm run, which is described as Case 1 in Table 4.6. For each point found along the Pareto front, only 27.7 high-fidelity calls are needed, while Case 1 needs 140 high-fidelity calls per point to generate the Pareto front.

4.6 Summary

This chapter explored further on surrogate-assisted multi-objective evolutionary optimization algorithm development. Surrogate modeling is developed in response to the various fidelity tools in structural simulations. A novel variable fidelity optimization (VFO) is presented where multiple surrogate models are built and managed on-line to improve the efficiency of the computational expensive multi-objective evolutionary optimizations.

A challenge of existing VFO approaches using a single surrogate is that as the optimization proceeds and the surrogate model size increases, the difficulty in solving the Kriging surrogate also increases. This significantly impacts the speed and numerical stability of the VFO method. A novel variable fidelity optimization scheme using

multiple local Kriging surrogate models are developed to elevate the performance of VFO. A clustering algorithm is employed to partition the Kriging sample points set into local, small sub-surrogates. In this way, the numerical difficulties in building large Kriging models can be avoided. Its applicability was tested on six standard two-objective test problems and a three-objective stiffened panel deisgn problem.

The clustered Kriging model VFO method proposed performed strongly on all problem types. On the FON, POL, and QV test problems it was able to successfully resolve complex Pareto fronts guided by approximate, low-fidelity versions of the Pareto front. On the ZDT1, ZDT2, and ZDT3 problems the method again found complex Pareto fronts. Additionally, the method was shown to be superior to the earlier single Kriging model version of the VFO algorithm in terms of accuracy and distribution of points on the Pareto font. For the complex three-objective stiffened panel problem, this improved performance was quantified. The method proposed was shown statistically to equal or exceed the single-Kriging model in terms of distance of found points from the Pareto front, error in objective function evaluations, span of the retuned Pareto front, and crowding distance along the Pareto front. This was achieved via Welchs T-Test over 25 replicates of each algorithm. Additionally, the revised method took only 21%-31% of the runtime of the single Kriging model method. Finally, the clustered Kriging model was computationally efficient and numerically stable enough to be run longer than the previous single Kriging model method, allowing a more accurate Pareto front to be evolved. It is believed that the proposed multiple Kriging modeling strategy enriched the on-line surrogate modeling techniques to better addresss structural design problems in multi-objective optimizations.

CHAPTER V

Efficient reliability-constrained structural design optimization including interval uncertainty

5.1 Introduction

Managing uncertainty, especially in the early stages of structural design, is critical to many novel designs and materials. While robust design and reliability-based optimization frameworks have been successfully developed (see e.g. *Richardson et al.* (2015); *Wang et al.* (2015); *Hardin et al.* (2015); *Meng et al.* (2015)) for recent work in this field), most of these formulations require a precise stochastic definition of the uncertainty involved. However, early stage structural design with novel materials or applications is marked by comparatively limited and vague information about the design and associated uncertainties. Uncertainty associated with limited knowledge is epistemic in nature and is normally reducible by investment in engineering investigation. Consequently, traditional reliability-based design tools may not be well suited to model the design situation as the epistemic uncertainty normally cannot be stochastically defined. This chapter presents the research work that enables interval uncertainty analysis to unveil the impact of epistemic uncertainty in Reliability based marine structural design optimization. The contents presented within this chapter are under consideration for publication in *Liu et al.* (2016). Dr. Hankoo Jeong pro-

vided the composite structural analysis tool for a case demonstration of the proposed framework.

Reliability analysis with non-probabilistic interval uncertainty (*Du*, 2007; *Jiang et al.*, 2011) has been studied as an approach to modeling epistemic uncertainty. While this removes the need to define a precise stochastic distribution, reasonable uncertainty bounds still must be developed. The present approach combines reliability-based design optimization and interval uncertainty. Then, the width of the uncertainty intervals are treated as an objective in a multi-objective optimization approach while maintaining consistent reliability levels. The resulting Pareto fronts show the impact of lack of knowledge on design performance and allow engineering design teams to prioritize where to invest time in reducing epistemic uncertainty. The core contribution required to make such an approach practical is a novel adaptive surrogate modeling technique for efficient interval reliability analysis. This surrogate approach allows the combination of reliability models, interval uncertainty, and multi-objective optimization to remain computationally feasible.

The central concept of the proposed framework is that interval uncertainty is a useful representation of uncertainty in early-stage design knowledge. Conventional structural optimization approaches (*Jin and Branke*, 2005) often use stochastic form to account for uncertainty in the design. However, in the field of marine structures, such models are difficult to apply early in the design process as precise stochastic uncertainty information does not yet exist. Any error in the assumption of the distribution can be harmful later in the design (*Ben-Haim and Elishakoff*, 2013). This is especially true for marine structures, where too low early structural weight estimates can cause extensive design re-work or in-service structural failures (*Keane*, 2012). It is argued that interval uncertainty can be used to address this concern where no assumption of distribution is needed (*Liu and Collette*, 2015). Among other non-traditional uncertainty models (*Möller and Beer*, 2008), interval uncertainty was selected as in

most cases the non-deterministic parameters and variables are only known within intervals (*Ferson et al.*, 2004). This work is the first to explore adaptive surrogate modeling to reveal the coupling between structural optimization with interval uncertainty while meeting reliability-based constraints. The interval uncertainty range will be treated as design variables, and the optimizer will determine feasible design configurations with respect to various uncertainty intervals. The aim is to resolve the trade-off between interval uncertainties and design performance through the optimization, thus unveiling the overall impact of early stage design uncertainty on the design.

A complication in selecting interval uncertainty is that working with intervals in optimization can be computationally expensive. In design optimization involving interval uncertainty, a max-min optimization (*Ong et al.*, 2006) is often needed. In min-max approaches, the optimizer searches for a solution that has the best worst-case performance in the uncertainty interval. In terms of reliability analysis involving interval uncertainty, the worst case of interval analysis needs to satisfy the specified reliability constraint. As reliability simulation itself normally involves a search for most probable point of failure (*Hasofer and Lind*, 1974), reliability analysis with interval uncertainty becomes a nested optimization in which the worst case performance needs to be located. Considering that such analysis is repeatedly requested in population-based metaheuristic design optimization approaches, the computation can soon become intractable. This chapter presents a method that is capable of locating the worst case reliability result while remaining computationally efficient.

Many previous authors have studied how to efficiently solve reliability-based design optimization (RBDO) problems. One way is focused on reliability problem formulation, where the performance measure approach (*Tu et al.*, 1999; *Youn et al.*, 2003), sequential RBDO (*Du and Chen*, 2004) and single-loop RBDO (*Liang et al.*, 2008) have been proposed. However, they are not ideal as solutions to interval uncertainty

problems where worst case reliability needs to be computed. A more direct way to reduce computational cost is to use surrogate models (*Kim and Choi*, 2008; *Kuczera and Mourelatos*, 2009; *Sudret*, 2012). Among the surrogate methods proposed, Kriging (*Sacks et al.*, 1989) shows promise as an approximation tool in reliability simulation. Kriging was first proposed for structural reliability problem by *Kaymaz* (2005), and recent development can be found in *Bichon et al.* (2008), *Echard et al.* (2011), and *Dubourg et al.* (2011). These studies on Kriging methods for reliability mainly focused on the approximation of the limit state function, and then applying Monte Carlo Simulation (MCS) method for reliability analysis. Such methodology has also been applied in solving reliability analysis with interval uncertainty problems (*Yang et al.*, 2014). While Kriging-assisted MCS methods may be a viable strategy for a single reliability analysis, adopting this methodology in population-based optimization algorithms, such as evolutionary algorithms, can be problematic, as the large number of reliability analyses required can quickly render MCS method extremely costly even with the help of Kriging.

This study introduced two new refinements to the interval uncertainty problem with reliability constraints first presented at the MARSTRUCT 2015 conference (*Liu and Collette*, 2015). First, a new online surrogate model construction technique is proposed, where the optimizer can dynamically refine an initial coarse surrogate model over the course of the investigation. Second, a quadratic approximation strategy is used to remove the innermost search on the surrogate model - that for worst-case reliability value in an interval. By coupling these strategies together, the computational burden of the proposed method is significantly reduced. Less time is spent building the surrogate model upfront. The quadratic approximation, when coupled with sequential refinement of the surrogate has proven reasonable in practice for cases where the worst-case performance is located in the interior of the interval search range, or at the boundary of the interval search range. The method proposed here can accurately

estimate worst-case reliability performance for multi-objective evolutionary algorithm with a high level of efficiency.

This chapter is organized as follows. Section 5.2 reviews the interval uncertainty and interval reliability analysis. Subsequently, Section 5.3 introduces the proposed surrogate modeling method and the multi-objective optimization framework for trade-off study. Next, the developed method is examined in an interval reliability benchmark problem (*Du*, 2007), and then the validated method is applied to CFRP top-hat stiffened panel design (*Maneepan et al.*, 2005). Discussion and concluding remarks are given in the end of the chapter.

5.2 Reliability analysis with interval uncertainty

This work treats uncertainties that are due to lack of information via an interval formulation. There are two critical components to such an approach: the definition of the interval model and the application of this model in reliability analysis. Each of these components is reviewed in turn in this section.

5.2.1 Interval uncertainty

Interval uncertainty provides an appropriate alternative from stochastic uncertainty, as no information regarding stochastic distribution is required. Interval modeling (*Moens and Vandepitte*, 2007) is usually applied in a simple closed form:

$$Y = [Y_l, Y_h] = \{Y \in I | Y_l \leq Y \leq Y_h\} \tag{5.1}$$

Interval uncertainty model is concerned with investigating the whole range of the potential values bounded by a higher bound and a lower bound. There is no assumption of probabilistic distribution within these bounds; specifically, a uniform distribution

is not assumed. The interval definition can also be interpreted in another form:

$$\forall Y \in [Y_0 - \Delta, Y_0 + \Delta] \tag{5.2}$$

where Δ is defined as the maximum deviation of uncertainty from nominal value, representing the range of the interval uncertainty. In this form, it is clear that interval uncertainty can reflect the tolerance for error in the design (*Jiang et al.*, 2015). The impact of interval ranges on the design performance is the aim of this work. This trade space can be studied by making a version of Δ one for the objectives of the optimization problem, as reviewed in Chapter III.

5.2.2 Worst-case reliability index with interval uncertainty

Without interval uncertainty, reliability analysis is concerned with calculating the probability of failure in a limit state function:

$$P_f = Pr(g(\mathbf{X}) \leq 0) = \int\limits_{g(\mathbf{X})\leq 0} f_{\mathbf{X}}(\mathbf{X}) d\mathbf{X} \tag{5.3}$$

where g denotes the limit state at which the system is safe if $g(\mathbf{X}) \geq 0$. $\mathbf{X}$ is a vector of random variables accounting for stochastic uncertainties, and $f_{\mathbf{X}}(\mathbf{X})$ is the joint probability density function. Required and achieved values of P_f are normally given in terms of the safety index, β, with a standard normal distribution Φ:

$$P_f = \Phi(-\beta) \tag{5.4}$$

As direct integration of the expression given in Equation 5.3 is difficult, many approximate methods of estimating β have been proposed. Here, for simplicity, the first order reliability method (FORM) *Hasofer and Lind* (1974) is used. In FORM, the

reliability index can be computed in the following procedure:

$$\begin{aligned} \beta &= \min_{U} \|U\| \\ \text{s.t. } & g(U) = 0 \end{aligned} \tag{5.5}$$

where U is the vector of transformed standard normal variables from random variables: $u_i = \Phi^{-1}(F_{x_i}(x_i))$. The reliability index can also be interpreted as the minimum distance from origin to limit state in U space.

Given that the information about stochastic uncertainty models are usually incomplete in the early stage of innovative marine structure design, the calculated reliability index may not be accurate enough to ensure sufficient safety in design (*Ben-Haim and Elishakoff*, 2013). Interval uncertainty can be introduced into reliability analysis to model the epistemic contribution to uncertainty and address this concern (*Du et al.*, 2005). Generally, when interval uncertainty parameter is involved in the limit state function, the interval reliability index output becomes an interval: $\beta \in [\beta_L, \beta_H]$. In this reliability index interval, the minimum and hence the worst-case reliability index β_L is used for safety examination.

Here, interval parameter Y is used to illustrate the process in computing β_L. After the transformation of random variables X to U space, the limit state function $g(U, Y) = 0$ is defined with both normal variables U and interval parameter Y. Though there is proposed work (*Jiang et al.*, 2012) conducting monotonicity analysis for reliability with interval parameter, most of the time an optimization search is needed to locate β_L in the interval reliability output. The procedure is defined with the equations below:

$$\textit{inner loop}: \begin{cases} \beta(Y) = \min\limits_{U} \|U\| \\ \text{s.t. } g(U, Y) = 0 \end{cases} \quad \textit{outer loop}: \begin{cases} \beta_L = \min\limits_{Y} \beta(Y) \\ \text{s.t. } Y_L \le Y \le Y_H \end{cases} \tag{5.6}$$

The above procedure is a nested optimization process where the outer loop minimizes the reliability index by locating the worst-case combination of interval variables Y^*, while the inner loop is a reliability determination via any standard reliability evaluation method. In this work, the FORM procedure defined in Equation 5.5, is used for the reliability index evaluation.

5.2.3 Impacts of interval reliability analysis

A reliability-based design optimization that accounts for worst-case reliability using FORM analysis can be expressed as follows:

$$\begin{aligned} &min\ f(\mu(\boldsymbol{X})) \\ &s.t.\ \beta_L = \min_{\boldsymbol{Y}} f_{FORM}(\boldsymbol{X}, \boldsymbol{Y}) \geq \beta_t \\ &\boldsymbol{Y} \in [\boldsymbol{Y}_0 - \Delta, \boldsymbol{Y}_0 + \Delta] \end{aligned} \tag{5.7}$$

where f is a performance function associated with design variables and β_t is the target reliability index constraint. f_{FORM} denotes the process by which the reliability index is evaluated. Additional constraints that do not involve Y or may be deterministic could also be added.

While most studies on interval reliability assumed fixed interval domain, this dissertation work argues that various range values Δ of interval uncertainty need to be examined. The range Δ reflects the epistemic uncertainty present. However, epistemic uncertainty is normally reducible by investing engineering time and expense to generate and analyze additional data. From the project management perspective, it is interesting to explore what the impact of epistemic uncertainty is on design performance and what performance gain could be achieved by reducing it. It is proposed that the width of Δ can be set as an objective in multi-objective optimization along with conventional objectives such as weight, cost etc. The resulting Pareto trade space will show the engineers the impact of epistemic uncertainty on the performance

of multi-objective optimization.

Computational efficiency is the primary issue that needs to be addressed in order for the proposed interval study. Due to the optimization search for worst-case combination of interval variables Y^*, the number of reliability analyses needed can increase significantly and makes the computation intractable for design optimization. This chapter proposes an advanced surrogate modeling technique that is capable of conducting the worst case search in interval domain and clearing the obstacle in proposed trade study on interval uncertainty.

5.3 Trade-off analysis with adaptive surrogate modeling

5.3.1 Overview

In this section, an efficient implementation of the interval reliability analysis is put forward in a multi-objective optimization framework. Since the design optimization cycle time is closely related to the number of reliability simulations in Equation 5.7, in this study, FORM reliability simulation is approximated by a proposed Kriging surrogate model in the optimization. Moreover, the surrogate method proposed here specifically address the problems where the worst-case search is needed in interval variable domain. In each surrogate prediction for the reliability performance of the individual, an estimated worst-case reliability index can also be provided simultaneously in the proposed method. Thus, a higher level of efficiency is achieved by removing the inner-loop search. The surrogate-assisted RBDO that accounts for worst-case reliability can be expressed as the following:

$$\begin{aligned} &min\ f(\mu(\boldsymbol{X})) \\ &s.t.\ \hat{\beta}_L(\boldsymbol{X}, \boldsymbol{Y}) = \hat{y}(\boldsymbol{X}, \boldsymbol{Y}^*) \geq \beta_t \end{aligned} \tag{5.8}$$

where $[\boldsymbol{X}, \boldsymbol{Y}^*]$ is the estimated worst-case scenario of design point $[\boldsymbol{X}, \boldsymbol{Y}]$ from the surrogate model. $\hat{y}$ means that the surrogate model is used to approximate FORM simulation. The Kriging theory and the derivation of worst-case estimator are presented in sequence. Afterwards, the multi-objective optimization for interval uncertainty trade study is introduced.

5.3.2 Kriging modeling

Kriging is a powerful surrogate model that has been widely used to approximate computationally expensive simulations. An in-depth Kriging theory can be found in works of Sacks *Sacks et al.* (1989) and Simpson *Simpson et al.* (2001). As a brief explanation, Kriging predicts the function value $\hat{y}$ at an unobserved point based on a set of sampled points through a realization of a regression model and a stochastic process:

$$\hat{y}(\boldsymbol{x}) = \boldsymbol{f}^T \boldsymbol{\beta} + z(\boldsymbol{x}) \tag{5.9}$$

where $\boldsymbol{f}$ is regression basis functions by user's choice and $\boldsymbol{\beta}$ are regressional coefficients. In this work, ordinary Kriging is used; thus, $\boldsymbol{f}$ is a vector of all 1.0 with length n_s, n_s is the number of initial sampled points. The stochastic process z is assumed to have zero mean and a covariance of:

$$E[z(x_i)z(x_j)] = \sigma^2 R(\theta, x_i, x_j) \tag{5.10}$$

where σ is the process variance and $R(\theta, x_i, x_j)$ is the correlation model. A commonly used Gaussian correlation model is adopted here:

$$R(\boldsymbol{\theta}, \boldsymbol{x_i}, \boldsymbol{x_j}) = exp(-\sum_{k=1}^{n_v} \theta_k(|x_k^i - x_k^j|^2) \tag{5.11}$$

where $\boldsymbol{\theta}$ is a correlation parameter vector that is found by optimizing a maximum likelihood function, n_v is the number of design variables. After that, the Kriging predictor for an unobserved point x can be expressed as the following:

$$\hat{y}(\boldsymbol{x}) = \boldsymbol{f}^T\boldsymbol{\beta} + \boldsymbol{r}^T(\boldsymbol{x})\boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{F}\boldsymbol{\beta}) \tag{5.12}$$

where $\boldsymbol{\beta}$ is computed by least square regression, the vector $\boldsymbol{r}$ measures the correlation between the prediction point $\boldsymbol{x}$ and the sampled points $[\boldsymbol{x_1}...\boldsymbol{x_m}]$.

The mean square error s^2 of the predictor can also be provided:

$$s^2(\boldsymbol{x}) = \sigma^2\left(1 + \boldsymbol{u}^T\left(\boldsymbol{F}^T\boldsymbol{R}^{-1}\boldsymbol{F}\right)^{-1}\boldsymbol{u} - \boldsymbol{r}^T\boldsymbol{R}^{-1}\boldsymbol{r}\right) \tag{5.13}$$

where $\boldsymbol{u} = \boldsymbol{F}^T\boldsymbol{R}^{-1}\boldsymbol{r} - \boldsymbol{f}$. As an indication of prediction quality, s^2 will be used later for adaptively improving the modeling.

From Equation 5.12 it follows that the gradient of the Kriging predictor can be derived as follows:

$$J_{\hat{y}} = J_f(\boldsymbol{x})^T\boldsymbol{\beta} + J_r(\boldsymbol{x})\boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{F}\boldsymbol{\beta}) \tag{5.14}$$

where J_f and J_r are the Jacobian of $\boldsymbol{f}$ and $\boldsymbol{r}$, respectively. J_f is $[O_{n\times 1}]$ as $\boldsymbol{f}$ is a vector of constants in ordinary Kriging.

The i^{th} row of the Hessian matrix of the Kriging predictor,

$$H_{i,:} = [\frac{\partial^2\hat{y}}{\partial x_i\partial x_1}, \frac{\partial^2\hat{y}}{\partial x_i\partial x_2}, ..., \frac{\partial^2\hat{y}}{\partial x_i\partial x_{n_v}}] \tag{5.15}$$

can also be derived as follows:

$$H_{i,:} = \frac{\partial J_r(\boldsymbol{x})^T}{\partial x_i}\boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{F}\boldsymbol{\beta}) \tag{5.16}$$

where:

$$\left(\frac{\partial J_r(\boldsymbol{x})}{\partial x_i}\right)_{mn} = \frac{\partial^2 R(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{x_m})}{\partial x_i \partial x_n}, m = 1, ..., n_s, n = 1, ..., n_v \tag{5.17}$$

These properties will be used to accelerate the worst-case performance prediction search, as explained in next section.

5.3.3 Worst-case prediction and adaptive refinement

In this study, the Kriging surrogate model is used as a worst-case estimator to replace the optimization search that is conventionally required in locating the worst case $\boldsymbol{Y}^*$. $\boldsymbol{Y}^*$ is expressed as follows:

$$\boldsymbol{Y}^* = \arg\min_{\boldsymbol{Y} \in \boldsymbol{I}} f(\boldsymbol{X}, \boldsymbol{Y}) \tag{5.18}$$

where $\boldsymbol{I} = [\boldsymbol{Y}^L, \boldsymbol{Y}^U]$ is the bounded interval variables domain. In the proposed surrogate method, $\boldsymbol{Y}^*$ is estimated in a single Kriging evaluation and therefore avoids an additional nested optimization search.

For every candidate point D that has n design variables and m interval variables $D = [X_1, ..., X_n, Y_1, ..., Y_m]$, the predicted response $\hat{y}_D$, Jacobian matrix J_D, and Hessian matrix H_D, can be provided by the surrogate. Within J_D and H_D, the components associated with interval variables $\boldsymbol{Y}$ will be sorted out and assembled to form local Jacobian and Hessian matrices for interval domain:

$$\begin{aligned} J_Y &= [\frac{\partial \hat{y}}{\partial Y_1}, ..., \frac{\partial \hat{y}}{\partial Y_m}]^T \\ H_Y &= \begin{bmatrix} \frac{\partial^2 \hat{y}}{\partial Y_1^2} & \cdots & \frac{\partial^2 \hat{y}}{\partial Y_1 \partial Y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \hat{y}}{\partial Y_m \partial Y_1} & \cdots & \frac{\partial^2 \hat{y}}{\partial Y_m^2} \end{bmatrix} \end{aligned} \tag{5.19}$$

Afterwards, a quadratic model for the interval domain can be established using

J_Y, H_Y and $\hat{y}_D$:

$$\hat{y}(\boldsymbol{Y} + \boldsymbol{s_Y}) \approx \frac{1}{2}\boldsymbol{s_Y}^T H_Y \boldsymbol{s_Y} + J_Y \boldsymbol{s_Y} + \hat{y}_D \tag{5.20}$$

where $\boldsymbol{s_Y}$ is called the Newton step to the minimum of the quadratic, which is the estimate of worst-case performance. $\boldsymbol{s_Y}$ can be obtained by:

$$\boldsymbol{s_Y} = -H_Y^{-1} J_Y \tag{5.21}$$

In many engineering applications, performance may degrade as uncertainty increases. In this case, the worse case performance will locate the maximum possible uncertainty. However, in the quadratic approximation proposed, this will result in the Hessian being negative, and the Newton step approach will not function. In implementing the algorithm, this particular situation is caught by program logic, and the Jacobian used to determine which interval boundary should be returned as the worse-case performance. The algorithm has been tested on problems where the worse performance falls in the middle and at the boundary of the interval, and in both cases the algorithm has proven robust.

Assuming the quadratic approximation implied by this approach matches the underlying Kriging model, the worst case can be directly determined as $\boldsymbol{Y}^* = \boldsymbol{Y} + \boldsymbol{s_Y}$. This method is much more efficient for the worst-case search as it eliminates the inner optimization run originally required in Equation 5.18.

However, there are two levels of nested assumptions in this approach. First, it is assumed that the quadratic approximation of the Kriging surface is accurate, and the Hessian is positive. Second, it is assumed that the Kriging model is a faithful model of the underlying reliability simulations. To ensure that these assumptions are both reasonable, two online refinement criteria are used. These criteria are evaluated every time a worst-case performance prediction is made, and if either is violated, an updated prediction is generated. This scheme will ensure that an accurate worst

case performance prediction is provided for optimization while refining the Kriging model to increase the quality of the surrogate model as the optimization runs. The refinement criteria are as follows:

1. Based on the initial predicted worst-case design $D^{(1)}_{worst}$, a repeated worst-case prediction $D^{(2)}_{worst}$ is made reestablishing the quadratic approximation at the previous worst-case point $D^{(1)}_{worst}$. The two outcomes of these two searches need to be sufficiently close;

2. The U metric proposed in *Echard et al.* (2011) for Kriging model accuracy is used here:

$$U = \frac{\hat{y}(D_{worst}) - \beta_t}{s(D_{worst})} \tag{5.22}$$

   where s is the Kriging variation derived in Equation 5.13. This metric compares the distance from the reliability constraint boundary to the predicted error in the Kriging model. Points close to a constraint boundary or with large prediction errors would be highlighted for refinement. A minimum value of 3.0 is required for U metric in this study.

The first criterion ensures that the worst-case prediction is stable by an iterated prediction check. The second criterion establishes a lower confidence bound (*Cox and John*, 1997) regarding the target value β_t. The U metric utilizes the Kriging variation information to ensure the accuracy of prediction. A minimum value of 3.0 suggests that the probability of making a mistake on $\hat{y} \geq \beta_t$ is $\Phi(-3) = 0.135\%$. In each worst-case estimation, these two criteria will be examined, and any violation will be used as a guidance to update the surrogate model. The procedure is outlined in Algorithm 3.

Algorithm 3 Worst case prediction and updating scheme.

BEGIN
Initialize: Train an initial Kriging model
for Candidate design $D_i = [\boldsymbol{X}, \boldsymbol{Y}]$ **do**
  Predict the worst case performance $D^{(1)}_{worst} = [\boldsymbol{X}, \boldsymbol{Y^*}]$of D_i;
  Repeat the process by predicting $D^{(2)}_{worst}$ of $D^{(1)}_{worst}$;
  Compute the U metric, and compare two predicted worst cases;
  if $|\hat{y}(D^{(2)}_{worst}) - \hat{y}(D^{(1)}_{worst})| \leq 10^{-3}$ *and* $U \geq 3.0$ **then**
    Do not update Kriging
  else
    Evaluate $\beta(D_i)$ using exact reliability analysis
    Update Kriging model
    Estimate D_{worst} again
  end if
end for
Set reliability index of D_i as $\hat{y}(D_{worst})$
END

5.3.4 Multi-objective optimization with interval variables

The goal of the proposed method is to explore the impact of reducing the width Δ of interval uncertainties on design performance. The resulting trade space will help prioritize areas for engineering investment to reduce epistemic uncertainty. With the developed worst-case estimation technique, this trade-off study of interval uncertainty can be resolved in a multi-objective optimization framework. This work proposes to couple interval uncertainty into structural design optimization by treating interval widths Δ as design variables, and an interval reduction metric as objective function. Some interval reduction measures can be found in the work of *Li et al.* (2009). Here, a simple inverse metric of interval variable ranges production is used as an indication of the cost needed to gain relevant information and reduce the interval uncertainty:

$$Cost = \frac{1}{\prod_{i=1}^{I} \Delta_i} \tag{5.23}$$

The interval reduction cost function and structural performance function will be optimized within NSGA-II - a multi-objective genetic algorithm optimizer proposed

by *Deb et al.* (2002). The NSGA-II algorithm is an elitist non-dominated sorting genetic algorithm that has proven to be powerful in capturing a set of diversified Pareto optimal solutions. In NSGA-II, children population is created from parent chromosomes by both crossover and mutation. The SBX method is adopted here for crossover with an exponent of 4.0. The probability of crossover operation is 0.8. The random mutation rate is set at a low probability of occurrence of 0.1%, exponent rate of mutation operator is 4.0. These genetic algorithm operators have proved to be robust in finding the Pareto front in the previous investigation of multi-objective optimization.

Relying on the proposed adaptive surrogate modeling strategy to estimate worst-case reliability, the optimizer can quickly evaluate the adequacy of structural design with different uncertainty interval widths, and thus keeps the interval computation tractable. The surrogate-assisted optimization process is summarized in Algorithm 4 below.

Algorithm 4 Adaptive surrogate-assisted MOGA with interval variables.

BEGIN
Initialize: Latin Hypercube Sampling method to collect initial sample points, evaluate sampled points using FORM and construct a Kriging surrogate model; Initialize NSGA-II.
while Termination criteria of GA is not met **do**
  for Individual D_i in population **do**
    Evaluate fitness values of D_i using objective functions $[f, Cost]$;
    Estimate the worst case reliability index $\hat{\beta}_L$ of D_i by surrogate
    Check updating criteria described in **Algorithm 3**
    if Status is updating Kriging **then**
      Evaluate $\beta(D_i)$ using exact reliability analysis
      Updating sampling database
      Refine Kriging model and estimate $\hat{\beta}_L$ again
    end if
    Set constraint violation as $max[0, \beta_t - \hat{\beta}_L]$
  end for
  Apply GA operators to create the next generation of population
end while
END

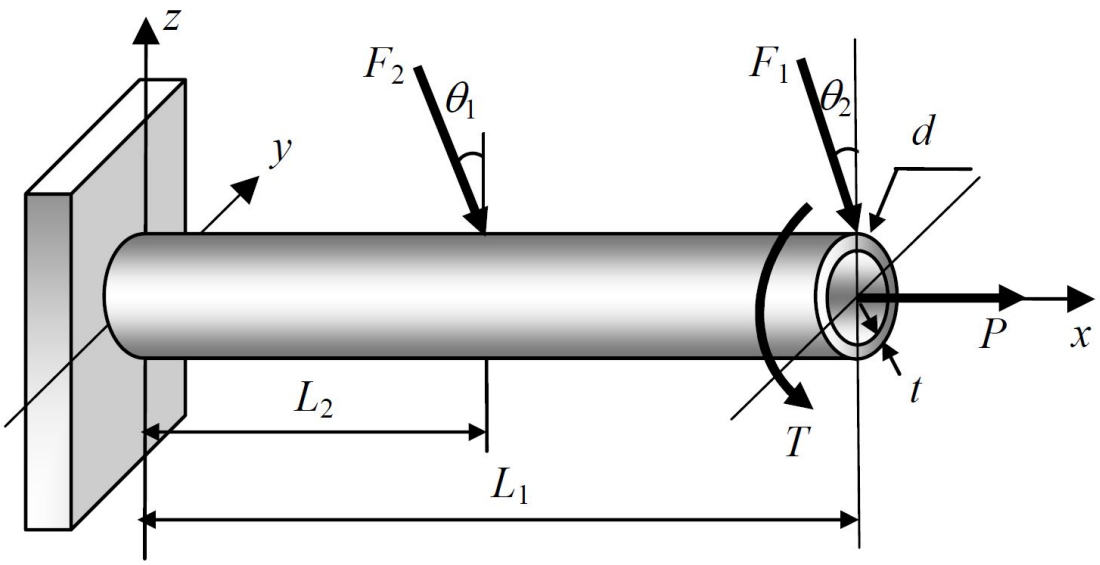


Figure 5.1: Cantilever tube problem (*Du*, 2007).

5.4 Interval reliability benchmark design problem

A cantilever tube problem proposed in *Du* (2007) for interval reliability study is revised here to test the framework. The cantilever tube shown in Figure 5.1 is subjected to external forces F_1, F_2, P and torsion T. The limit state function is expressed as follows:

$$g(\boldsymbol{X}, \boldsymbol{Y}) = S_y - \sigma_{max} + w \tag{5.24}$$

where S_y is the yield strength, and σ_{max} is the maximum von Mises stress given by:

$$\sigma_{max} = \sqrt{\sigma_x^2 + 3\tau_{xy}^2} \tag{5.25}$$

While w is added Gaussian noise ($w \sim \mathcal{N}(0, 0.001)$) to simulate a more complex simulation-based (e.g. implicit) limit state function than the simple equations used here. Such noise contamination was added to assess if the method would be robust if extended to more complex limit states.

Table 5.1: Random variables of the cantilever tube problem

Variables	Parameter 1	Parameter 2	Distribution
$X_1(t)$	5mm	0.1mm	Normal
$X_2(d)$	42mm	0.5mm	Normal
$X_3(L_1)$	119.75mm	120.25mm	Uniform
$X_4(L_2)$	59.75mm	60.25mm	Uniform
$X_5(F_1)$	3.0kN	0.3kN	Normal
$X_6(F_2)$	3.0kN	0.3kN	Normal
$X_7(P)$	12.0kN	1.2kN	Gumbel
$X_8(T)$	90.0Nm	9.0Nm	Normal
$X_9(S_y)$	220.0Nm	22.0Nm	Normal

*:For uniform distributions, Parameter 1 is the low bound, Parameter 2 is the upper bound. For other distributions, Parameter 1 is the mean value, Parameter 2 is standard deviation.

Table 5.2: Interval variables of the cantilever tube problem

Variables	Intervals
$Y_1(\theta_1)$	$[0°, 10°]$
$Y_2(\theta_2)$	$[5°, 15°]$

The calculation of stress is determined from classical structural mechanics:

$$\begin{aligned} \sigma_x &= \frac{P + F_1 \sin\theta_1 + F_2 \sin\theta_2}{A} + \frac{Mc}{I} \\ M &= F_1 L_1 \cos\theta_1 + F_2 L_2 \cos\theta_2 \\ A &= \frac{\pi}{4}[d^2 - (d-2t)^2],\ c = d/2, \\ I &= \frac{\pi}{64}[d^4 - (d-2t)^4],\ \tau_{xy} = \frac{Td}{4I}. \end{aligned} \tag{5.26}$$

In the computation of the reliability index, the random variables settings are given in Table 5.1, and interval variables settings are given in Table 5.2. Note that the interval variable ranges are fixed in the original problem (*Du*, 2007) for a single analysis. In this work, the test multi-objective optimization problem with varying

Table 5.3: Design variables of the cantilever tube problem

Design variables	Description	Lower bound	Upper bound
t	Mean value of thickness X_1	3mm	6mm
d	Mean value of diameter X_2	38mm	44mm
Δ_1	Interval variable range of Y_1	0	10
Δ_2	Interval variable range of Y_2	0	10

interval ranges is formulated as follows:

$$\begin{aligned} &min\ Volume(t,d) = \pi(d-t)tL_1 \\ &min\ Cost(\Delta_1, \Delta_2) \\ &s.t.\ \hat{\beta}_L(\boldsymbol{X}, \boldsymbol{Y}) = \hat{y}(\boldsymbol{X}, \boldsymbol{Y}^*) \geq \beta_t \\ &where\ [t,d] = [\mu(\boldsymbol{X_1}), \mu(\boldsymbol{X_2})] \\ &\boldsymbol{Y_0} \leq \boldsymbol{Y} \leq \boldsymbol{Y_0} + \boldsymbol{\Delta} \end{aligned} \tag{5.27}$$

In this optimization, the mean value of thickness and diameter are set as design variables along with two interval variable ranges. The design variable domain is defined in Table 5.3. The interval uncertainty lengths Δ_1, Δ_2 range from 0 to 10, where zero ranges means the interval variables are reduced to the deterministic values: $[Y_1, Y_2] = [0°, 5°]$, and the largest ranges mean the interval variables cover the same range shown in Table 5.2. In conducting the analysis, the initial surrogate model is constructed from 150 points sampled by the Latin Hyper-square Method evenly in the design space. Reliability simulations are conducted using the PyRe (python Reliability) module version 5.0.2 (*Hackl*, 2013).

To validate the proposed method, initial, the worst reliability in the interval range is located by Algorithm 3 directly (e.g. no optimization). This computation is also performed by *Du* (2007) which allows for comparison with the current method. The design point $[t, d, \Delta_1, \Delta_2] = [5.0, 42.0, 10, 10]$ is chosen as the prediction point, which means the random variables X and interval variables Y are the same as Tables 5.1 and 5.2 in the literature. The predicted worst-case interval variable combination is

Table 5.4: Determination of worst-case reliability for the cantilever tube problem

Worst case reliability	Proposed surrogate prediction	Du *Du* (2007)
$p_f^{max}(=\Phi^{-1}(-\beta_L))$	1.63×10^{-4}	1.63×10^{-4}
function evaluations of g	1092*	147

*:Note this surrogate was built for optimization, hence it contains more function calls than Du's single-value computation.

$Y^* = [3.9203, 7.8132]$ and the associated reliability index is 3.5964. A comparison of the result with Du's method is shown in Table 5.4. It is clear that the worst-case prediction successfully estimated the worst-case design compared to nonlinear optimization results from Du's study. The computational efficiency of the surrogate model is also promising, the 1092 limit state function evaluations are made for 150 reliability analyses during the sampling stage. Though this is higher than the 147 function calls used by Du, the surrogate is now ready for optimization while Du's work only finds a single value.

Next, the test problem stated in Equation 5.27 is solved using the proposed multi-objective optimization framework in Algorithm 4. As a reference, the test problem is also optimized within NSGA-II with all the reliability analysis conducted directly via FORM simulation. The two optimizations both used 100 generations and a population size of 100, and a target reliability index constraint of 3.0, these parameters were chosen for experimental study here.

Due to the stochastic nature of the NSGA-II algorithm, the optimization problem was run 10 replicates using different pseudo-random number generator seeds. All the Pareto fronts generated from the optimizations are compared with the Pareto front from exact reliability analysis.

On visual inspection, all the 10 Pareto fronts generated from the proposed method fall very close to the exact solution. In some cases, the optimization process struggled to resolve the entire span of the front, however, the portion of the front found was always accurate. One of the worse-case independent run results in terms of the

Table 5.5: Score metrics from 10 independent optimization runs

Cases	Score	Cases	Score
1	0.1364	6	0.0648
2	0.0033	7	0.0122
3	0.1091	8	0.0851
4	0.0913	9	0.1521
5	0.1388	10	0.1148
Mean	0.0907	Std	0.0512

distance metric is shown in Figure 5.2. Here the Pareto front is compared with the Pareto front from exact FORM reliability analysis. Note that the product of the interval width is plotted directly in Figure 5.2. This is the inverse of the objective space the problem was solved in via Equation 5.23. The approximation in this case is still reasonable in terms of both the magnitude of the values and spacing along the Pareto front.

Both the approximated Pareto front and exact Pareto front show a clear trade-off between interval uncertainty range and design performance, though the absolute impact on structural material volume is small. The increase of interval range multiples $\Delta_1 * \Delta_2$ causes penalty on design performance measured by volume of cantilever tube. The penalty stops after reaching a certain value of interval range. This is due to the nonlinearity of interval uncertainties in the original problem: if a local minimum of performance is already contained in the interval, expanding the interval does not further reduce performance unless a new minimum is included.

To quantify the convergence of the 10 independent runs, the Distance Score metric described in Equation 4.17 that measures the average normalized Euclidean distances between two Pareto fronts is employed here. The score metric is calculated for each of the 10 independent runs, the results are listed in Table 5.5.

The metrics indicate that even though the Pareto front found by the surrogate closely follows the front found by exact FORM simulation, the diversity of the approximated front is still affected by the stochastic nature of genetic algorithm.

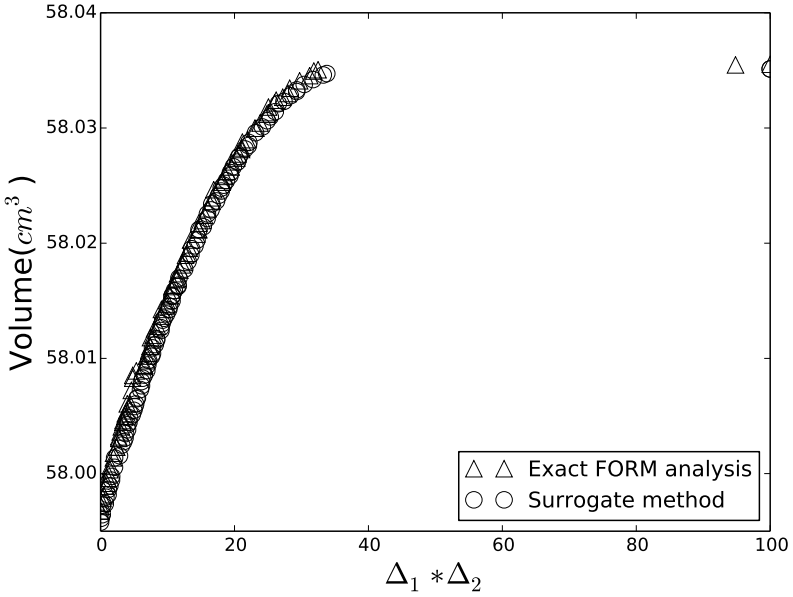


Figure 5.2: Cantilever tube problem: comparison of Pareto fronts for exact and approximate reliability analysis (Case 5).

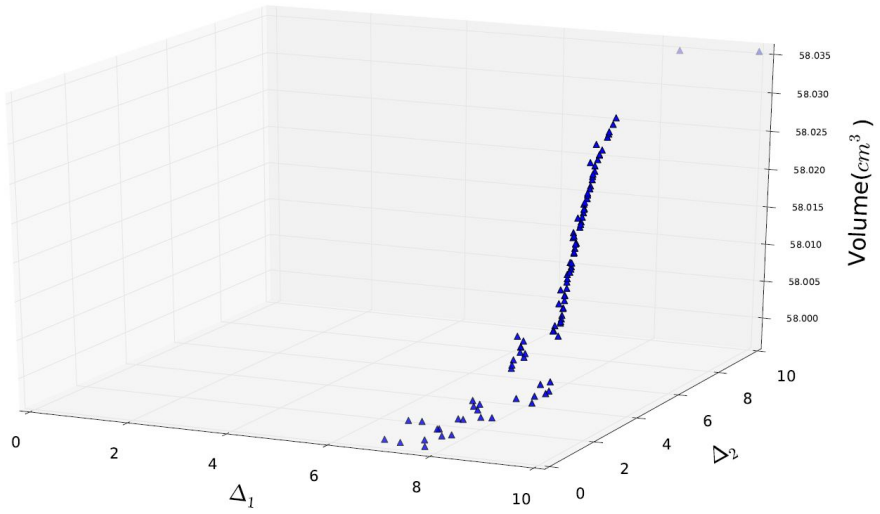


Figure 5.3: Cantilever tube problem: effect of interval range on the performance of multi-objective optimization.

To disclose the relative importance between the two interval uncertainties involved, a detailed comparison is illustrated in Figure 5.3. It can be seen that Pareto optimal points favor large Δ_1 values and small Δ_2 values, which indicates that reducing interval uncertainty Y_2 is more worthwhile in improving design performance compared

to Y_1. It also indicates that there is a critical value of Δ_2 at which point the design performance worsens rapidly.

This study reveals the necessity of the proposed multi-objective optimization framework for interval uncertainty problems. The closely converged Pareto front from surrogate method in Figure 5.2 suggests that the proposed surrogate model is capable in locating the worst-case performance in an efficient way for the multi-objective optimization, also it can provide accurate worst-case estimations in a noisy limit state function environment. Overall, the quadratic approximation to locate the worst performing point directly from the Kriging surrogate, and the coupling of this technique with the NSGA-II appears to successfully solve this two-interval problem.

5.5 Composite top-hat stiffened panel structural design

Marine composite structures are often used for high-speed, innovative vessels. Compared to conventional steel vessels, there is larger uncertainty in both the loading of the structure and the material variability of the composites for such vessels. Marine composite structural design then requires significantly larger loading and material uncertainty factors than steel structures. In this case study, it is attempted to apply interval uncertainties to represent large variabilities in loading component and composite material, while using conventional reliability methods to capture the configuration variability of the structure. Through the proposed method, a trade study between interval uncertainty ranges and design performance of a hollow rectangular, or top-hat, stiffened panel structure is performed. Such an analysis is similar to an early-stage vessel design problem, in which a robust estimation of structural weight is needed while loading remains uncertain.

A simplified grillage structure with Carbon Fiber-Reinforced Plastic (CFRP) materials (*Maneepan et al.*, 2005) is adopted here to demonstrate an application of proposed method. The typical configuration of this type of structure used in marine

(a) Picture of a FRP top-hat stiffened grillage.

(b) Picture of a FRP top-hat stiffened plate.

Figure 5.4: Composite FRP top-hat stiffened grillage plate.

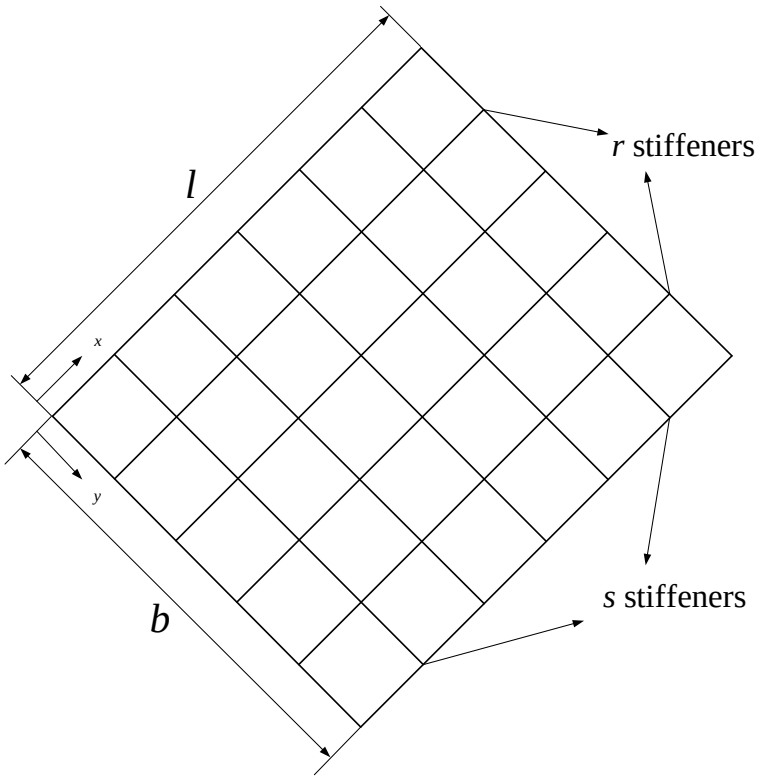


Figure 5.5: Schematic of a multi-stiffener grillage.

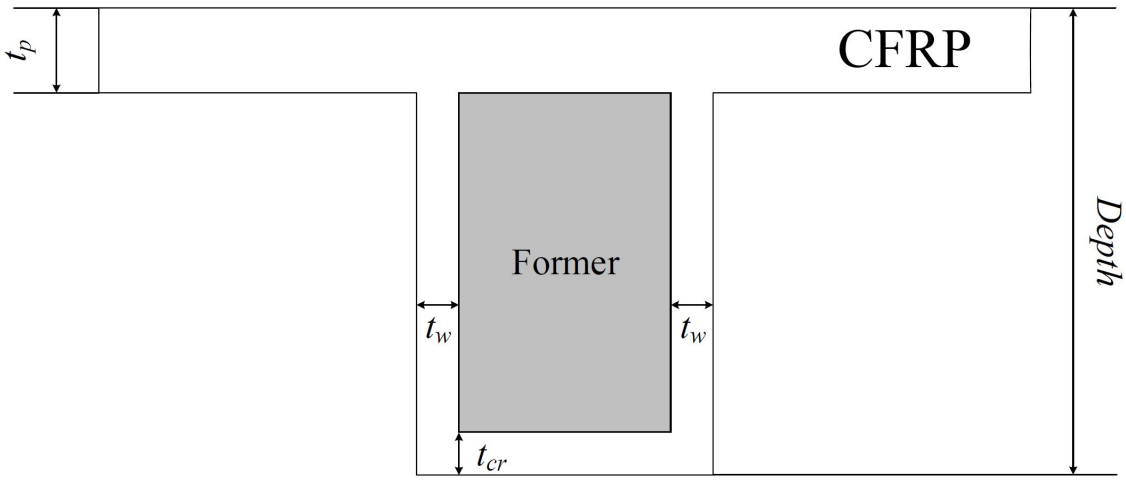


Figure 5.6: A sectional view of the CFRP single skin with top-hat stiffeners.

structures is shown in Figure 5.4. The size of this top-hat stiffened grillage plate is 4.2 m in length and 3.2 m in width, and spacing values for longitudinal and transverse frames are 800 mm and 700 mm, respectively. This gives the grillage plate three longitudinal and five transverse top-hat stiffeners.

A cross-section view of a top-hat shape stiffener with the attached plating is shown in Figure 5.6. This top-hat stiffener including base plate is consisted of crown, web, non-structural former. It is assumed that CFRP laminates are monolithic; hence no attempt is made to ascribe thickness of ply details, make-up, fiber volume fractions, etc. Such decisions would be taken up at a later stage after gross decisions about the choice of particular structural topology and material are made.

Under lateral pressure, both the strength and the deflection of such panels are important to structural design. For this work, the maximum panel deflection is considered as the limit state equation for the structure. The deflection limit state function is defined as follows:

$$g(\boldsymbol{X}, \boldsymbol{Y}) = w_{max} - w(\boldsymbol{X}, \boldsymbol{Y}) \tag{5.28}$$

where w_{max} is the allowed maximum deflection, taken as $L/150$ in this study. The deflection w of the grillage is estimated by using Navier's energy method *Bedair* (1997), where the deflection is determined by equating total strain energy to the work done by the load. The equation is given below:

$$w = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{16Pl^4b/EI_r}{\pi^6 mn[m^4(r+1) + \frac{I_s}{I_r}\frac{l^3}{b^3}n^4(s+1)]} \sin\frac{m\pi x}{l} \sin\frac{n\pi y}{b} \tag{5.29}$$

where I_r and I_s are moments of inertia for the longitudinal stiffeners and transverse stiffeners, respectively; m and n are wave numbers; E is equivalent Young's moduli; a reasonable pressure load P is estimated by the ABS High Speed Craft rules *American Bureau of Shipping* (2015)

The distribution descriptions for each components in Equation 5.28 are presented in Table 5.6. In this work, the geometric design variables are treated as random variables $\boldsymbol{X}$, while the interval variables $\boldsymbol{Y}$ are used to change the COVs of the loading estimate and Young's modulus in the limit state equation. The mean value of equivalent Young's modulus μ_E is taken as 140GPa.

Table 5.6: Distributions of variables in CFRP problem

Variables	Description	Mean	COV	Distribution
t_w^l	Longitudinal Web thickness	μ_1	0.03	Normal
t_{cr}^l	Longitudinal Crown thickness	μ_2	0.03	Normal
t_p	Plate thickness	μ_3	0.03	Normal
t_w^r	Transverse Web thickness	μ_4	0.03	Normal
t_{cr}^r	Transverse Crown thickness	μ_5	0.03	Normal
$Depth^l$	Longitudinal Depth	μ_6	0.03	Normal
$Depth^r$	Transverse Depth	μ_7	0.03	Normal
E	Equivalent Young's modulus	μ_E	$0.15 \pm \Delta_1$	Normal
P	Pressure Load	μ_P	$0.15 \pm \Delta_2$	Normal

The multi-objective optimization problem for interval trade-off study is formulated in Equation 5.30, within which the independent design variables are defined in

Table 5.7.

$$
\begin{aligned}
&min\ Weight(\boldsymbol{X}) \\
&min\ Cost(\boldsymbol{\Delta}) \\
&s.t.\ \hat{\beta}_L(\boldsymbol{X}, \boldsymbol{Y}) = \hat{y}(\boldsymbol{X}, \boldsymbol{Y}^*) \geq \beta_t \\
&where\ \boldsymbol{X} = [\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6, \mu_7, \Delta_1, \Delta_2]^T \\
&\quad \boldsymbol{Y_0} - \boldsymbol{\Delta} \leq \boldsymbol{Y} \leq \boldsymbol{Y_0} + \boldsymbol{\Delta}
\end{aligned}
\tag{5.30}
$$

Table 5.7: Design variables of CFRP problem

Design variables	Description	Lower bound	Upper bound
μ_1	Mean value of t_w^l	1.0mm	15mm
μ_2	Mean value of t_{cr}^l	1.8mm	15mm
μ_3	Mean value of t_p	5mm	45mm
μ_4	Mean value of t_w^r	1.0mm	15mm
μ_5	Mean value of t_{cr}^r	1.8mm	15mm
μ_6	Mean value of $Depth^l$	150mm	180mm
μ_7	Mean value of $Depth^r$	150mm	180mm
Δ_1	Interval variable Δ_1 range	0.01	0.1
Δ_2	Interval variable Δ_2 range	0.01	0.1

The above problem is optimized in both the proposed surrogate method presented in Algorithm 4, and the all FORM analysis in terms of different ways in computing β_L. The initial surrogate model is constructed using 200 sampling points. The NSGA-II parameters are set as 100 generations and a population size of 100. The target reliability criteria β_t are investigated at 3.0 as in the cantilever tube problem. The corresponding optimized results are shown in Figure 5.7. The Cost objective (Equation 5.23) is normalized and inverted to interval range for a better trade space view.

From the Pareto fronts in Figure 5.7, we can observe that for this composite structure case study, the absolute value of the penalty for information uncertainty is higher than in the cantilever tube case. The increased interval uncertainty range causes approximately a 40% weight increase on the design in the worst case scenario. The potential weight penalty is critical for composite material application in naval ship

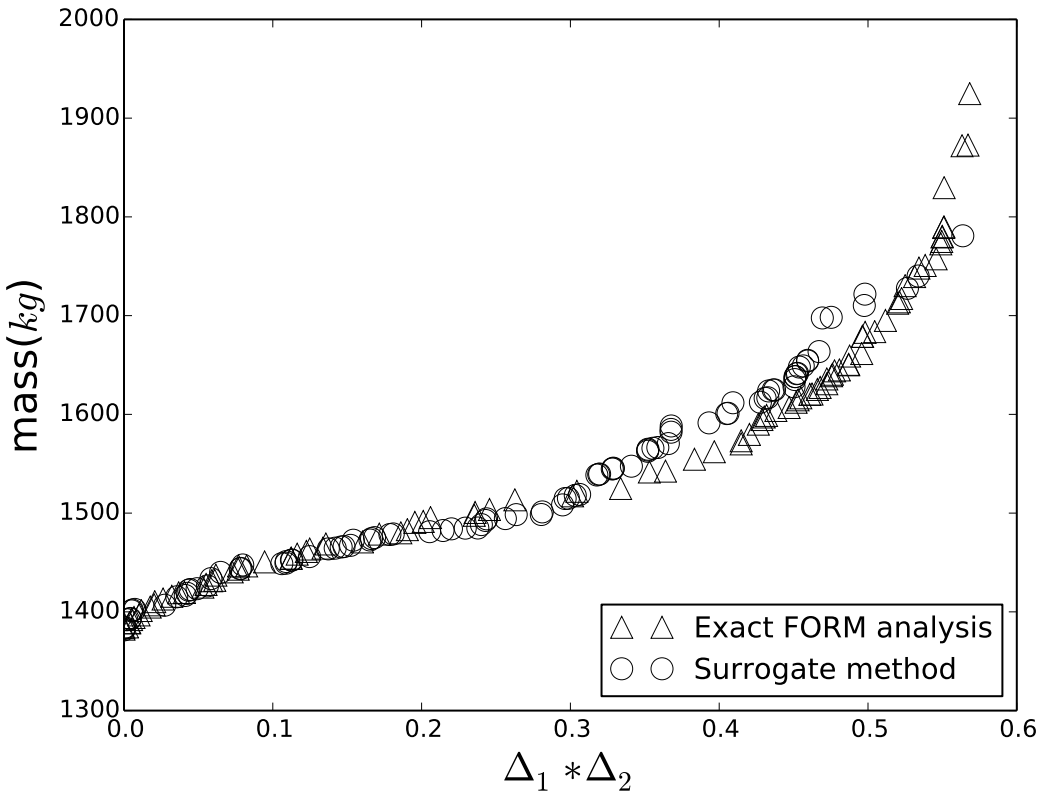


Figure 5.7: CFRP panel problem: comparison of Pareto fronts for exact and approximate reliability analysis.

designs where weight requirement is stringent. Therefore, this trade-off information can be highly valuable in early stage structural weight estimation.

However, in this more complex study the proposed method does not converge as closely to the high-fidelity results as it did in the simpler tube example. The approximation is more effective for lower value interval ranges on the left-hand side of the figure. In the region of larger interval range values, the surrogate predicts slightly worse (heavier) solutions, though it still resolves the shape and extent of the front accurately. A likely limit here is the suitability of a single Kriging model (in this case approximately 300 points at completion of the algorithm) to capture the entire design space of the more complex problem. The capacity of Kriging is limited by

the number of infill sampling points. The previous chapter has suggested that such problems can be overcome by adapting a surrogate model management framework, where clustering is used to build multiple surrogates in local regions of the design space. These results do suggest that absent such techniques, the scale of the problems solvable by the method may be limited. However, the scale of panel design problem here, with nine design variables and two uncertainty intervals is certainly relevant for industrial applications.

The proposed methodology was also computationally efficient for a problem of this size. The total number of reliability evaluations for this problem with the proposed algorithm was 298 - 200 for the initial surrogate, and 98 additional samples via refinement. This number is reasonable even if the simple reliability formulation used here is replaced with a more complex simulation-based method. However, when directly evaluating the reliability without a surrogate (e.g. performing a second nested optimization to determine the worse performance for each individual) over 10^6 reliability evaluations are needed. While some improvement in efficiency could undoubtedly be gained by tuning this inner search more than was done in the present study, the proposed method reduces the need for reliability evaluations by several orders of magnitude. Most importantly, this large reduction in reliability evaluations would allow uncertainty trade-off information to be generated with a computational burden compatible with the time scale of early-stage design.

5.6 Summary

This chapter presented and demonstrated a novel interval-uncertainty optimization approach to reduce epistemic uncertainty in design. Initially, a combined interval-reliability model was presented. In this approach, the limit state equation can contain both stochastic (aleatory) variables and interval-uncertainty variables. Then, a surrogate modeling approach was demonstrated that extended existing Kriging modeling

approaches by directly locating minima on the Kriging surface via the Jacobian and Hessian of the model. This approach was used to remove the innermost loop in the optimization - the location of the worst reliability performance within a given interval. Adaptive refinement criteria to ensure the accuracy of this approximation were also presented. Next, the approach was coupled to a conventional NSGA-II optimizer, and an information cost function defined in terms of the product of interval uncertainty widths. By including this information cost function as an objective, the impact of epistemic uncertainty on the design performance was revealed in the Pareto trade-spaces that resulted from the NSGA-II. The proposed approach was demonstrated on a cantilever tube and composite panel design problem. The results from the benchmark cantilever tube problem showed that the proposed approach was accurate and efficient in obtaining the Pareto fronts similar to those found from exact reliability simulations. Also in the composite panel design problem, the proposed surrogate method showed promising in capturing a good overview of the Pareto front.

With the help of the presented method, the decision makers are better equipped to explore novel structural designs in early-stage design. As shown in the composite panel case study result, the interval uncertainty study can fully prepare the designer to design against the worst-case scenario and make structural weight estimations under incomplete information. Additionally, this interval uncertainty trade study can quickly inform the designer regarding whether engineering improvement is critically needed for the design project. Further improvements on this surrogate method strategy is under investigation. The application of multiple Kriging surrogate models is envisioned to improve the performance of worst-case prediction for more practical applications.

CHAPTER VI

Summary

6.1 Conclusions

The objective of this research is to incorporate early stage design uncertainty into structural design optimization using a novel surrogate modeling strategy for rapid trade-off analysis. Structural decision making in early stage design is subject to inadequate information and knowledge about the design. The existence of early stage epistemic uncertainty complicates the decision maker's judgment in early stage structural decisions. In the worst case, the design under epistemic uncertainty can fail after the physical prototype is built and put into service. It was proposed in this dissertation to tackle this design challenge early in the design stage planning through a trade study between performance and the epistemic uncertainty. This dissertation work explored the interval-type epistemic uncertainty in structural design optimization, and gradually carry out a numerical efficient surrogate-assisted optimization framework to reveal the overall impact of epistemic uncertainty on the structural configuration designs.

Interval uncertainty is proposed to study information-based epistemic uncertainty in the early stage design. Interval uncertainty model well represents the non-probabilistic, reducible characteristic of epistemic uncertainty in the early stage. This dissertation introduced a new way of treating uncertainty in optimization problem by defining

uncertainty intervals as design variables. The presented optimization framework delivered a quantified multi-objective optimization analysis regarding both interval uncertainties and design performance objectives. The developed trade space indicated how interval uncertainty influenced the structural design performance. Interval uncertainty model is adaptable to other epistemic uncertainty forms. As introduced in Chapter II, the other epistemic uncertainty models all involve the concept of uncertain interval to some degree. Therefore, this research work laid a foundation for further study on early stage epistemic design uncertainty by first investigating interval-type uncertainty.

Introducing interval uncertainty into structural design optimization poses challenges in the numerical efficiency on the optimization implementations. Despite the promising insights gained through the interval uncertainty analysis in structural design optimization, the higher computational effort prevents the wide application of the interval uncertainty analysis. This dissertation presented a surrogate-assisted optimization strategy to develop efficient solutions for the proposed uncertainty analysis and optimization problems. The structural design case studies discussed in this dissertation validated the necessities of applying presented surrogate models for a rapid and reliable analysis. Moreover, the developed surrogate modeling techniques apply to more types of optimization problems, they can find far more widely applications for structural design and optimizations beyond the uncertainty in this dissertation. Surrogate modeling techniques are studied both in terms of efficiency and updating strategy throughout the interval uncertainty studies presented.

A robust design optimization framework considering interval uncertainty was first proposed to show the difference between robust optimal solution and deterministic optimal design solution. The robust optimal solution was designed to consider a interval uncertainty parameter in the ship's compartment length. By doing so, the robust solution had designed out the epistemic uncertainty early in the design stage,

and established a robust baseline structural weight estimation for later stage design process. The robust solution was inferior to the deterministic solution as it had 15% more structural weight. However, the design solution that is feasible over a range of interval length values provides flexibility for the designer considering later stage design changes. The robust design optimization was carried out in a surrogate-assisted evolutionary optimization framework. A RBF surrogate was build off-line to significantly reduce the number of fitness evaluations in the optimization process. The surrogate method proved efficient and accurate for the compartmentalization structural design study. The off-line construction methodology was sufficient for this single-objective design optimization considering fixed interval uncertainty length. This dissertation continued to explore advanced surrogate techniques in allocating limited computational simulations for structural design optimization.

Early stage design is marked by the limited time to generate high-fidelity experimental data or simulation results. Considering this situation, a surrogate modeling approach was investigated that enables interaction of high-fidelity and low-fidelity numerical models for early stage structural design optimization. The presented multiple surrogate based variable fidelity optimization framework was an extension of the VFO framework proposed by *Zhu et al.* (2014). This novel VFO surrogate structure manages multiple surrogates in an on-line construction scheme. A clustering algorithm was implemented to dynamically partition large sampling size data to local regions for Kriging surrogate modelings. These multiple surrogates worked effectively in scaling the low-fidelity Pareto front to the high-fidelity Pareto optimal front. The proposed surrogate algorithm was vigorously tested through benchmark numerical multi-objective optimization problems and a stiffened panel marine structural design problem. These case studies all demonstrated the superior performance of the proposed surrogate modeling approach. The on-line model of surrogate construction revealed itself as a more efficient computational scheme in allocating a limited number

of expensive simulations.

Following this strategy of building on-line surrogate models to enable efficient multi-objective optimization runs, this research continue to investigate the implication of interval uncertainty in marine structural designs. This research presented a means to quantify the trade-off analysis between structural designs with various levels of interval uncertainty. In this proposed framework, the epistemic uncertainty interval was treated as a design variable. Accordingly, an information cost function was defined to represent the effort needed in reducing the interval uncertainty. The uncertainty cost function is optimized along with performance function, thus making the uncertainty study a multi-objective optimization problem. The previous study showed that computation problems with interval uncertainty analysis is complicated as they normally involve an additional computation loop. This research presented a unique surrogate modeling solution that allows the extensive interval analysis in the MOGA while remains tractable. The proposed surrogate model extended the existing Kriging modeling approaches by directly locating the minima on the Kriging surface via utilizing higher order information of the surrogate. In other words, this surrogate approach is novel in that it can directly predict the extreme case in the interval domain, therefore, the innermost loop associated with interval analysis is removed. Adaptive refinement of the surrogate ensured the quality of the prediction and also gradually updated the surrogate on-line in the optimization process.

The proposed surrogate approach was tested in a novel composite panel structural design study. Its ability to predict the correct worst-case performance in various interval uncertainty domains was proved throughout the study. The optimization was able to quickly resolve the Pareto optimal front with the help of the developed surrogate model. The presented efficient optimization framework was shown to be effective for the early stage structural decision making process. By seeing the impact of various levels of interval uncertainty on the design performance, the designers are fully

prepared to make structural weight estimations under incomplete information at the early stage. The interval uncertainty trade-off study can also quickly inform the decision makers regarding where data is critically needed to increase the understanding in the early stage design. In these regards, the reported surrogate developments in this dissertation contributed to improve our capability for dealing with marine design under complex, uncertain early stage design environments.

6.2 Recommendations for future work

The research findings throughout this dissertation work motivate future research work in expanding the current framework. The multiple surrogate variable fidelity framework proposed in this dissertation can be used to refine the on-line surrogate modeling in the interval uncertainty trade-off study. The multiple Kriging construction methodology can significantly expand the applicability of surrogate modeling for larger design spaces. This could allow the interval uncertainty study applied in a larger problem domain such as an entire vessel structural design problem.

In the interval uncertainty study presented in Chapter V, the reliability simulation was considered as the computational-intensive part of the design problem. The proposed surrogate modeling technique can deal with simulation problems involving interval uncertainty analysis. In principle, the methodology can be used with other computationally-intensive simulations, such as Underwater Explosion (UNDEX) analysis or manufacturing analysis. This can broaden up the influence of the proposed uncertainty study in the structural settings considering other design priorities.

To achieve an overall understanding of epistemic uncertainty impact in marine design, some other epistemic uncertainty models are also worth investigations. The interval uncertainty study presented in this dissertation demonstrated valuable research findings for early stage design study. A comparison of interval-type uncertainty trade-off study to that of fuzzy logic, or evidence theory would help to gain

more insight into how to treat uncertainty in early stage design.

Apart from the uncertainty problem domain, surrogate modeling also needs increased capability to deal with common structural design challenges. The Kriging and Radial Basis Function models currently used in this technique assume smooth, continuous variables. Many early-stage structural design decisions are categorical, e.g. aluminum, titanium, or discrete in nature. Extension of the modeling techniques to include such discrete variables without simply repeating the analysis for each discrete value would allow the designer to better explore design options, and further broaden the applicability of the research.

BIBLIOGRAPHY

BIBLIOGRAPHY

Alefeld, G., and J. Herzberger (1983), *Introduction to interval computation*, New York: Academic press.

American Bureau of Shipping (2015), *Rules for Building and Classing High-Speed Naval Craft*, Houston.

Arnold, D. V., and H.-G. Beyer (2002), Local performance of the (1+ 1)-es in a noisy environment, *Evolutionary Computation, IEEE Transactions on*, *6*(1), 30–41.

Bedair, O. (1997), Analysis of stiffened plates under lateral loading using sequential quadratic programming (sqp), *Computers & Structures*, *62*(1), 63–80.

Ben-Haim, Y., and I. Elishakoff (2013), *Convex models of uncertainty in applied mechanics*, Elsevier.

Beyer, H.-G., and B. Sendhoff (2007), Robust optimization–a comprehensive survey, *Computer methods in applied mechanics and engineering*, *196*(33), 3190–3218.

Bichon, B. J., M. S. Eldred, L. P. Swiler, S. Mahadevan, and J. M. McFarland (2008), Efficient Global Reliability Analysis for Nonlinear Implicit Performance Functions, *AIAA Journal*, *46*(10), 2459–2468, doi:10.2514/1.34321.

Bishop, C. M. (1995), *Neural networks for pattern recognition*, Oxford university press.

Breitung, K. W. (1994), Asymptotic approximations for probability integrals.

Buhmann, M. D. (2000), Radial basis functions, *Acta Numerica 2000*, *9*, 1–38.

Cornell, C. A. (1969), A probability-based structural code*, *66*(12), doi: 10.14359/7446.

Cox, D. D., and S. John (1997), SDO: A Statistical Method for Global Optimization, in *in Multidisciplinary Design Optimization: State-of-the-Art*, pp. 315–329, SIAM, Philadelphia.

Cuneo, B. J., K. J. Maki, and D. J. Singer (2011), Applying soft constraints to multi-disciplinary design optimization using fuzzy logic, in *11th International Conference on Fast Sea Transportation (FAST 2011), Honolulu, Hawaii*.

Deb, K. (2001), *Multi-objective optimization using evolutionary algorithms*, vol. 16, John Wiley & Sons.

Deb, K., and R. B. Agrawal (1995), Simulated binary crossover for continuous search space, *Complex systems*, *9*(2), 115–148.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan (2002), A fast and elitist multiobjective genetic algorithm: Nsga-ii, *Evolutionary Computation, IEEE Transactions on*, *6*(2), 182–197.

Dempster, A. P. (1967), Upper and lower probabilities induced by a multivalued mapping, *The annals of mathematical statistics*, pp. 325–339.

Der Kiureghian, A., and O. Ditlevsen (2009), Aleatory or epistemic? does it matter?, *Structural Safety*, *31*(2), 105–112.

Ditlevsen, O., and P. Bjerager (1989), Plastic reliability analysis by directional simulation, *Journal of engineering mechanics*, *115*(6), 1347–1362.

Du, X. (2007), Interval reliability analysis, in *ASME 2007 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pp. 1103–1109, American Society of Mechanical Engineers.

Du, X., and W. Chen (2004), Sequential optimization and reliability assessment method for efficient probabilistic design, *Journal of Mechanical Design*, *126*(2), 225–233.

Du, X., A. Sudjianto, and B. Huang (2005), Reliability-based design with the mixture of random and interval variables, *Journal of Mechanical Design*, *127*(6), 1068–1076.

Dubourg, V., B. Sudret, and J.-M. Bourinet (2011), Reliability-based design optimization using kriging surrogates and subset simulation, *Structural and Multidisciplinary Optimization*, *44*(5), 673–690, doi:10.1007/s00158-011-0653-8, arXiv: 1104.3667.

Eberhart, R. C., J. Kennedy, et al. (1995), A new optimizer using particle swarm theory, in *Proceedings of the sixth international symposium on micro machine and human science*, vol. 1, pp. 39–43, New York, NY.

Echard, B., N. Gayton, and M. Lemaire (2011), AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation, *Structural Safety*, *33*(2), 145–154, doi:10.1016/j.strusafe.2011.01.002.

Enevoldsen, I., and J. D. Sørensen (1994), Reliability-based optimization in structural engineering, *Structural safety*, *15*(3), 169–196.

Evans, J. H. (1959), Basic design concepts, *Journal of the American Society for Naval Engineers*, *71*(4), 671–678.

Fang, K.-T., D. K. Lin, P. Winker, and Y. Zhang (2000), Uniform design: theory and application, *Technometrics*, *42*(3), 237–248.

Faulkner, D. (1987), Toward a better understanding of compression induced tripping, *Steel structures. London: Elsevier Applied Science*, pp. 159–75.

Faulkner, D., J. Adamchak, G. Snyder, and M. Vetter (1973), Synthesis of welded grillages to withstand compression and normal loads, *Computers & Structures*, *3*(2), 221–246.

Ferson, S., C. A. Joslyn, J. C. Helton, W. L. Oberkampf, and K. Sentz (2004), Summary from the epistemic uncertainty workshop: consensus amid diversity, *Reliability Engineering & System Safety*, *85*(13), 355–369, doi:10.1016/j.ress.2004.03.023.

Fonseca, C. M., and P. J. Fleming (1998), Multiobjective optimization and multiple constraint handling with evolutionary algorithms. ii. application example, *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, *28*(1), 38–47.

Forrester, A., A. Sobester, and A. Keane (2008), *Engineering design via surrogate modelling: a practical guide*, John Wiley & Sons.

Forrester, A. I., and A. J. Keane (2009), Recent advances in surrogate-based optimization, *Progress in Aerospace Sciences*, *45*(1), 50–79.

Frangopol, D. M., and R. B. Corotis (1996), Reliability-based structural system optimization: state-of-the-art versus state-of-the-practice, in *Analysis and Computation (1996)*, pp. 67–78, ASCE.

Gale, P. A. (2003), *The Ship Design Process*, vol. I, chap. 5, New Jersey : Society of Naval Architects and Marine Engineers.

Gano, S. E., J. E. Renaud, J. D. Martin, and T. W. Simpson (2006), Update strategies for kriging models used in variable fidelity optimization, *Structural and Multidisciplinary Optimization*, *32*(4), 287–298.

Goel, T., R. T. Haftka, W. Shyy, and N. V. Queipo (2007), Ensemble of surrogates, *Structural and Multidisciplinary Optimization*, *33*(3), 199–216.

Goldberg, D. E. (1989), Genetic algorithms in search, optimization, and machine learning, *Addion wesley*, *1989*.

Gray, A. W., A. S. Daniels, and D. J. Singer (2010), Impacts of fuzzy logic modeling for constraints optimization, *Naval Engineers Journal*, *122*(2), 121–132.

Hackl, J. (2013), PyRe 5.0.2 https://github.com/hackl/pyre.

Haftka, R. T. (1991), Combining global and local approximations, *AIAA journal*, *29*(9), 1523–1525.

Hamza, K., and K. Saitou (2005), Vehicle crashworthiness design via a surrogate model ensemble and a co-evolutionary genetic algorithm, in *ASME 2005 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pp. 899–907, American Society of Mechanical Engineers.

Hannapel, S., and N. Vlahopoulos (2014), Implementation of set-based design in multidisciplinary design optimization, *Structural and Multidisciplinary Optimization*, *50*(1), 101–112.

Hardin, R., K. Choi, N. Gaul, and C. Beckermann (2015), Reliability based casting process design optimisation, *International Journal of Cast Metals Research*, *28*(3), 181–192.

Hasofer, A. M., and N. C. Lind (1974), Exact and invariant second-moment code format, *Journal of the Engineering Mechanics division*, *100*(1), 111–121.

Hedayat, A., N. Sloane, and J. Stufken (1999), *Orthogonal Arrays: Theory and Applications*, Springer Science & Business Media.

Helton, J. C., and D. E. Burmaster (1996), Guest editorial: treatment of aleatory and epistemic uncertainty in performance assessments for complex systems, *Reliability Engineering & System Safety*, *54*(2), 91–94.

Herrera, F., M. Lozano, and J. L. Verdegay (1998), Tackling real-coded genetic algorithms: Operators and tools for behavioural analysis, *Artificial intelligence review*, *12*(4), 265–319.

Isaacs, A., T. Ray, and W. Smith (2007), An evolutionary algorithm with spatially distributed surrogates for multiobjective optimization, in *Progress in Artificial Life*, pp. 257–268, Springer.

Jiang, C., W. X. Li, X. Han, L. X. Liu, and P. H. Le (2011), Structural reliability analysis based on random distributions with interval parameters, *Computers & Structures*, *89*(2324), 2292–2302, doi:10.1016/j.compstruc.2011.08.006.

Jiang, C., X. Han, W. Li, J. Liu, and Z. Zhang (2012), A hybrid reliability approach based on probability and interval for uncertain structures, *Journal of Mechanical Design*, *134*(3), 031,001.

Jiang, C., H. C. Xie, Z. G. Zhang, and X. Han (2015), A new interval optimization method considering tolerance design, *Engineering Optimization*, *47*(12), 1637–1650, doi:10.1080/0305215X.2014.982632.

Jin, R., W. Chen, and T. W. Simpson (2001), Comparative studies of metamodelling techniques under multiple modelling criteria, *Structural and Multidisciplinary Optimization*, *23*(1), 1–13.

Jin, Y. (2005), A comprehensive survey of fitness approximation in evolutionary computation, *Soft computing*, *9*(1), 3–12.

Jin, Y. (2011), Surrogate-assisted evolutionary computation: Recent advances and future challenges, *Swarm and Evolutionary Computation*, *1*(2), 61–70.

Jin, Y., and J. Branke (2005), Evolutionary optimization in uncertain environments-a survey, *Evolutionary Computation, IEEE Transactions on*, *9*(3), 303–317.

Jin, Y., and B. Sendhoff (2004), Reducing fitness evaluations using clustering techniques and neural network ensembles, in *Genetic and Evolutionary Computation–GECCO 2004*, pp. 688–699, Springer.

Kaymaz, I. (2005), Application of kriging method to structural reliability problems, *Structural Safety*, *27*(2), 133–151.

Keane, R. G. (2012), Reducing total ownership cost: Designing inside-out of the hull# 169, *Naval Engineers Journal*, *124*(4), 67–80.

Kim, C., and K. K. Choi (2008), Reliability-based design optimization using response surface method with prediction interval estimation, *Journal of Mechanical Design*, *130*(12), 121,401.

Kirkpatrick, S., M. P. Vecchi, et al. (1983), Optimization by simmulated annealing, *science*, *220*(4598), 671–680.

Klanac, A., and J. Jelovica (2009), Vectorization and constraint grouping to enhance optimization of marine structures, *Marine Structures*, *22*(2), 225–245.

Klanac, A., S. Ehlers, and J. Jelovica (2009), Optimization of crashworthy marine structures, *Marine Structures*, *22*(4), 670–690.

Kleijnen, J. P., S. M. Sanchez, T. W. Lucas, and T. M. Cioppa (2005), State-of-the-art review: a users guide to the brave new world of designing simulation experiments, *INFORMS Journal on Computing*, *17*(3), 263–289.

Kleijnen, J. P. C. (2009), Kriging metamodeling in simulation: A review, *European Journal of Operational Research*, *192*(3), 707–716, doi:10.1016/j.ejor.2007.10.013.

Klir, G. J. (2005), *Uncertainty and information: foundations of generalized information theory*, John Wiley & Sons.

Kosko, B. (1992), Neural networks and fuzzy systems: a dynamical approach to machine intelligence, *Englewood Cliffs. NJ: PrenticeHall*, *1*(99), 1.

Kuczera, R. C., and Z. P. Mourelatos (2009), On estimating the reliability of multiple failure region problems using approximate metamodels, *Journal of Mechanical Design*, *131*(12), 121,003.

Lee, S., P. Von Allmen, W. Fink, A. E. Petropoulos, and R. J. Terrile (2005), Comparison of multi-objective genetic algorithms in optimizing q-law low-thrust orbit transfers, in *GECCO Conference Late-breaking Paper, Washington, DC*.

Li, J., M. G. Parsons, and R. GUTIERREZ-FRAILE (1996), The use of fuzzy logic in shipping and shipbuilding market modeling, analysis and forecasting. discussion, *Journal of ship production*, *12*(2), 85–98.

Li, M., N. Williams, and S. Azarm (2009), Interval Uncertainty Reduction and Single-Disciplinary Sensitivity Analysis With Multi-Objective Optimization, *Journal of Mechanical Design*, *131*(3), 031,007–031,007, doi:10.1115/1.3066736.

Liang, J., Z. P. Mourelatos, and J. Tu (2008), A single-loop method for reliability-based design optimisation, *International Journal of Product Development*, *5*(1), 76–92.

Liu, Y., and M. Collette (2014a), Improving surrogate-assisted variable fidelity multi-objective optimization using a clustering algorithm, *Applied Soft Computing*, *24*, 482–493.

Liu, Y., and M. Collette (2014b), Surrogate-assisted robust design optimization considering interval-type uncertainty, in *Maritime Technology and Engineering*, pp. 287–293, CRC Press.

Liu, Y., and M. Collette (2015), Efficient optimization framework for robust and reliable structural design considering interval uncertainty, in *Analysis and Design of Marine Structures V*, pp. 555–563, CRC Press.

Liu, Y., H. Jeong, and M. Collette (2016), Efficient optimization of reliability-constrained structural design problems including interval uncertainty, submitted in *Computers&Structures*.

MacQueen, J., et al. (1967), Some methods for classification and analysis of multivariate observations, in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, pp. 281–297, Oakland, CA, USA.

Madsen, H. O., S. Krenk, and N. C. Lind (2006), *Methods of structural safety*, Courier Corporation.

Mahadevan, S., and A. Haldar (2000), *Probability, reliability and statistical method in engineering design*, John Wiley & Sons.

Maneepan, K., H. Jeong, and R. Shenoi (2005), Optimization of frp top-hat stiffened single skin and monocoque sandwich plates using genetic algorithm, in *The Fifteenth International Offshore and Polar Engineering Conference*, International Society of Offshore and Polar Engineers.

Mansour, A. (1993), *Probability-based ship design procedures: a demonstration*, Ship Structure Committee.

McKay, M. D., R. J. Beckman, and W. J. Conover (1979), A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics*, *21*(2), 239–245.

Melchers, R. E. (1999), *Structural reliability analysis and prediction*, John Wiley & Son Ltd.

Mendel, J. M. (2001), Uncertain rule-based fuzzy logic system: introduction and new directions.

Meng, Z., G. Li, B. P. Wang, and P. Hao (2015), A hybrid chaos control approach of the performance measure functions for reliability-based design optimization, *Computers & Structures*, *146*, 32–43, doi:10.1016/j.compstruc.2014.08.011.

Miettinen, K. (2012), *Nonlinear multiobjective optimization*, vol. 12, Springer Science & Business Media.

Moens, D., and D. Vandepitte (2007), Interval sensitivity theory and its application to frequency response envelope analysis of uncertain structures, *Computer Methods in Applied Mechanics and Engineering*, *196*(21), 2486–2496.

Möller, B., and M. Beer (2008), Engineering computation under uncertainty–capabilities of non-traditional models, *Computers & Structures*, *86*(10), 1024–1041.

Moore, R. E. (1966), *Interval analysis*, vol. 4, Prentice-Hall Englewood Cliffs.

Mourelatos, Z. P., and J. Zhou (2006), A design optimization method using evidence theory, *Journal of mechanical design*, *128*(4), 901–908.

Müller, W. G. (2007), *Collecting spatial data: optimum design of experiments for random fields*, Springer Science & Business Media.

Myers, R. H., D. C. Montgomery, and C. M. Anderson-Cook (2016), *Response surface methodology: process and product optimization using designed experiments*, John Wiley & Sons.

Nelder, J. A., and R. Mead (1965), A simplex method for function minimization, *The computer journal*, *7*(4), 308–313.

Oberkampf, W. L., and J. C. Helton (2005), Evidence theory for engineering applications.

Oberkampf, W. L., S. M. DeLand, B. M. Rutherford, K. V. Diegert, and K. F. Alvin (2002), Error and uncertainty in modeling and simulation, *Reliability Engineering & System Safety*, *75*(3), 333–357.

Oberkampf, W. L., J. C. Helton, C. A. Joslyn, S. F. Wojtkiewicz, and S. Ferson (2004), Challenge problems: uncertainty in system response given uncertain parameters, *Reliability Engineering & System Safety*, *85*(1), 11–19.

Ong, Y.-S., P. B. Nair, and K. Lum (2006), Max-min surrogate-assisted evolutionary algorithm for robust design, *Evolutionary Computation, IEEE Transactions on*, *10*(4), 392–404.

Paik, J. K., and A. K. Thayamballi (2003), *Ultimate limit state design of steel-plated structures*, John Wiley & Sons.

Parsons, M. G., and D. Singer (2000), A fuzzy logic agent for design team communications and negotiations, in *Conference on Computer Applications and Information Technology in the Maritime Industries, COMPIT 2000*.

Poloni, C. (1995), Hybrid ga for multi objective aerodynamic shape optimization, *Genetic algorithms in engineering and computer science*, *33*, 397–415.

Powell, M. J. D. (1981), *Approximation theory and methods*, Cambridge university press.

Quagliarella, D. (1998), *Genetic algorithms and evolution strategy in engineering and computer science: recent advances and industrial applications*, John Wiley & Son Ltd.

Rahman, M., and J. Caldwell (1995), Ship structures: improvement by rational design optimisation, *International shipbuilding progress*, *42*(429), 61–102.

Ray, T., and W. Smith (2006), A surrogate assisted parallel multiobjective evolutionary algorithm for robust engineering design, *Engineering Optimization*, *38*(8), 997–1011.

Richardson, J. N., R. Filomeno Coelho, and S. Adriaenssens (2015), Robust topology optimization of truss structures with random loading and material properties: A multiobjective perspective, *Computers & Structures*, *154*, 41–47, doi: 10.1016/j.compstruc.2015.03.011.

Rigterink, D., M. Collette, and D. J. Singer (2013), A method for comparing panel complexity to traditional material and production cost estimating techniques, *Ocean Engineering*, *70*, 61–71.

Rubinstein, R. Y., and D. P. Kroese (2011), *Simulation and the Monte Carlo method*, vol. 707, John Wiley & Sons.

Sacks, J., W. J. Welch, T. J. Mitchell, and H. P. Wynn (1989), Design and analysis of computer experiments, *Statistical Science*, *4*(4), 409–423.

Sanchez, E., S. Pintos, and N. V. Queipo (2008), Toward an optimal ensemble of kernel-based approximations with engineering applications, *Structural and multidisciplinary optimization*, *36*(3), 247–261.

Sekulski, Z. (2009), Least-weight topology and size optimization of high speed vehicle-passenger catamaran structure by genetic algorithm, *Marine Structures*, *22*(4), 691–711.

Shafer, G. (1976), *A mathematical theory of evidence*, vol. 1, Princeton university press Princeton.

Simpson, T. W., J. Poplinski, P. N. Koch, and J. K. Allen (2001), Metamodels for computer-based engineering design: survey and recommendations, *Engineering with computers*, *17*(2), 129–150.

Singer, D. J., N. Doerry, and M. E. Buckley (2009), What is set-based design?, *Naval Engineers Journal*, *121*(4), 31–43.

Stadler, W. (1984), Applications of multicriteria optimization in engineering and the sciences (a survey), *Multiple Criteria*.

Stein, M. (1987), Large sample properties of simulations using latin hypercube sampling, *Technometrics*, *29*(2), 143–151.

Sudret, B. (2012), Meta-models for structural reliability and uncertainty quantification, *arXiv preprint arXiv:1203.2062*.

Taguchi, G. (1986), *Introduction to quality engineering: designing quality into products and processes*, The Organization.

Temple, D. W. (2015), A multi-objective collaborative optimization framework to understand trade-offs between naval lifetime costs considering production, operation, and maintenance, Ph.D. thesis, University of Michigan.

Tsutsui, S. (1999), A comparative study on the effects of adding perturbations to phenotypic parameters in genetic algorithms with a robust solution searching scheme, in *Systems, Man, and Cybernetics, 1999. IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on*, vol. 3, pp. 585–591, IEEE.

Tsutsui, S., and A. Ghosh (1997), Genetic algorithms with a robust solution searching scheme, *Evolutionary Computation, IEEE Transactions on*, *1*(3), 201–208.

Tu, J., K. K. Choi, and Y. H. Park (1999), A new study on reliability-based design optimization, *Journal of Mechanical Design*, *121*(4), 557–564.

Viana, F. A. (2011), Multiple surrogates for prediction and optimization, Ph.D. thesis, University of Florida.

Wang, P., Z. Wang, B. D. Youn, and S. Lee (2015), Reliability-based robust design of smart sensing systems for failure diagnostics using piezoelectric materials, *Computers & Structures*, *156*, 110–121, doi:10.1016/j.compstruc.2015.04.012.

Wu, X., et al. (2008), Top 10 algorithms in data mining, *Knowledge and information systems*, *14*(1), 1–37.

Yang, X., Y. Liu, Y. Gao, Y. Zhang, and Z. Gao (2014), An active learning kriging model for hybrid reliability analysis with both random and interval variables, *Structural and Multidisciplinary Optimization*, *51*(5), 1003–1016.

Youn, B. D., K. K. Choi, and Y. H. Park (2003), Hybrid analysis method for reliability-based design optimization, *Journal of Mechanical Design*, *125*(2), 221–232.

Zadeh, L. A. (1965), Fuzzy sets, *Information and control*, *8*(3), 338–353.

Zhu, J., Y.-J. Wang, and M. Collette (2014), A multi-objective variable-fidelity optimization method for genetic algorithms, *Engineering Optimization*, *46*(4), 521–542.

Zitzler, E., K. Deb, and L. Thiele (2000), Comparison of multiobjective evolutionary algorithms: Empirical results, *Evolutionary computation*, *8*(2), 173–195.