UNIVERSIDADE FEDERAL DO CEARÁ

CENTRO DE TECNOLOGIA

DEPARTAMENTO DE ENGENHARIA ESTRUTURAL E CONSTRUÇÃO CIVIL

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL

MESTRADO ACADÊMICO EM ENGENHARIA CIVIL

LEONARDO GONÇALVES RIBEIRO

EFFICIENT OPTIMIZATION OF COMPOSITE STRUCTURES USING MULTI-FIDELITY MODELS

FORTALEZA

2022

LEONARDO GONÇALVES RIBEIRO

EFFICIENT OPTIMIZATION OF COMPOSITE STRUCTURES USING MULTI-FIDELITY MODELS

Dissertation submitted to the Programa de Pós-Graduação em Engenharia Civil of the Centro de Tecnologia of the Universidade Federal do Ceará, as a partial requirement for obtaining the title of Master in Civil Engineering. Concentration Area: Structural Engineering

Advisor: Prof. D. Sc. Evandro Parente Junior

Co-advisor: Antônio Macário Cartaxo de Melo

FORTALEZA

2022

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

R369e Ribeiro, Leonardo Gonçalves.
Efficient Optimization of Composite Structures using Multi-Fidelity Models / Leonardo Gonçalves Ribeiro. – 2022.
219 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia Civil: Estruturas e Construção Civil, Fortaleza, 2022.
Orientação: Prof. Dr. Evandro Parente Junior.
Coorientação: Prof. Dr. Antônio Macário Cartaxo de Melo.

1. Composite Materials. 2. Structural Optimization. 3. Surrogate Models. 4. Sequential Approximate Optimization. 5. Multi-Fidelity Models. I. Título.

CDD 624.1

LEONARDO GONÇALVES RIBEIRO

EFFICIENT OPTIMIZATION OF COMPOSITE STRUCTURES USING MULTI-FIDELITY MODELS

Dissertation submitted to the Programa de Pós-Graduação em Engenharia Civil of the Centro de Tecnologia of the Universidade Federal do Ceará, as a partial requirement for obtaining the title of Master in Civil Engineering. Concentration Area: Structural Engineering

Approved on: 31 de Março de 2022

EXAMINATION BOARD

Prof. D. Sc. Evandro Parente Junior (Advisor)
Universidade Federal do Ceará (UFC)

Antônio Macário Cartaxo de Melo (Co-advisor)
Universidade Federal do Ceará (UFC)

Prof. Dr. Marcelo Silva Medeiros Júnior
Universidade Federal do Ceará (UFC)

Prof. Dr. Marco Antônio Luersen
Universidade Tecnológica Federal do Paraná (UTFPR)

ACKNOWLEDGEMENTS

A meus pais, Ricardo e Solange Ribeiro, e minha irmã, Ana Beatriz, pela confiança, dedicação e compreensão mesmo nos momentos em que eu estou tão absorto no meu mundo que esqueço de falar o quanto eu gosto deles.

À Carolina Trompieri, por me acompanhar nesse caminho que traçamos juntos. Que eu possa continuar aprendendo contigo por muito tempo ainda. Obrigado por tudo.

Aos colegas do Laboratório de Mecânica Computacional e Visualização da UFC, aos mestrandos da minha turma e ao doutorando Elias Barroso, pela amizade e pelos contínuos ensinamentos ofertados. Um agradecimento especial também à Marina Maia que, mesmo estando em outros caminhos, sempre se manteve disposta a ajudar.

Ao professor Evandro Parente Junior, pela orientação impecável e a constante disponibilidade para retirar minha dúvidas sobre o tema.

Aos outros professores do Departamento de Engenharia Estrutural e Construção Civil, em especial ao professor Antônio Macário Cartaxo de Melo, meu co-orientador, pelos ensinamentos prestados nesses mais de dois anos de mestrado.

Aos professores Marcelo Silva Medeiros Junior e Marco Antônio Luersen, por aceitarem participar da banca.

À CNPQ, UFC, DEECC e FUNCAP pelo aporte financeiro ofertado durante o mestrado e ao apoio à pesquisa acadêmica brasileira.

ABSTRACT

Composite structures are receiving increasing interest in the last few decades. These often require the use of numerical analysis methods, such as the Finite Element Method (FEM) or the Isogeometric Analysis (IGA). Due to their high design flexibility, the optimization of composites is very promising, as it may provide more efficient structures. This work uses Surrogate Based Optimization (SBO) to make the process more efficient. Examples of robust surrogate modeling techniques are Radial Basis Functions (RBF) and Kriging. For an efficient optimization process, one may use the model to locate promising regions in the design space and add new data points. This way, the approximation quality in the regions of interest is improved. Another way of improving the model quality is to consider information from low-fidelity sampling points, which are often cheaper and easier to assess. Here, a low-fidelity point refers to data evaluated using lower fidelity sources, such as using a coarser mesh or a simplified theory. If low-fidelity and high-fidelity sources are well-correlated, the low-fidelity sample may capture the general behavior of the function in the design space, thus greatly improving the model prediction while also allowing for a lower computational cost. These are denominated Multi-Fidelity Models (MFMs). This work aims at employing these techniques in the optimization of laminated composites and functionally graded structures, mainly plates and shells. The use of adaptive sampling is integrated into MFMs, where error-based exploration is employed to further improve the model. Different surrogate modeling approaches and adaptive sampling criteria are tested. Different aspects of multi-fidelity modeling are discussed, such as importance of correlation between sources, analyses cost, and ratio between low and high-fidelity samples. Our proposed methodology is able to solve both functionally graded and laminate problems, and very good results are also obtained when expensive constraints are considered. Results show that MFMs are able to significantly reduce the number of expensive analyses required to find the optimum in most optimization problems. Accuracy is also improved, especially in complex multi-modal optimization problems. At the same time, as MFMs present higher model complexity, model building and evaluation costs are more restrictive than those found for usual single-fidelity models.

Keywords: Composite materials. Structural optimization. Surrogate models. Sequential Approximate Optimization. Multi-Fidelity Models.

RESUMO

Estruturas de material compósito vêm recebendo cada vez mais interesse nas últimas décadas. A análise destas requer o uso de métodos numéricos, como o Método dos Elementos Finitos (MEF) ou a Análise Isogeométrica (AIG). Devido à grande flexibilidade destes materiais, a otimização dessas estruturas é desejável. Neste trabalho, a Otimização baseada em Modelos Substitutos (*Surrogate-Based Optimization*, SBO) será utilizada. Exemplos de abordagens robustas de modelagem são as Funções de Base Radial (*Radial Basis Functions*, RBF) e o Kriging. Para um processo de otimização mais eficiente, o projetista pode utilizar o modelo para auxiliar na seleção de novos pontos amostrais em regiões de interesse. Estes pontos amostrais podem ser inseridos, melhorando a aproximação do modelo na região. Outra forma de melhorar a qualidade do modelo é considerar informação de pontos amostrais de baixa fidelidade, que são normalmente mais baratos de avaliar. Esses pontos correspondem, por exemplo, a análises feitas utilizando malhas mais grosseiras ou teorias simplificadas. Se as fontes de baixa e alta fidelidade apresentarem uma boa correlação, a amostra de menor fidelidade pode ser capaz de capturar o comportamento geral da função e, desse modo, melhorar a precisão do modelo. Esses modelos são denominados Modelos Multi-Fidelidade (*Multi-Fidelity Models*, MFMs). Este trabalho visa empregar tais técnicas na otimização de estruturas laminadas e estruturas com gradação funcional, em especial placas e cascas. A amostragem adaptativa será integrada aos MFMs, e a exploração baseada em medidas de erro será utilizada para melhorar o modelo pela adição de novos pontos. Diferentes abordagens de modelagem e critérios de inserção de pontos serão testados. Diferentes aspectos da modelagem multi-fidelidade serão discutidos, como a importância da correlação entre as fontes, o custo da análise, e a razão entre o número de amostras de baixa e alta fidelidade. A metodologia proposta é capaz de resolver problemas de otimização de compósitos laminados e graduados funcionalmente, e resultados promissores também são encontrados ao considerar restrições caras. Os resultados mostram que os MFMs são capazes de reduzir significativamente o número de análises de alta fidelidade necessárias para achar o ótimo na maioria dos problemas de otimização. A precisão do processo também é melhorada, em especial em problemas multi-modais mais complexos. Ao mesmo tempo, os MFMs apresentam uma maior complexidade, e o custo de treiná-los e avaliá-los é mais restritivo.

Palavras-chave: Materiais compósitos. Otimização estrutural. Modelos substitutos. Otimização Sequencial Aproximada. Modelos Multi-Fidelidade.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AEI	Augmented Expected Improvement
ANN	Artifical Neural Network
BK	Blind Kriging
CAD	Computer Aided Design
CAE	Computer Aided Engineering
CDF	Cumulative Distribution Function
CEI	Constrained Expected Improvement
CFD	Computational Fluid Dynamic
CLT	Classical Lamination Theory
Co-Kriging	Cooperative Kriging
Co-RBF	Cooperative Radial Basis Functions
Co-SVR	Cooperative Support Vector Regression
CPT	Classical Plate Theory
CV	Cross Validation
DE	Differential Evolution
DNN	Deep Neural Network
DoE	Design of Experiments
EGO	Efficient Global Optimization
EI	Expected Improvement
FAST	Finite element AnalysiS Tool
FEM	Finite Element Method
FG	Functionally Graded
FGM	Functionally Graded Material
FGP	Functionally Graded Plate
FPF	First Ply Failure
FRC	Fiber-Reinforced Composite
FSDT	First-order Shear Deformation Theory
GA	Genetic Algorithm
GP	Gaussian Process
HF	High-Fidelity

HK	Hierarchical Kriging
HSDT	Higher-order Shear Deformation Theory
HSS	Hammersley Sequence Sampling
ICK	Improved Cooperative Kriging
IGA	Isogeometric Analysis
k-FCV	*k*-Fold Cross Validation
LCB	Lower Confidence Bound
LF	Low-Fidelity
LHS	Latin Hypercube Sampling
LOOCV	Leave-One-Out Cross Validation
MARS	Multivariate Adaptive Regression Splines
MFM	Multi-Fidelity Model
MISO	Mixed Integer Surrogate Optimization
MLE	Maximum Likelihood Estimator
MSE	Mean Squared Error
NRMSE	Normalized Root Mean Squared Error
NURBS	Non-Uniform Rational B-Splines
OK	Ordinary Kriging
OLHS	Optimized Latin Hypercube Sampling
PDF	Probability Density Function
PF	Probability of Feasibility
PI	Probability of Improvement
PSO	Particle Swarm Optimization
RBF	Radial Basis Function
RS	Random Sampling
SAO	Sequential Approximate Optimization
SBO	Surrogate-Based Optimization
SF	Safety Factor
SO-MI	Surrogate Optimization-Mixed Integer
SQP	Sequential Quadratic Programming
SVM	Support Vector Machine
SVR	Support Vector Regression

TTO	Tamura-Tomota-Ozawa
UCB	Upper Confidence Bound
UK	Universal Kriging
VF	Variable Fidelity
VF-EI	Variable Fidelity Expected Improvement
VF-LCB	Variable Fidelity Lower Confidence Bound
VF-PI	Variable Fidelity Probability of Improvement
VF-WEI	Variable Fidelity Weighted Expected Improvement
WEI	Weighted Expected Improvement

CONTENTS

1 INTRODUCTION

Composite materials have brought a significant improvement in a variety of fields, such as naval, aeronautical, and civil engineering[4]. These materials are formed by the combination of two or more constituents, providing enhanced properties compared to each component individually[5, 6]. In this work, the focus will be given to two types of composite materials: fiber-reinforced and functionally graded materials.

For structural applications, Fiber-Reinforced Composites (FRCs) are usually manufactured as thin layers, also denominated as *laminas* or *plies*. These materials often present high stiffness-to-weight and strength-to-weight ratios, and the orientation of each ply and the stacking sequence can be chosen to provide the required strength for a specific application[7]. However, the sudden change in material properties in laminate structures may lead to various damage modes, such as delamination and debonding[5].

On the other hand, Functionally Graded Materials (FGMs) are known for their smooth and continuous variation of material properties in the structure, which eliminates interface problems and stress concentrations, especially for problems with high temperatures and thermal gradients[8, 9]. This way, FGMs provide a higher fracture toughness and a better distribution of residual stresses[8, 10]. These materials were initially developed in Japan in the 1980s to serve as a thermal barrier for the aerospace industry[11], but now are used in several applications[12, 13].

These two materials introduce different parameters which can act as design variables for optimization problems, such as the orientation and stacking sequence for laminates or the material gradation for FGMs[5, 14]. Due to manufacturing constraints, design variables are often discrete for laminate structures[7, 15, 16]. On the other hand, on FG structures, material gradation can usually be defined by continuous variables[17, 1, 18]. Most researchers deal with the optimization of beams, plates, and shells while trying to optimize aspects such as the critical buckling load, fundamental frequency, weight, and cost of the structure[5, 14].

For composite structures, analytical solutions for displacements, strains, stresses, buckling load and frequencies are available only for very simple problems in terms of geometry, loading, and boundary conditions[19, 20, 21, 22]. In most cases, responses should be obtained by an approximated computational method, such as the Finite Element Method (FEM) or the Isogeometric Analysis (IGA)[23]. However, these methods often demand a significant computational time to perform a structural analysis[24], especially for complex models.

The most popular algorithms for structural optimization are nature-inspired methods[25].

These are accurate population-based zero-order optimization algorithms and, thus, do not require the computation of gradients, which is especially important for discrete problems. Due to their global search capabilities, these algorithms are much more capable of finding the global optimum in multimodal problems, even for continuous optimization[26, 27].

In that matter, a variety of nature-inspired optimization algorithms are employed in the optimization of composite structures, such as the GA[28, 15, 29], the PSO[26, 30, 16, 17], and the Differential Evolution (DE)[31, 32, 33, 34]. A major problem with these is that they often require hundreds or even thousands of function evaluations to find the optimum design[35, 36, 37]. At the same time, structural analyses from simulation-based methods such as FEM and IGA can become very expensive. Thus, the high computational cost is a significant hindrance for optimization of real-life structures[38].

When numerical analyses are too expensive, a common approach to deal with the high computational cost of heuristic optimization algorithms is parallelization, taking advantage of a powerful machine with multiple cores[27, 15, 16]. However, such machines are not always available to the user. Furthermore, while computing power significantly increased in the last few decades, so did the required fidelity and model complexity for engineering problems. Thus, time constraints are still a major barrier for structural optimization[39].

Alternatively, for a cheaper optimization process, one may approximate the structural response using an appropriate response surface method. In this approach, a surrogate model is built from the structural responses at a set of sampling points, providing an easier and faster, but approximated, assessment for the true, expensive function[40, 41]. For engineering applications, commonly employed models are Artifical Neural Networks (ANNs)[1, 18, 42], Support Vector Regression (SVR)[43, 44, 45], Radial Basis Functions (RBFs)[46, 47, 48], and Kriging[49, 50, 51].

For optimization problems, the model should aim at being more accurate in regions closer to the global optimum. This way, the model itself may be used to guide the further addition of new sampling points, thus improving its accuracy in regions of interest. This technique is known as the adaptive sampling[52], or Sequential Approximate Optimization (SAO). This approach was popularized after Jones, Schonlau and Welch[53] presented the Efficient Global Optimization (EGO) algorithm, which can solve optimization problems by iteratively improving a Kriging model using the Expected Improvement (EI) criterion. Soon, other researchers proposed different models and infill criteria that are competitive with EGO in terms of the computational cost and accuracy[54, 55, 56, 37, 57].

It is important to note that, to this day, there are still some open issues in SAO, mainly regarding discrete optimization, constraint-handling, and the choice of the model to be employed[58]. Even though a variety of techniques have been proposed to deal with these aspects, there are no general guidelines for when each should be employed[37]. This work aims at tackling these issues for the optimization of composite structures.

In the last few years, the use of modeling techniques for discrete optimization is gaining interest. Bartz-Beielstein and Zaefferer[59] present a review of the main approaches to tackle these problems. A common method is to simply ignore the discrete structure, as long as data can be represented as a vector. However, other approaches may perform better in most problems, e.g. the use of inherently discrete models, mapping approaches, or similarity-based modeling[59, 37, 58]. The selection of infill points should also be performed using an appropriate algorithm that allows for the optimization in a discrete space, such as bio-inspired optimization algorithms.

Regarding constraint-handling methods, most approaches often work by employing a surrogate model to approximate an expensive constraint function and, then, the user may consider a given feasibility function, based on the model prediction, to penalize infeasible designs. Researchers proposed a variety of feasibility functions[60, 61, 55, 62, 63], but it is still unclear which one performs better in most problems.

Concerning the choice of the model to be employed, Radial Basis Function (RBF) and Kriging models are often among the best-performing methods for mathematical and engineering problems[64, 65, 66]. However, some improvements have been proposed to these methods in the last few years[52], which have yet to be addressed for most engineering optimization problems. In particular, Multi-Fidelity Models (MFMs) seem to provide a very accurate and robust approximation, as they may consider sampling points derived from both a High-Fidelity (HF) and a Low-Fidelity (LF) source[67, 68, 52, 69, 70, 45] (e.g. using finer and coarser meshes). This allows for a better exploration of the design space, as it is cheaper to evaluate multiple Low-Fidelity (LF) sampling points. If sources are well-correlated, the surrogate model may better guide the further addition of new sampling points.

In that matter, this work compares two different Multi-Fidelity Models (MFMs): The Cooperative Kriging (Co-Kriging), proposed by Kennedy and O'Hagan[71], and the Hierarchical Kriging (HK), proposed by Han and Görtz[72]. These will be employed considering different adaptive sampling criteria. It is important to note that few papers employed MFMs with adaptive

sampling techniques[73, 74, 75, 76, 77], which is very interesting from an optimization standpoint, and even fewer employed this approach to the optimization of composite structures[78, 79].

1.1 Aims and scope

This work aims at proposing and implementing a methodology for the Sequential Approximate Optimization (SAO) of laminate and functionally graded composite structures. Single and multi-fidelity models will be employed, and a comparison will be performed between different techniques to assess which approach is better suited to the optimization of composite structures, focusing on plates and shallow shells. For this purpose, this work aims:

a) to develop a methodology for surrogate-based optimization of laminate and FG structures;
b) to implement existing and develop new SAO methods based on the use of multi-fidelity models;
c) to apply the proposed methodology to the optimization of composite structures;
d) to compare the performance of single-fidelity and multi-fidelity models in terms of accuracy, efficiency, and robustness;
e) to compare the performance of SAO with conventional optimization approaches.

1.2 Organization of the text

The remainder of this work is organized as follows. Chapter 2 describes the main concepts of laminated and functionally graded structures, including aspects important to the optimization of these structures. Chapter 3 presents details about the structural analysis of composite plates and shallow shells, such as the kinematic assumptions and internal forces evaluation.

Chapter 4 further discusses the optimization of composite structures, and a quick review over state-of-the-art algorithms for structural optimization is presented. Three robust algorithms are described, GA, PSO, and DE, as well as appropriate methods to handle constraints and discrete variables.

Chapter 5 presents the Sequential Approximate Optimization (SAO) and its main aspects, such as model building, infill criteria, and constraint-handling techniques, focusing on single-fidelity models. Chapter 6 then addresses multi-fidelity techniques and their use in SAO.

Chapter 7 presents the numerical examples, and the efficiency and accuracy of

the proposed algorithms are demonstrated by a set of problems regarding the optimization of laminate and FG plates and shallow shells. Finally, Chapter 8 presents conclusions of the work and suggestions for future research.

2 COMPOSITE MATERIALS

On a macroscopic scale, composite materials are formed by the combination of two or more materials in a way to create a new material that presents desirable properties that could not be achieved from any of the constituents alone[7, 6]. This is often accomplished by the addition of reinforcement to a matrix material. The reinforcement will act as the principal load-carrying member, while the matrix serves as a load-transfer medium.

Composites may be divided into three types, depending on how the combination of materials is performed: particulate, fibrous, or laminated composites. Particulate composites are formed by the addition of particles of reinforcement in a matrix of another. Fibrous composites consist of the addition of fiber reinforcement. Laminated composites are made of layers of one or more materials stacked, embedded in a matrix material.

Composite material can also be built by a specific combination of these three materials[20, 7]. For instance, Fiber-Reinforced Composites (FRCs) are often employed in the form of laminated components. These are usually built as thin layers known as plies. Due to fiber orientation, these plies present orthotropic behavior[16], showing a much higher strength and stiffness in the fiber direction. However, these plies can be stacked in different orientations to provide resistance in different directions. Thus, the designer may pick fiber orientations, thickness of each lamina, constituent materials, and stacking sequence (also known as *layup*) to provide the required strength for a particular application[7].

Even though these materials provide high strength-to-weight and stiffness-to-weight ratios, the stacking of plies incurs a sudden change in material properties, which may lead to various damage modes such as delamination and debonding[5]. For example, these materials can become severely compromised in high-temperature environments due to high residual inter-layer stresses caused by thermal expansion.

In the 1980s, a group of Japanese researchers developed a novel material, denominated as Functionally Graded Material (FGM), to serve as a thermal barrier for the aerospace industry[11, 8]. The use of a smooth and continuous gradation between ceramic and metallic materials provided both a great thermal and mechanical resistance, while not presenting specific failure modes from laminate composites. Nowadays, FGMs are used in a variety of applications[12], where continuous changes in their composition, microstructurem and porosity are employed to improve the structural performance in different environments.

The following sections present a brief discussion about these materials, especially

regarding important aspects for the analysis and optimization of composite structures.

2.1 Fiber-Reinforced Composites

In laminated Fiber-Reinforced Composites (FRCs), plies are stacked using different materials and orientations to provide the required stiffness and strength to the structure. There is a very close relationship between how the laminate is manufactured and its end use[20]. Laminate structures can be applied to civil and military aircraft, automotive structures, and marine risers[20, 4].

Since the 1940s, composite materials have been employed in many engineering applications due to their high specific stiffness and strength along with their low weight[9]. In some applications, such as in aircraft structures, the weight savings also help to reduce fuel costs, which justified further research in the area[4]. To compete with homogeneous metallic materials, composites must be affordable and allow for some manufacturing flexibility.

Usually, due to manufacturing constraints, fiber orientation is given by discrete values. In many applications, to provide strength in two main directions, cross-ply laminates are employed, where the structure has only 0° and 90° plies alternately (e.g. $[90^\circ, 0^\circ, 90^\circ, 0^\circ]$). One may also employ angle-ply laminates, where any fiber orientation angle can be employed (e.g. $[0^\circ, 90^\circ, 45^\circ]$)[80].

FRCs can also be classified regarding their mid-plane symmetry. In symmetric laminates, the upper half of the laminate is a mirror of the lower half in terms of fiber orientation, materials, and thickness. These are identified by the subscript s, as in $[45^\circ, 90^\circ, 30^\circ]_s$, which refers to the layup $[45^\circ, 90^\circ, 30^\circ, 30^\circ, 90^\circ, 45^\circ]$. Anti-symmetric laminates are similar, but the fiber orientations in the lower half are the negative of the upper half, as in $[45^\circ, 90^\circ, 30^\circ, -30^\circ, -90^\circ, -45^\circ]$.

Asymmetric laminates, on the other hand, present no mid-plane symmetry. Furthermore, balanced laminates present a negative fiber orientation for each positive one, as in $[90^\circ, -45^\circ, 0^\circ, 90^\circ, -45^\circ, 0^\circ]$.

2.1.1 Constitutive equations

On a microscopic scale, the interactions between constituents are very important to determine the failure modes in FRCs. However, on a macroscopic scale, the composite constituents can be considered as a homogeneous orthotropic material[81]. Here, the mechanical

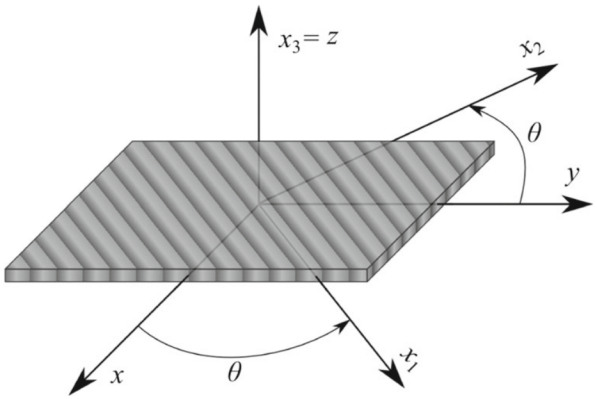
formulation of laminate composites will be derived, while kinematics and analysis methods will be presented in the next chapter.

In this work, two-dimensional formulations will be adopted for the analysis of laminated composites. This way, their mechanical behavior can be studied by two basic approaches: Equivalent Single Layer Theories, e.g. Classical Lamination Theory (CLT) and First-order Shear Deformation Theory (FSDT), or 3D elasticity theory.

Here, the FSDT will be adopted, which is an extension of the Reissner-Mindlin plate theory to laminate composite plates[7, 82]. The FSDT requires that normal lines to the mid-surface remain straight but, different from the CLT, these are not necessarily perpendicular to the mid-surface. Thus, shear deformations are considered in a simplified way.

For laminates with unidirectional fibers, the material can be considered, in a macroscopic scale, as a homogeneous orthotropic material in the local axis of the ply[20], denoted as x_1, x_2 and x_3 in Figure 1, where θ is the fiber orientation. These do not necessarily coincide with the global axis, which is usually related to the geometry of the structure.

Figure 1 – Local axis of a single ply.



Source: Barroso, Parente and Melo[16]

It is fair to assume that FRCs present a linear elastic behavior until very close to failure[20]. Thus, considering the FSDT, the plane-stress state and the symmetry of the constitutive matrix, the generalized Hooke's law can be written as:

$$\boldsymbol{\sigma} = \mathbf{C}\boldsymbol{\varepsilon} \quad \Rightarrow \quad \begin{Bmatrix} \sigma_1 \\ \sigma_2 \\ \tau_{13} \\ \tau_{23} \\ \tau_{12} \end{Bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} & 0 & 0 & 0 \\ Q_{21} & Q_{22} & 0 & 0 & 0 \\ 0 & 0 & Q_{44} & 0 & 0 \\ 0 & 0 & 0 & Q_{55} & 0 \\ 0 & 0 & 0 & 0 & Q_{66} \end{bmatrix} \begin{Bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \gamma_{13} \\ \gamma_{23} \\ \gamma_{12} \end{Bmatrix} \tag{2.1}$$

Usually, this equation is written as:

$$\begin{aligned}
\boldsymbol{\sigma}_1 = \mathbf{Q}\,\boldsymbol{\varepsilon}_1 \quad &\Rightarrow \quad \begin{Bmatrix} \sigma_1 \\ \sigma_2 \\ \tau_{12} \end{Bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} & 0 \\ Q_{21} & Q_{22} & 0 \\ 0 & 0 & Q_{66} \end{bmatrix} \begin{Bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \gamma_{12} \end{Bmatrix} \\
\boldsymbol{\tau}_1 = \mathbf{Q}_s\,\boldsymbol{\gamma}_1 \quad &\Rightarrow \quad \begin{Bmatrix} \tau_{13} \\ \tau_{23} \end{Bmatrix} = \begin{bmatrix} Q_{44} & 0 \\ 0 & Q_{55} \end{bmatrix} \begin{Bmatrix} \gamma_{13} \\ \gamma_{23} \end{Bmatrix}
\end{aligned} \tag{2.2}$$

where $\boldsymbol{\tau}_1$ and $\boldsymbol{\gamma}_1$ refers to the out-of-plane stresses and strains and the subscript 1 to the local system of the ply. In terms of the orthotropic mechanical properties, the coefficients Q_{ij} are given by:

$$\begin{aligned}
&Q_{11} = \frac{E_1}{1-\nu_{12}\,\nu_{21}}, \quad Q_{12} = \frac{\nu_{21}\,E_1}{1-\nu_{12}\,\nu_{21}}, \quad Q_{22} = \frac{E_2}{1-\nu_{12}\,\nu_{21}} \\
&Q_{66} = G_{12}, \quad Q_{44} = G_{13}, \quad Q_{55} = Q_{44}
\end{aligned} \tag{2.3}$$

When considering thermal effects, the evaluation of the normal stresses should consider the thermal strain vector, as in:

$$\boldsymbol{\sigma}_1 = \mathbf{Q}\,\boldsymbol{\varepsilon}_1 \quad \Rightarrow \quad \begin{Bmatrix} \sigma_1 \\ \sigma_2 \\ \tau_{12} \end{Bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} & 0 \\ Q_{21} & Q_{22} & 0 \\ 0 & 0 & Q_{66} \end{bmatrix} \begin{Bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \gamma_{12} \end{Bmatrix} \tag{2.4}$$

To derive strains in the global axis, one should employ a transformation matrix, which depends only on the fiber orientation:

$$\begin{aligned}
\boldsymbol{\varepsilon}_1 = \mathbf{T}\,\boldsymbol{\varepsilon} \quad &\Rightarrow \quad \begin{Bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \gamma_{12} \end{Bmatrix} = \begin{bmatrix} \cos^2\theta & \sin^2\theta & \sin\theta\cos\theta \\ \sin^2\theta & \cos^2\theta & -\sin\theta\cos\theta \\ -2\sin\theta\cos\theta & 2\sin\theta\cos\theta & \cos^2\theta - \sin^2\theta \end{bmatrix} \begin{Bmatrix} \varepsilon_x \\ \varepsilon_y \\ \gamma_{xy} \end{Bmatrix} \\
\boldsymbol{\gamma}_1 = \mathbf{T}_s\,\boldsymbol{\gamma} \quad &\Rightarrow \quad \begin{Bmatrix} \varepsilon_{13} \\ \varepsilon_{23} \end{Bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{Bmatrix} \gamma_{xz} \\ \gamma_{yz} \end{Bmatrix}
\end{aligned} \tag{2.5}$$

Finally, replacing Eq. (2.5) in Eq. (2.2), we may obtain the global stresses by:

$$\begin{aligned}
&\boldsymbol{\sigma} = \mathbf{T}^T\,\mathbf{Q}\,\mathbf{T}\,\boldsymbol{\varepsilon} \quad \Rightarrow \quad \boldsymbol{\sigma} = \overline{\mathbf{Q}}\,\boldsymbol{\varepsilon} \\
&\boldsymbol{\tau} = \mathbf{T}_s^T\,\mathbf{Q}_s\,\mathbf{T}_s\,\boldsymbol{\gamma} \quad \Rightarrow \quad \boldsymbol{\tau} = \overline{\mathbf{Q}}_s\,\boldsymbol{\gamma}
\end{aligned} \tag{2.6}$$

where $\overline{\mathbf{Q}}$ and $\overline{\mathbf{Q}}_s$ are the constitutive matrices in the global system.

According to the FSDT, there is a linear strain variation along with the thickness of the structure. Since material properties change for each ply, the change in the characteristic

moduli is often discontinuous, and the stress distribution is not necessarily linear[20]. This discontinuity is also responsible for additional failure modes in laminate structures, such as delamination and debonding[5].

The macroscopic representation of failure for laminate composites demands an adequate failure criterion to ensure structural safety. Most failure criteria are derived from their homogeneous isotropic counterparts[83, 84]. For laminate composites, there are three different types of failure criteria[84]: non-interactive criteria, such as Maximum Strain Criterion or Maximum Stress Criterion; interactive criteria, such as the Tsai-Hill or Tsai-Wu; and failure mode based theories, such as the Puck failure criterion.

The simplest approach is the use of non-interactive criteria. For instance, in the Maximum Stress Criterion, the layer fails when one of its stresses exceeds their corresponding ultimate stress. For each ply k, this criterion can be written as:

$$\mathrm{SF}_k = \min\left(\frac{\sigma_1^u}{\sigma_1^k}, \frac{\sigma_2^u}{\sigma_2^k}, \frac{\tau_{12}^u}{\tau_{12}^k}\right) \tag{2.7}$$

where σ_1^k, σ_2^k and τ_{12}^k are the maximum stresses in the k-th ply (which will be located either on the top or the bottom of the ply), σ_1^u, σ_2^u and τ_{12}^u are the ultimate strains in each direction, and SF_k is the Safety Factor (SF) for the k-th ply.

This approach is often regarded as too simplistic, even for isotropic materials. Alternatively, an interactive criterion may be employed, such as the Tsai-Hill criterion[84]:

$$\mathrm{SF}_k = \left(\frac{\sigma_1^k}{\sigma_1^u}\right)^2 + \left(\frac{\sigma_2^k}{\sigma_2^u}\right)^2 + \left(\frac{\tau_{12}^k}{\tau_{12}^u}\right)^2 - \left(\frac{\sigma_1^k}{\sigma_1^u}\right)\left(\frac{\sigma_2^k}{\sigma_1^u}\right) \tag{2.8}$$

where σ_1^u, σ_2^u and τ_{12}^u are the ultimate stresses in each direction. Different ultimate stresses can be considered if the material is under tension or compression. Alternatively, the Tsai-Wu criterion is also widely used in practice[84].

The actual failure of the laminate can be defined by the First Ply Failure (FPF) criterion, where the structure SF is given by:

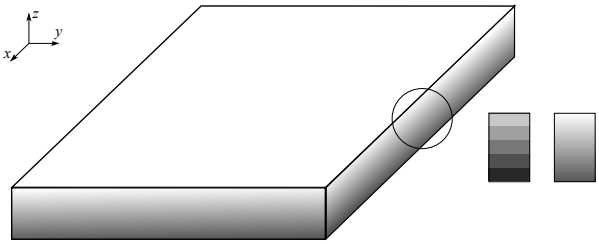
$$\mathrm{SF} = \min_{k=1}^{N_p}(\mathrm{SF}_k) \tag{2.9}$$

where N_p is the total number of plies. It is worth noting that this criterion is conservative since it neglects the stress redistribution to the remaining plies once failure occurs on a single ply. Techniques that aim at considering the progressive failure of the laminate are more complex and are out of the scope of this work.

2.2 Functionally Graded Materials

Functionally Graded Materials (FGMs) are composites where the proportion of the constituents present a smooth gradation. This feature allows Functionally Graded (FG) structures to present a better performance when subjected to thermal effects since stress concentrations are mostly eliminated[85, 86]. As shown in Figure 2, material gradation can take place in a step-wise manner, which gives rise to a multilayered structure with an interface between layers, or in a continuous variation. In this work, the latter will be considered.

Figure 2 – Functionally Graded (FG) Plate.



Source: the author

FGMs are relatively novel materials, first employed by Japanese researchers in 1984 to serve as a thermal barrier in aerospace structures[11, 8]. Nowadays, applications can be found in the fields of defense, energy, aerospace, biomedical, electronics, and optoelectronics[12, 13, 87]. Due to their excellent performance when subjected to thermal loadings, FGMs have also been employed in combustion chambers in a variety of industries[12]. Usually, such materials are made by the combination of metal (which presents better ductility and toughness) and ceramic (which presents high stiffness and low thermal conductivity). Examples of metals include aluminum (Al), stainless steel (SUS304), and titanium (Ti), while ceramic materials include silicon carbide (SiC), aluminum oxide (Al_2O_3), and silicon nitride (Si_3N_4).

Udupa, Rao and Gangadharan[12] presented a review over the different uses, engineering applications, and methods for fabricating FG structures. The authors state that the main drawback for these structures is that the total preparation cost is still high since more advanced manufacturing processes are required[12]. The most common methods for fabricating a continuous FGM are liquid phase processes, such as the centrifugal casting[88], even though gas-based methods and solid phase processes can also be employed. Naebe and Shirvanimoghaddam[13] also presented a vast review over fabrication methods for these materials.

The volume fraction distribution of each constituent is often defined by a closed-form equation, and the derivation of the constitutive relations requires the use of an appropriate

homogenization technique. Those aspects will be further discussed in the following sections.

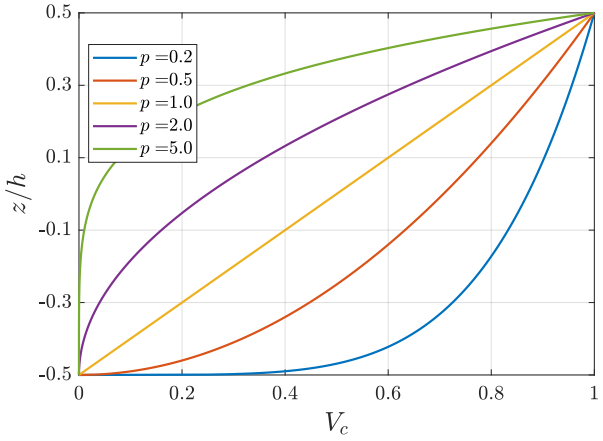
2.2.1 Volume fraction distribution

In Functionally Graded Materials (FGMs), material composition, defined by the volume fraction of each component, vary smoothly in the structure domain. The choice for a proper distribution is very important, since it should allow for a certain degree of flexibility. The most popular equation is the power-law rule (P-FGM)[8, 85]:

$$V_c = V_{c,b} + (V_{c,t} - V_{c,b}) \left(\frac{1}{2} + \frac{z}{h}\right)^p$$
$$V_m = 1 - V_c \tag{2.10}$$

where the subscripts c and m refer to the ceramic and the metal, respectively, and $V_{c,t}$ and $V_{c,b}$ are the ceramic volume fractions on the top and on the bottom of the structure, respectively. The gradation is controlled by the exponent p, and Figure 3 depicts how the ceramic volume fraction changes through the thickness for different values of p considering $V_{c,b} = 0.0$ and $V_{c,t} = 1.0$.

Figure 3 – Ceramic volume fraction gradation for different p using the power-law rule.



Source: the author

The main advantage of this gradation lies in its simplicity since there are very few effective parameters. On the other hand, this also means that there is less flexibility for defining the FG structure. For instance, there is no way for the user to design a symmetrical gradation using the power-law rule from Eq. (2.10). Some researchers proposed similar methods, which aim at introducing more parameters to the power-law to increase its flexibility[30, 89, 90]. Also, researchers often make use of other simple rules, such as sigmoid (S-FGM) and exponential (E-FGM) functions[14].

Alternatively, one may employ an approach based on a set of control points to achieve the desired gradation, such as cubic Hermite polynomials[91, 14] and B-Spline functions[1, 17, 2, 42]. These are particularly interesting for optimization problems, as the increased flexibility allows for more optimal designs.

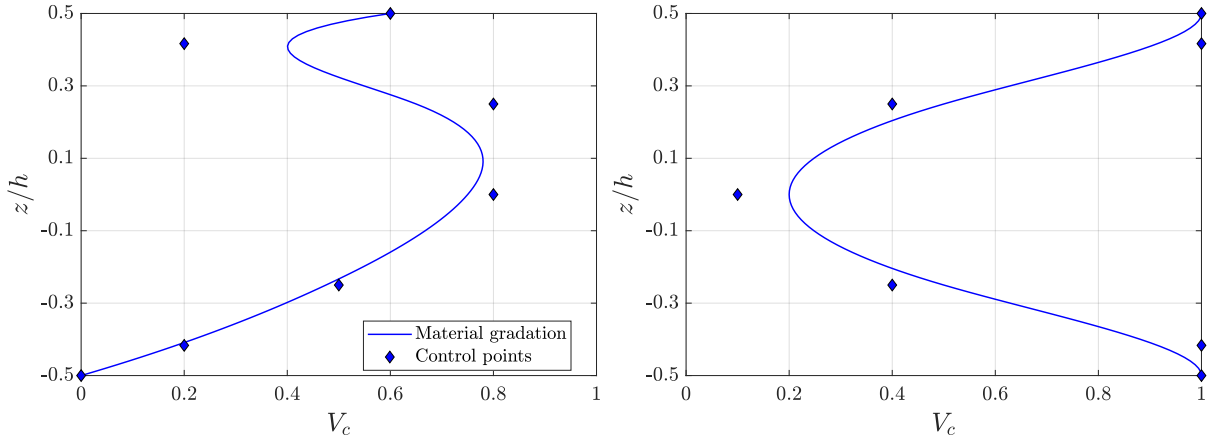
In this work, B-Spline functions will be employed. These are non-interpolating functions with a high degree of continuity[92] which allow for a continuous and smooth variation of material properties. For instance, using univariate B-Spline functions, the volume fraction is given by:

$$V_c(\xi) = \sum_{i=1}^{n_{cp}} B_{i,p}(\xi)\, V_{c,i}, \quad \xi \in [0,1]$$
$$V_m(\xi) = 1 - V_c(\xi) \tag{2.11}$$

where n_{cp} is the number of control points, $V_{c,i}$ is the ceramic volume fraction for the i-th control point, $B_{i,p}$ is the corresponding B-Spline basis, p is the basis degree, and ξ is the parametric coordinate.

Figure 4 presents an example of a gradation represented by a B-Spline curve. In cases where a symmetrical gradation is desired, one may impose that the control points are symmetric with respect to the mid-plane. Also, a designer can use bivariate or trivariate B-Splines so that the gradation can be defined in two or three directions. This method will be further discussed in Section 3.1.1.

Figure 4 – Ceramic volume fraction gradation using B-Spline functions.



(a) Asymmetric (b) Symmetric

Source: the author

2.2.2 Effective material properties

After defining the material gradation, one should use appropriate techniques to evaluate the effective material properties, which are often denominated homogenization techniques. Since FGMs are usually employed in high-temperature environments, it might be important to consider the change in material properties due to the temperature. This variation can be represented by the Touloukian's equation[93]:

$$P = P_0 \left(P_{-1} T^{-1} + 1 + P_1 T + P_2 T^2 + P_3 T^3\right) \tag{2.12}$$

where P_i are unique coefficients for the properties of each constituent and P is the estimated property for a temperature T (in Kelvin) [93, 8].

The homogenization may be performed by the rule of mixtures, also known as the Voigt model[8, 94]:

$$P_f = \sum_{j=1}^{n_m} P_j V_j \tag{2.13}$$

which is a weighted average between the properties of each material. This method is very simple and efficient, but numerical and experimental studies show that this approach may lead to very poor results. Akbarzadeh, Abedini and Chen[94] presented numerical comparisons between different micromechanical models on the structural response of Functionally Graded Plates (FGPs), showing that the rule of mixtures may present a very large discrepancy in results: the maximum deflection was up to 23% lower using the Voigt model, and the buckling load up to 30% higher. Medeiros, Parente and Melo[95] compared results obtained in the analysis of pressurized hollow cylinders with experimental data, also arguing that the Voigt model should be avoided due to its large discrepancies with experimental results.

Alternatively, many other homogenization methods can be found in the literature such as Reuss, Hashin-Shtrikman bounds, Tamura, and Self-Consistent Method[94, 95, 14]. In this work, the Hashin-Shtrikman lower bound will be employed, also known as the Mori-Tanaka scheme. The method assumes that a matrix phase is reinforced by spherical particles, taking into account the elastic fields among neighboring inclusions[8]. This is a simple approach that can achieve very good results for the estimation of mechanical properties by[95, 8]:

$$\frac{K - K_m}{K_c - K_m} = \frac{V_c}{1 + \dfrac{3V_m(K_c - K_m)}{3K_m + 4G_m}} \quad \text{and} \quad \frac{G - G_m}{G_c - G_m} = \frac{V_c}{1 + \dfrac{V_m(G_c - G_m)}{G_m + f_m}} \tag{2.14}$$

where the subscripts m and c refer to the metallic and ceramic materials, K is the homogenized bulk modulus, G is the homogenized shear modulus, and the parameter f_m is given by:

$$f_m = \frac{G_m\,(9\,K_m + 8\,G_m)}{6(K_m + 2G_m)} \tag{2.15}$$

Then, the homogenized Young's modulus and Poisson's coefficient can be computed from:

$$E = \frac{9KG}{3K+G} \quad \text{and} \quad \nu = \frac{3K-2G}{2\,(3K+G)} \tag{2.16}$$

Thermal properties, on the other hand, may be evaluated by[8]:

$$\frac{\alpha-\alpha_m}{\alpha_c-\alpha_m} = \frac{(1/K)-(1/\mathrm{K}_m)}{(1/\mathrm{K}_c)-(1/\mathrm{K}_m)} \quad \text{and} \quad \frac{\kappa-\kappa_m}{\kappa_c-\kappa_m} = \frac{V_c}{1+(1-V_c)\,[(\kappa_c-\kappa_m)\,/3\kappa_c]} \tag{2.17}$$

where α is the homogenized thermal expansion coefficient, and κ is the homogenized thermal conductivity. Other properties, such as the material density, are often estimated by the Voigt model.

2.2.3 Constitutive equations

Since FG composites are formed by the inclusion of small particulate reinforcement, they often present isotropic non-homogeneous behavior. Thus, constitutive equations are very similar to that of isotropic materials, but with material properties changing along the structure. For 3D models, the constitutive matrix is given by:

$$\boldsymbol{\sigma} = \mathbf{C}\,\boldsymbol{\varepsilon} \quad \Rightarrow \quad \begin{Bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \tau_{12} \\ \tau_{13} \\ \tau_{23} \end{Bmatrix} = \begin{bmatrix} 2\,\mu+\lambda & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & 2\,\mu+\lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & 2\,\mu+\lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{bmatrix} \begin{Bmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_z \\ \gamma_{xy} \\ \gamma_{xz} \\ \gamma_{yz} \end{Bmatrix} \tag{2.18}$$

where λ and μ are the Lamé constants, given by:

$$\lambda = \frac{E\,\nu}{(1+\nu)\,(1-2\,\nu)} \text{ and } \mu = \frac{E}{2\,(1+\nu)} \tag{2.19}$$

where E and ν are the homogenized Young's modulus and Poisson's ratio, which can be evaluated by Eq. (2.16). Note that, for FGMs, material properties vary in the structure's domain. For 2D models, considering the FSDT, constitutive equations may be defined as shown in Eq. (2.2), but assessing the coefficients Q_{ij} by:

$$Q_{11} = Q_{22} = \frac{E}{1-\nu^2}, \quad Q_{12} = \frac{\nu\,E}{1-\nu^2}, \quad Q_{44} = Q_{55} = Q_{66} = G = \frac{E}{2\,[1+\nu]} \tag{2.20}$$

Similar to laminated composites, the failure of a FGM may be defined by the Tsai-Hill criterion[85, 96]. However, the method is more appropriate for anisotropic materials. Alternatively, one may use the Tamura-Tomota-Ozawa (TTO) model to define the yield stress, as proposed by Tamura[97]. The model assumes that the overall failure of the material is governed by the ductile phase, which is a reasonable consideration in the case of FGMs made of ceramic and metal[9, 98]. The method defines the yield stress by:

$$\sigma_y = \sigma_{y,m}\left[V_m + \left(\frac{q+E_m}{q+E_c}\right)\frac{E_c}{E_m}(1-V_m)\right] \tag{2.21}$$

where $\sigma_{y,m}$ is the ductile material equivalent stress, and q is a stress transfer parameter which depends on the constituent materials. For example, for Ni-Al_2O_3[99] and TiB-Ti[100], $q = 4.5$ GPa, while, for Al-SiC[101], $q = 91.6$ GPa. The Safety Factor (SF) can be computed by:

$$\text{SF} = \frac{\sigma_y}{\sigma_{vm}}, \quad \sigma_{vm} = \sqrt{\sigma_1^2 + \sigma_2^2 - \sigma_1\sigma_2 + 3\tau_{12}^2 + 3\tau_{13}^2 + 3\tau_{23}^2} \tag{2.22}$$

where σ_{vm} is the equivalent von Mises stress. The TTO model also allows for the consideration of material non-linearity on FGMs by evaluating the plastic tangent modulus H by:

$$H = \frac{V_m H_m \dfrac{q+E_c}{q+H_m} + (1-V_m)E_c}{V_m \dfrac{q+E_c}{q+H_m} + (1-V_m)} \tag{2.23}$$

where H_m is the ductile material tangent modulus. It should be noted that studies focused on the material nonlinearity on FGMs are still very scarce in the literature[98].

3 STRUCTURAL ANALYSIS

In this work, the analysis of functionally graded and laminate structures will be performed via the Isogeometric Analysis (IGA). This is an approximated method employed to solve solid mechanics problems described by differential equations, such as static, dynamic, buckling, and modal analysis[102]. For that end, geometrically nonlinear analysis of plates and shallow-shells will be conducted considering both 3D and 2D theories, where the latter provides an approximate response that is accurate when the thickness-to-length ratio is sufficiently small. Isogeometric Analysis (IGA) and the structural theories employed will be briefly discussed in the following.

3.1 Isogeometric Analysis

Isogeometric Analysis (IGA) was first proposed by Hughes, Cottrell and Bazilevs[102] as a way to model an exact geometry described by Non-Uniform Rational B-Splines (NURBS). These functions were already vastly employed in Computer Aided Design (CAD), and IGA extends them to Computer Aided Engineering (CAE), where NURBS are also employed to approximate the displacement field. The method can represent exact geometries even using coarse meshes and mesh refinement is straightforward.

The method has been employed to structural analysis using 3D theories [103, 104], as well as various plate theories, such as the Classical Plate Theory (CPT)[104], the First-order Shear Deformation Theory (FSDT)[82, 105, 17, 2], and Higher-order Shear Deformation Theories (HSDT)[106, 1].

3.1.1 B-Splines

A B-Spline curve is defined by the linear combination of basis functions $B_{i,p}$ and control points $\mathbf{p}_i$:

$$C(\xi) = \sum_{i=1}^{n_b} B_{i,p}(\xi)\,\mathbf{p}_i \tag{3.1}$$

where n_b is the number of basis functions, p is the basis degree, and ξ is the parametric coordinate. The B-Spline curve depends on the location of the control points and the spatial distribution of the parametric coordinates.

A knot span is defined by the knot vector, which contains non-negative and non-

decreasing parametric values bounded by the interval in which the B-Spline is defined. Given a knot vector $\Xi = [\xi_1, \xi_2, ..., \xi_{n+p+1}]$ and considering a parametric interval in the range $[0, 1]$, the B-Spline basis functions can be evaluated by the recursive Cox-de Boor formula[92]:

$$B_{i,p}(\xi) = \frac{\xi - \xi_i}{\xi_{i+p} - \xi_i} B_{i,p-1}(\xi) + \frac{\xi_{i+p+1} - \xi}{\xi_{i+p+1} - \xi_{i+1}} B_{i+1,p-1}(\xi)$$
$$B_{i,0}(\xi) = \begin{cases} 1, & \xi_i \leq \xi < \xi_{i+1} \\ 0, & \text{otherwise} \end{cases} \tag{3.2}$$

where $p \geq 1$. The number of times a value is repeated in the knot vector defines the knot multiplicity m. In that parametric coordinate, the curve is C^{p-m} continuous. Usually, the knots are equally spaced in the parametric space, except for the first and last parametric values, which are often repeated $p+1$ times ($m = p+1$), which means that these points are interpolated. The size of the knot vector n_k also defines the number of basis n_b (which is the same as the number of control points):

$$n_b = n_k - p - 1 \tag{3.3}$$

Piegl and Tiller[92] point out some important characteristics of B-Spline functions, such as:

(a) Non-negative: $B_{i,p}(\xi) \geq 0$.

(b) Partition of unity: $\sum_{i=1}^{n_b} B_{i,p}(\xi) = 1.0$.

(c) Compact suport: Each basis $B_{i,p}(\xi)$ only contributes in the interval $[\xi_i, \xi_{i+p+1}]$. Otherwise, $B_{i,p}(\xi) = 0$.

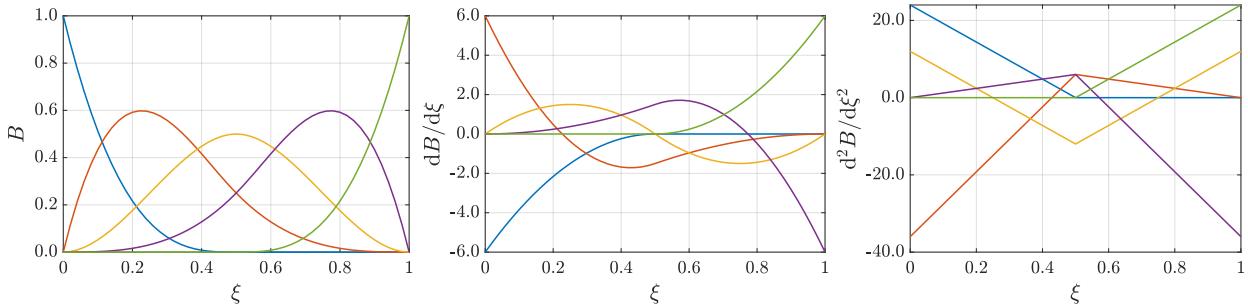
(d) In each knot span, only $p+1$ basis are non-zero.

(e) All derivatives of $B_{i,p}$ exist inside the knot spans, but, at the knots, bases are only $m - p$ differentiable.

Figure 5 depicts the B-Spline basis functions and their derivatives for a knot vector $\Xi = [0, 0, 0, 0, 0.5, 1, 1, 1, 1]$ and $p = 3$. Since $n_k = 9$, there are $n_b = 5$ basis. Note how the basis functions have only C^2 continuity, as the second derivative is non-differentiable at $\xi = 0.5$.

Similar to the conventional Finite Element Method (FEM), B-Splines allow for the p-refinement, related to the degree elevation, and for the h-refinement, related to the knot insertion. However, B-Splines also allow for a third refinement, known as the k-refinement, which has no analog in FEM and is related to the increase in the multiplicity of unique knots. The k-refinement is given by a subsequent combination of the h and p refinements, in that order,

Figure 5 – B-Spline basis functions and their derivatives.

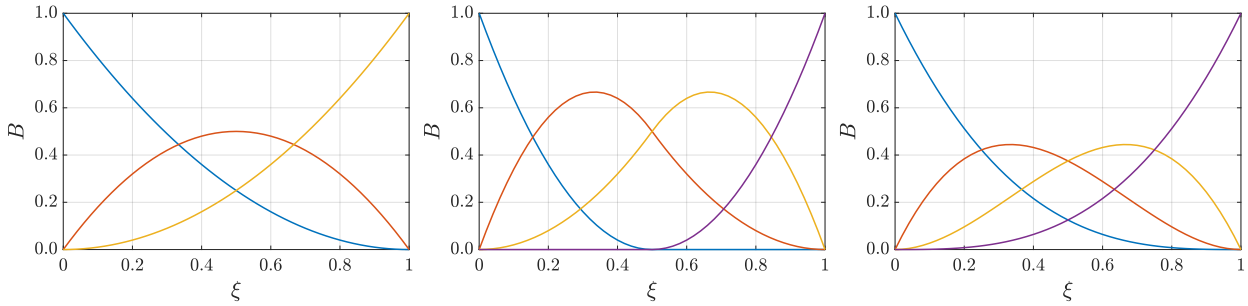


(a) Basis functions $B_{i,p}$ (b) First derivatives (c) Second derivatives

Source: the author

and takes advantage of the fact that these processes do not commute[102]. Figure 6 depicts the h and k-refinements for an initial set of basis functions defined by $\Xi = [0,0,0,1,1,1]$ and $p = 2$.

Figure 6 – B-Spline refinements.



(a) Initial basis functions (b) h-refinement (c) k-refinement

Source: the author

B-Spline functions can also be employed to define the volume fraction gradation on FGMs, as described in Section 2.2.1. In that context, to ensure a smooth gradation, this work will employ cubic B-Splines and open knot vectors with equally spaced interior knots[2]. It is important to note that, for a gradation in multiple directions, one may use bivariate or trivariate B-Splines, which are given by the tensor product between two or three univariate basis[18].

3.1.2 NURBS

Since B-Spline basis functions are polynomials, they are not able to exactly represent conical geometries (e.g. circles and ellipses). This issue can be solved by employing Non-Uniform Rational B-Splines (NURBS), which are defined by the linear combination of rational basis $R_{i,p}$ and control points $\mathbf{p}_i$:

$$C(\xi) = \sum_{i=1}^{n_b} R_{i,p}(\xi)\,\mathbf{p}_i \tag{3.4}$$

The rational basis are given in terms of the basis function $B_{i,p}$ by:

$$R_{i,p}(\xi) = \frac{w_i B_{i,p}(\xi)}{\sum_{i=1}^{n_b} w_i B_{i,p}(\xi)} \tag{3.5}$$

where w_i is the basis weight, which weight function $W(\xi)$. Note that, if weights $w_i = 1.0$, NURBS become the same as usual B-Spline basis functions.

3.2 3D analysis

In general, every structure can be analyzed using a 3D model based on the continuum mechanics equations, although these might not be the most efficient way of performing its analysis. A NURBS solid can be defined as a tensor product between three univariate bases:

$$V(\xi,\eta,\zeta) = \sum_{i=1}^{n_{b1}} \sum_{j=1}^{n_{b2}} \sum_{k=1}^{n_{b3}} R_{ijk,p}(\xi,\eta,\zeta)\,\mathbf{p}_{ijk} \tag{3.6}$$

where $R_{ij,p}(\xi,\eta,\zeta)$ is the trivariate rational basis, given by:

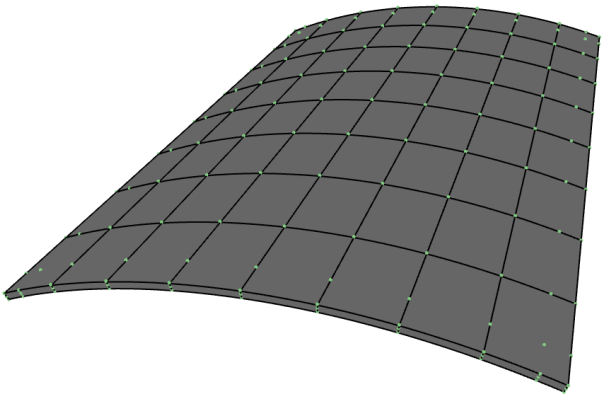
$$R_{ijk,p} = \frac{w_{ijk} N_{i,p}(\xi) N_{j,q}(\eta) N_{k,l}(\zeta)}{\sum_{i=1}^{n_{b1}} \sum_{j=1}^{n_{b2}} \sum_{k=1}^{n_{b3}} w_{ijk} N_{i,p}(\xi) N_{j,q}(\eta) N_{k,l}(\zeta)} \tag{3.7}$$

Thus, the geometry of a solid model can be given by:

$$x = \sum_{i=1}^{n_{cp}} R_i x_i, \quad y = \sum_{i=1}^{n_{cp}} R_i y_i, \quad z = \sum_{i=1}^{n_{cp}} R_i z_i \tag{3.8}$$

where n_{cp} is the number of control points. For instance, Figure 7 shows how a 3D isogeometric model may be used to represent a shallow shell.

Figure 7 – Example of a 3D model for a shallow shell.



Source: the author

In this model, the displacement field is described considering three components[107]:

$$\mathbf{u} = \begin{Bmatrix} u(x,y,z) \\ v(x,y,z) \\ w(x,y,z) \end{Bmatrix} \tag{3.9}$$

and the Green-Lagrange strains are given by[104]:

$$\boldsymbol{\varepsilon} = \begin{Bmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_z \\ \gamma_{xy} \\ \gamma_{xz} \\ \gamma_{yz} \end{Bmatrix} = \underbrace{\begin{Bmatrix} u_{,x} \\ v_{,y} \\ w_{,z} \\ u_{,y} + v_{,x} \\ u_{,z} + w_{,x} \\ v_{,z} + w_{,y} \end{Bmatrix}}_{\boldsymbol{\varepsilon}_0} + \underbrace{\frac{1}{2} \begin{Bmatrix} u_{,x}^2 + v_{,x}^2 + w_{,x}^2 \\ u_{,y}^2 + v_{,y}^2 + w_{,y}^2 \\ u_{,z}^2 + v_{,z}^2 + w_{,z}^2 \\ 2\left(u_{,x}u_{,y} + v_{,x}v_{,y} + w_{,x}w_{,y}\right) \\ 2\left(u_{,x}u_{,z} + v_{,x}v_{,z} + w_{,x}w_{,z}\right) \\ 2\left(u_{,y}u_{,z} + v_{,y}v_{,z} + w_{,y}w_{,z}\right) \end{Bmatrix}}_{\boldsymbol{\varepsilon}_L} \tag{3.10}$$

where $\boldsymbol{\varepsilon}_0$ and $\boldsymbol{\varepsilon}_L$ are the linear and nonlinear terms, respectively.

In this work, we deal with composite materials, where mechanical properties change in the structure. This variation should be considered when evaluating the constitutive matrix $\mathbf{C}$ via Eq. (2.1) or Eq. (2.18). For FG structures, this is performed by evaluating effective material properties as discussed in Section 2.2.2, and then evaluating the $\mathbf{C}$ matrix as show in Section 2.2.3. For laminate structures, due to their orthotropic behavior, the evaluation of the $\mathbf{C}$ matrix is different, as shown in Section 2.1.1.

Using IGA, similar to the coordinates, displacements are given by:

$$u = \sum_{i=1}^{n_{cp}} R_i u_i, \quad v = \sum_{i=1}^{n_{cp}} R_i v_i, \quad w = \sum_{i=1}^{n_{cp}} R_i w_i \tag{3.11}$$

or, in matrix form:

$$\mathbf{u} = \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \sum_{i=1}^{n_{cp}} \begin{bmatrix} R_i & 0 & 0 \\ 0 & R_i & 0 \\ 0 & 0 & R_i \end{bmatrix} = \mathbf{N}\mathbf{u} \tag{3.12}$$

Then, strains can be evaluated by:

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_0 + \boldsymbol{\varepsilon}_L = \mathbf{H}\boldsymbol{\beta} + \frac{1}{2}\mathbf{A}\boldsymbol{\beta} \tag{3.13}$$

where:

$$\boldsymbol{\beta} = \begin{bmatrix} u_{,x} \\ u_{,y} \\ u_{,z} \\ v_{,x} \\ v_{,y} \\ v_{,z} \\ w_{,x} \\ w_{,y} \\ w_{,z} \end{bmatrix} = \sum_{i=1}^{n_p} \begin{bmatrix} R_{i,x} & 0 & 0 \\ R_{i,y} & 0 & 0 \\ R_{i,z} & 0 & 0 \\ 0 & R_{i,x} & 0 \\ 0 & R_{i,y} & 0 \\ 0 & R_{i,z} & 0 \\ 0 & 0 & R_{i,x} \\ 0 & 0 & R_{i,y} \\ 0 & 0 & R_{i,z} \end{bmatrix} \begin{bmatrix} u_1 \\ v_1 \\ w_1 \\ \vdots \\ u_{n_p} \\ v_{n_p} \\ w_{n_p} \end{bmatrix} = \mathbf{G}\,\mathbf{u} \tag{3.14}$$

and:

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} u_{,x} & 0 & 0 & v_x & 0 & 0 & w_x & 0 & 0 \\ 0 & u_{,y} & 0 & 0 & v_{,y} & 0 & 0 & w_{,y} & 0 \\ 0 & 0 & u_{,z} & 0 & 0 & v_{,z} & 0 & 0 & w_{,z} \\ u_{,y} & u_{,x} & 0 & v_{,y} & v_{,x} & 0 & w_{,y} & w_{,x} & 0 \\ u_{,z} & 0 & u_{,x} & v_{,z} & 0 & v_{,x} & w_{,z} & 0 & w_{,x} \\ 0 & u_{,z} & u_{,y} & 0 & v_{,z} & v_{,y} & 0 & w_{,z} & w_{,y} \end{bmatrix} \tag{3.15}$$

Substituting Eq. (3.14) in Eq. (3.13):

$$\boldsymbol{\varepsilon} = \mathbf{H}\,\mathbf{G}\,\mathbf{u} + \frac{1}{2}\mathbf{A}\,\mathbf{G}\,\mathbf{u} = \left(\mathbf{B}_0 + \frac{1}{2}\mathbf{B}_L\right)\mathbf{u} = \mathbf{B}\,\mathbf{u} \tag{3.16}$$

where:

$$\mathbf{B}_0 = \mathbf{H}\,\mathbf{G}, \text{ and } \mathbf{B}_L = \mathbf{A}\,\mathbf{G} \tag{3.17}$$

Thus, strains can be found based on matrix $\mathbf{B}$, which has a linear and a non-linear term ($\mathbf{B}_0$ and $\mathbf{B}_L$, respectively). For the incremental formulation, the strain increment is given by:

$$\delta\hat{\boldsymbol{\varepsilon}} = (\mathbf{B}_0 + \mathbf{B}_L)\,\delta\mathbf{u}_e = \overline{\mathbf{B}}\,\delta\mathbf{u}_e \tag{3.18}$$

3.2.1 Equilibrium equations

The dynamic equilibrium equations of the model can be obtained using the D'Alembert and virtual work principles:

$$\int_V \delta\mathbf{u}^T\,\overline{\mathbf{M}}\,\ddot{\mathbf{u}}\,dV + \int_V \delta\boldsymbol{\varepsilon}^T\,\boldsymbol{\sigma}\,dV = \int_V \delta\mathbf{u}^T\,\mathbf{b}\,dV + \int_S \delta\mathbf{u}^T\,\mathbf{q}\,dS \tag{3.19}$$

where $\delta\boldsymbol{\varepsilon}$ is the virtual displacement vector, $\boldsymbol{\sigma}$ is the stress vector, and $\mathbf{q}$ and $\mathbf{b}$ are the surface and body loads, respectively. For solid elements, $\overline{\mathbf{M}}$ is given by:

$$\overline{\mathbf{M}} = \begin{bmatrix} \rho & 0 & 0 \\ 0 & \rho & 0 \\ 0 & 0 & \rho \end{bmatrix} \tag{3.20}$$

Displacements can be given in terms of the displacements in the control points, as described in Eq. (3.36). Thus, considering arbitrary displacements, Eq. (3.19) can be written as:

$$\mathbf{M}\,\ddot{\mathbf{u}}_e + \mathbf{g}(\mathbf{u}_e) = \mathbf{f}(t) \tag{3.21}$$

where $\mathbf{M}$ is the mass matrix, $\mathbf{g}(\mathbf{u}_e)$ is the internal forces vector, which depends on $\mathbf{u}_e$ due to the nonlinear term, and $\mathbf{f}(t)$ is the external load vector. These terms are given by:

$$\begin{aligned} \mathbf{M} &= \int_V \mathbf{R}^T\,\overline{\mathbf{M}}\,\mathbf{R}\,\mathrm{d}V \\ \mathbf{g} &= \int_V \overline{\mathbf{B}}^T\,\hat{\boldsymbol{\sigma}}\,\mathrm{d}V \\ \mathbf{f} &= \int_V \mathbf{R}^T\,\mathbf{b}\,\mathrm{d}V + \int_S \mathbf{R}^T\,\mathbf{q}\,\mathrm{d}S \end{aligned} \tag{3.22}$$

The nonlinear problem can be solved by an incremental and iterative approach, such as the Newton-Raphson method. In each step, for a given displacement $\mathbf{u}_e$, the stiffness matrix is given by the differentiation of the internal forces vector:

$$\mathbf{K}_T = \frac{\partial \mathbf{g}}{\partial \mathbf{u}} = \mathbf{K}_L + \mathbf{K}_\sigma \tag{3.23}$$

where the material stiffness matrix $\mathbf{K}_L$ and the geometric stiffness matrix $\mathbf{K}_\sigma$ are given by:

$$\begin{aligned} \mathbf{K}_L &= \int_A \overline{\mathbf{B}}^T \frac{\partial \hat{\boldsymbol{\sigma}}}{\partial \mathbf{u}_e}\,\mathrm{d}A = \int_A \overline{\mathbf{B}}^T\,\mathbf{C}\,\overline{\mathbf{B}}\,\mathrm{d}A \\ \mathbf{K}_\sigma &= \int_A \frac{\partial \overline{\mathbf{B}}^T}{\partial \mathbf{u}_e}\,\boldsymbol{\sigma}\,\mathrm{d}A = \int_A \mathbf{G}^T\,\mathbf{S}\,\mathbf{G}\,\mathrm{d}A \end{aligned} \tag{3.24}$$

where $\mathbf{C}$ is the constitutive matrix, presented in Sections 2.1.1 and 2.2.3 for laminate and FG composites respectively. For a more efficient implementation, instead of computing the integral in the entire volume, one may perform integration in the isogeometric element, defined by the knot span. This is similar to what is performed in FEM for the assembly of the global stiffness matrix. More details can be found in Barroso[104].

For displacement-independent loads, the nonlinear equilibrium can be written as:

$$\mathbf{r}(\mathbf{u},\lambda) = \mathbf{g}(\mathbf{u}) - \mathbf{f} = \mathbf{g}(\mathbf{u}) - \lambda\,\mathbf{q} \tag{3.25}$$

where $\mathbf{q}$ is a reference load vector and λ is the load factor. The equation is solved in each step for $\mathbf{r} = 0$ using an appropriate path-following method, such as the Load Control, Displacements Control, or the Arc-Length Method, which are based on Newton-Raphson method iterations. The detailed methodology can be found in Praciano[108].

3.2.2 Eigenvalue problems

For some structures, we may solve buckling and vibration problems based on the initial geometry by the linearization of the problem. These are performed by the solution of a generalized eigenvalue problem.

For structures with negligible pre-buckling displacements, the stability analysis can be solved by the generalized eigenproblem[107]:

$$\left(\mathbf{K}_L + \lambda\,\overline{\mathbf{K}}_\sigma\right)\boldsymbol{\phi} = \mathbf{0} \tag{3.26}$$

where $\boldsymbol{\phi}$ are the buckling modes, represented by the eigenvectors, and the buckling loads are given by the eigenvalues λ. The geometric stiffness matrix should be defined according to a reference load vector, where:

$$\mathbf{K}_\sigma = \lambda\,\overline{\mathbf{K}}_\sigma \tag{3.27}$$

It should be noted that, when composites are not symmetric, the membrane-bending coupling should be considered. Thus, the application of an axial load may incur relevant transverse displacements, and the linearized theory should not be considered.

In the same way, the free vibration eigenproblem can be solved by:

$$\left(\mathbf{K}_L + \omega^2\,\mathbf{M}\right)\boldsymbol{\phi} = \mathbf{0} \tag{3.28}$$

where $\boldsymbol{\phi}$ are the vibration modes and ω are the natural frequencies.

3.3 First-order Shear Deformation Theory

The First-order Shear Deformation Theory (FSDT) is a simplified theory for the analysis of 2D structures, such as plates and shallow shells. For the representation of those, one must use a NURBS surface, which can be obtained by the tensor product between two univariate bases:

$$S(\xi,\eta) = \sum_{i=1}^{n_{b1}}\sum_{j=1}^{n_{b2}} R_{ij,p}(\xi,\eta)\,\mathbf{p}_{ij} \tag{3.29}$$

where $R_{ij,p}(\xi,\eta)$ is the bivariate rational basis, given by:

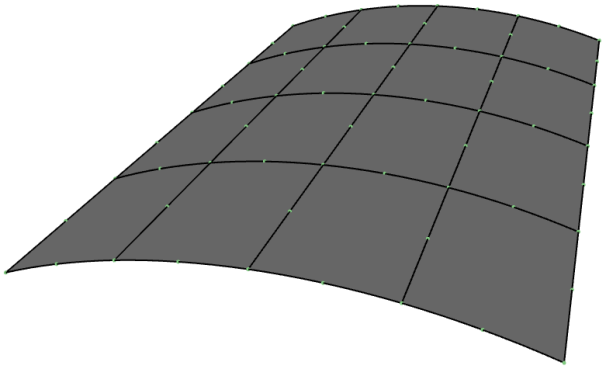
$$R_{ij,p} = \frac{w_{ij} N_{i,p}(\xi) N_{j,q}(\eta)}{\sum_{i=1}^{n_{b1}} \sum_{j=1}^{n_{b2}} w_{ij} N_{i,p}(\xi) N_{j,q}(\eta)} \tag{3.30}$$

For the representation of plates and shallow-shells, coordinates are given by:

$$x = \sum_{i=1}^{n_{cp}} R_i x_i, \quad y = \sum_{i=1}^{n_{cp}} R_i y_i, \quad z_0 = \sum_{i=1}^{n_{cp}} R_i z_{0i} \tag{3.31}$$

where n_{cp} is the number of control points. For plates with no imperfections, $z_0 = 0$. For instance, Figure 8 shows how a 2D isogeometric model can represent a shallow shell. Here, since there are no elements in the thickness of the structure, variation of displacements and stresses in the thickness direction may not be captured correctly, especially for high thickness-to-length ratios.

Figure 8 – Example of a 2D model for a shallow shell.



Source: the author

In this work, the First-order Shear Deformation Theory (FSDT) will be considered, where shear strains are considered to vary linearly in the thickness direction[108, 3]. According to the First-order Shear Deformation Theory (FSDT), also known as the Reissner-Mindlin plate theory, normal lines to the mid-surface remain straight, but not necessarily perpendicular to the mid-surface. Considering small strains and rotations, displacements in any point of the plate can be written as:

$$\begin{Bmatrix} \overline{u} \\ \overline{v} \\ \overline{w} \end{Bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & z \\ 0 & 1 & 0 & -z & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{Bmatrix} u \\ v \\ w \\ \theta_x \\ \theta_y \end{Bmatrix} \Rightarrow \overline{\mathbf{u}} = \mathbf{Z}\mathbf{u} \tag{3.32}$$

where u, v, and w are the mid-surface displacements, and θ_x and θ_y are the rotations about x and y axes. The coordinate z is the distance from a given point to the mid-surface.

Based on the displacement field and on the Marguerre theory, in-plane strains are given by:

$$\boldsymbol{\varepsilon} = \begin{Bmatrix} \varepsilon_x \\ \varepsilon_y \\ \gamma_{xy} \end{Bmatrix} = \boldsymbol{\varepsilon}^m + z\,\boldsymbol{\kappa} \tag{3.33}$$

where the superscript m refers to the membrane strains and κ to the curvature. In matricial form, the membrane strains are given by[82]:

$$\boldsymbol{\varepsilon}^m = \begin{Bmatrix} \varepsilon_x^m \\ \varepsilon_y^m \\ \gamma_{xy}^m \end{Bmatrix} = \underbrace{\begin{Bmatrix} u_{,x} \\ v_{,y} \\ u_{,y} + v_{,x} \end{Bmatrix} + \begin{Bmatrix} w_{,x}\, z_{0,x} \\ w_{,y}\, z_{0,y} \\ w_{,x}\, z_{0,y} + w_{,y}\, z_{0,x} \end{Bmatrix}}_{\boldsymbol{\varepsilon}_0^m} + \underbrace{\frac{1}{2} \begin{Bmatrix} w_{,x}^2 \\ w_{,y}^2 \\ 2\, w_{,x}\, w_{,y} \end{Bmatrix}}_{\boldsymbol{\varepsilon}_L^m} \tag{3.34}$$

Bending and transverse shear strains are given by:

$$\begin{aligned} \boldsymbol{\kappa} &= \begin{Bmatrix} \kappa_x \\ \kappa_y \\ \kappa_{xy} \end{Bmatrix} = \begin{Bmatrix} \theta_{y,x} \\ -\theta_{x,y} \\ \theta_{y,y} - \theta_{x,x} \end{Bmatrix} \\ \boldsymbol{\gamma} &= \begin{Bmatrix} \gamma_{xz} \\ \gamma_{yz} \end{Bmatrix} = \begin{Bmatrix} w_{,x} + \theta_y \\ w_{,y} - \theta_x \end{Bmatrix} \end{aligned} \tag{3.35}$$

Thus, both the curvatures and the transverse shear strains are constant through the shell thickness. However, by the Elasticity Theory, this is not true. Thus, the Reissner-Mindlin theory considers these in a simplified way, which is still better than the Classical Plate Theory (CPT), where $\boldsymbol{\gamma} = \mathbf{0}$.

Using IGA, displacements are represented by:

$$u = \sum_{i=1}^{n_{cp}} R_i\, u_i, \quad v = \sum_{i=1}^{n_{cp}} R_i\, v_i, \quad w = \sum_{i=1}^{n_{cp}} R_i\, w_i, \quad \theta_x = \sum_{i=1}^{n_{cp}} R_i\, \theta_{x,i}, \quad \theta_y = \sum_{i=1}^{n_{cp}} R_i\, \theta_{y,i} \tag{3.36}$$

or, in matrix notation, by:

$$\mathbf{u} = \mathbf{R}\,\mathbf{u}_e \tag{3.37}$$

where $\mathbf{u}_e$ is the vector of degrees of freedom, which are the displacements at control points, and $\mathbf{R}$ is the matrix of shape functions:

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{R}_2 & \dots & \mathbf{R}_{nc} \end{bmatrix} \tag{3.38}$$

where

$$\mathbf{R}_k = R_k \, \mathbf{I}_{5\times 5} \tag{3.39}$$

where $\mathbf{I}$ is the identity matrix.

For the analysis of 2D structures using the FSDT, generalized strains can be obtained similar to those from 3D models, as a function of the degrees of freedom. Thus, rewriting Eq. (3.33) in matricial form:

$$\hat{\boldsymbol{\varepsilon}} = \begin{bmatrix} \varepsilon_0^m \\ \kappa \\ \gamma \end{bmatrix} + \begin{bmatrix} \varepsilon_L^m \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{B}_0^m \\ \mathbf{B}_0^b \\ \mathbf{B}_0^s \end{bmatrix} \mathbf{u}_e + \frac{1}{2} \begin{bmatrix} \mathbf{B}_L^m \\ 0 \\ 0 \end{bmatrix} \mathbf{u}_e = \left(\mathbf{B}_0 + \frac{1}{2} \mathbf{B}_L \right) \mathbf{u}_e = \mathbf{B} \, \mathbf{u}_e \tag{3.40}$$

By the Marguerre theory, these matrices are given by[82]:

$$\begin{aligned}
\mathbf{B}_0^m &= \begin{bmatrix} R_{k,x} & 0 & Z_x R_{k,x} & 0 & 0 \\ 0 & R_{k,y} & Z_y R_{k,y} & 0 & 0 \\ R_{k,y} & R_{k,x} & Z_x R_{k,y} + Z_y R_{k,x} & 0 & 0 \end{bmatrix} \\
\mathbf{B}_0^b &= \begin{bmatrix} 0 & 0 & 0 & 0 & R_{k,x} \\ 0 & 0 & 0 & -R_{k,y} & 0 \\ 0 & 0 & 0 & -R_{k,x} & R_{k,y} \end{bmatrix} \\
\mathbf{B}_0^s &= \begin{bmatrix} 0 & 0 & R_{k,x} & 0 & R_k \\ 0 & 0 & R_{k,y} & -R_k & 0 \end{bmatrix} \\
\mathbf{B}_L^m &= \begin{bmatrix} 0 & 0 & W_x R_{k,x} & 0 & 0 \\ 0 & 0 & W_y R_{k,y} & 0 & 0 \\ 0 & 0 & W_x R_{k,y} + W_y R_{k,x} & 0 & 0 \end{bmatrix}
\end{aligned} \tag{3.41}$$

where:

$$Z_x = \sum_{k=1}^{n_{cp}} R_{k,x} \, z_{0,k}, \quad Z_y = \sum_{k=1}^{n_{cp}} R_{k,y} \, z_{0k} \quad W_x = \sum_{k=1}^{n_{cp}} R_{k,x} \, w_k, \quad W_y = \sum_{k=1}^{n_{cp}} R_{k,y} \, w_k \tag{3.42}$$

It is important to note that $\mathbf{B}_L$ depends on the displacements w, thus denoting a nonlinear behavior. This is similar to the solid elements. This feature allows for the nonlinear analysis and the study of the stability of such structures. For the incremental formulation, the strain increment is given by:

$$\delta \hat{\boldsymbol{\varepsilon}} = (\mathbf{B}_0 + \mathbf{B}_L) \, \delta \mathbf{u}_e = \overline{\mathbf{B}} \, \delta \mathbf{u}_e \tag{3.43}$$

3.3.1 Internal forces

For plates and shallow shells, internal forces can be written in terms of membrane, shear, and bending strains, resulting in the matrices **A**, **B**, **D**, and **G**. The evaluation of internal forces requires a stress-strain relation, which was presented in Sections 2.1.1 and 2.2.3 for laminated and FG composites, respectively. Since these present important differences, they will be described in different sections.

3.3.1.1 Laminated composites

Forces and moments can be evaluated by integrating stresses over the structure thickness h. In laminated composites, each ply is considered as a homogeneous and orthotropic material. Due to the discontinuous change in material properties, this may be performed by a summation:

$$\begin{aligned}
\mathbf{N} &= \begin{Bmatrix} N_x \\ N_y \\ N_{xy} \end{Bmatrix} = \int_{-h/2}^{h/2} \begin{Bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{Bmatrix} \mathrm{d}z = \sum_{k=1}^{N_p} \int_{z_k}^{z_{k+1}} \overline{\mathbf{Q}}^{(k)} \boldsymbol{\varepsilon}^{(k)} \, \mathrm{d}z \\
\mathbf{M} &= \begin{Bmatrix} M_x \\ M_y \\ M_{xy} \end{Bmatrix} = \int_{-h/2}^{h/2} \begin{Bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{Bmatrix} z \, \mathrm{d}z = \sum_{k=1}^{N_p} \int_{z_k}^{z_{k+1}} \overline{\mathbf{Q}}^{(k)} \boldsymbol{\varepsilon}^{(k)} \, z \, \mathrm{d}z \\
\mathbf{V} &= \begin{Bmatrix} V_{xz} \\ V_{yz} \end{Bmatrix} = \int_{-h/2}^{h/2} \begin{Bmatrix} \tau_{xz} \\ \tau_{yz} \end{Bmatrix} \mathrm{d}z = k_s \sum_{k=1}^{N_p} \int_{z_k}^{z_{k+1}} \overline{\mathbf{Q}}_s^{(k)} \boldsymbol{\gamma}^{(k)} \, \mathrm{d}z
\end{aligned} \tag{3.44}$$

where k refers to the k-th ply, and k_s is a shear correction factor, employed to adjust the shear factor to a value closer to the parabolic shear-stress distribution from the Elasticity Theory. Usually, k_s is considered as $5/6$, as derived by Reissner[109] for isotropic rectangular plates.

By considering membrane, bending and shear strains separately, we may define the

constitutive relation by:

$$\begin{Bmatrix} N_x \\ N_y \\ N_{xy} \\ M_x \\ M_y \\ M_{xy} \\ V_{xz} \\ V_{yz} \end{Bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{16} & B_{11} & B_{12} & B_{16} & 0 & 0 \\ A_{21} & A_{22} & A_{26} & B_{21} & B_{22} & B_{26} & 0 & 0 \\ A_{61} & A_{62} & A_{66} & B_{61} & B_{62} & B_{66} & 0 & 0 \\ B_{11} & B_{12} & B_{16} & D_{11} & D_{12} & D_{16} & 0 & 0 \\ B_{21} & B_{22} & B_{26} & D_{21} & D_{22} & D_{26} & 0 & 0 \\ B_{61} & B_{62} & B_{66} & D_{61} & D_{62} & D_{66} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & G_{44} & G_{45} \\ 0 & 0 & 0 & 0 & 0 & 0 & G_{54} & G_{55} \end{bmatrix} \begin{Bmatrix} \varepsilon_x^m \\ \varepsilon_y^m \\ \gamma_{xy}^m \\ \kappa_x \\ \kappa_y \\ \kappa_{xy} \\ \gamma_{xz} \\ \gamma_{yz} \end{Bmatrix} \tag{3.45}$$

which can be summarized by the **ABDG** matrix:

$$\begin{Bmatrix} \mathbf{N} \\ \mathbf{M} \\ \mathbf{V} \end{Bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{0} \\ \mathbf{B} & \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{G} \end{bmatrix} \begin{Bmatrix} \boldsymbol{\varepsilon}^m \\ \boldsymbol{\kappa} \\ \boldsymbol{\gamma} \end{Bmatrix} \Rightarrow \hat{\boldsymbol{\sigma}} = \mathbf{C}\,\hat{\boldsymbol{\varepsilon}} \tag{3.46}$$

where **A** is the membrane stiffness matrix, **D** is the bending stiffness matrix, **B** is the membrane-bending coupling matrix, and **G** is the shear stiffness matrix. These terms may be evaluated by numerical or analytical integration:

$$\begin{aligned} \mathbf{A} &= \sum_{k=1}^{N_p} \int_{z_k}^{z_{k+1}} \overline{\mathbf{Q}}^{(k)}\,\mathrm{d}z \Rightarrow A_{ij} = \sum_{k=1}^{N_p} \overline{Q}_{ij}^{(k)}\,(z_{k+1} - z_k) \\ \mathbf{B} &= \sum_{k=1}^{N_p} \int_{z_k}^{z_{k+1}} \overline{\mathbf{Q}}^{(k)}\,z\,\mathrm{d}z \Rightarrow B_{ij} = \sum_{k=1}^{N_p} \frac{1}{2}\overline{Q}_{ij}^{(k)}\,\left(z_{k+1}^2 - z_k^2\right) \\ \mathbf{D} &= \sum_{k=1}^{N_p} \int_{z_k}^{z_{k+1}} \overline{\mathbf{Q}}^{(k)}\,z^2\,\mathrm{d}z \Rightarrow D_{ij} = \sum_{k=1}^{N_p} \frac{1}{3}\overline{Q}_{ij}^{(k)}\,\left(z_{k+1}^3 - z_k^3\right) \\ \mathbf{G} &= \sum_{k=1}^{N_p} \int_{z_k}^{z_{k+1}} k_s\,\overline{\mathbf{Q}}_s^{(k)}\,\mathrm{d}z \Rightarrow G_{ij} = \sum_{k=1}^{N_p} k_s\,\overline{Q}_{s,ij}^{(k)}\,(z_{k+1} - z_k) \end{aligned} \tag{3.47}$$

where $i, j = 4,5$ for the **G** matrix and $i, j = 1,2,6$ for all other matrices. For symmetric laminates, there is no membrane-bending coupling, and $\mathbf{B} = \mathbf{0}$.

3.3.1.2 Functionally graded composites

FG composites are considered as non-homogeneous, isotropic materials, where properties vary through the thickness due to the change in volume fractions. Thus, internal forces

may be computed by through-thickness integration:

$$\begin{aligned}
\mathbf{N} &= \begin{Bmatrix} N_x \\ N_y \\ N_{xy} \end{Bmatrix} = \int_{-h/2}^{h/2} \begin{Bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{Bmatrix} \mathrm{d}z = \int_{-h/2}^{h/2} \mathbf{Q}(z)\, \boldsymbol{\varepsilon}(z)\, \mathrm{d}z \\
\mathbf{M} &= \begin{Bmatrix} M_x \\ M_y \\ M_{xy} \end{Bmatrix} = \int_{-h/2}^{h/2} \begin{Bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{Bmatrix} z\, \mathrm{d}z = \int_{-h/2}^{h/2} \mathbf{Q}(z)\, \boldsymbol{\varepsilon}(z)\, z\, \mathrm{d}z \\
\mathbf{V} &= \begin{Bmatrix} V_{xz} \\ V_{yz} \end{Bmatrix} = \int_{-h/2}^{h/2} \begin{Bmatrix} \tau_{xz} \\ \tau_{yz} \end{Bmatrix} \mathrm{d}z = k_s \int_{-h/2}^{h/2} \mathbf{Q}_s(z)\, \boldsymbol{\gamma}(z)\, \mathrm{d}z
\end{aligned} \tag{3.48}$$

Similar to laminate composites, we may consider membrane, bending and shear terms separately by defining a **ABDG** matrix, as in Eq. (3.46). For FG structures, its components may be evaluated by:

$$\begin{aligned}
\mathbf{A} &= \int_{-h/2}^{h/2} \mathbf{Q}(z)\, \mathrm{d}z \\
\mathbf{B} &= \int_{-h/2}^{h/2} \mathbf{Q}(z)\, z\, \mathrm{d}z \\
\mathbf{D} &= \int_{-h/2}^{h/2} \mathbf{Q}(z)\, z^2\, \mathrm{d}z \\
\mathbf{G} &= \int_{-h/2}^{h/2} k_s\, \mathbf{Q}_s(z)\, \mathrm{d}z
\end{aligned} \tag{3.49}$$

These integrals are usually evaluated numerically, using an appropriate quadrature rule. For a symmetrical material distribution, there is no membrane-bending coupling, and $\mathbf{B} = \mathbf{0}$.

3.3.2 Equilibrium equations

For the FSDT, dynamic equilibrium equations are given by:

$$\int_A \delta\mathbf{u}^T\, \overline{\mathbf{M}}\, \ddot{\mathbf{u}}\, dA + \int_A \delta\hat{\boldsymbol{\varepsilon}}^T\, \hat{\boldsymbol{\sigma}}\, \mathrm{d}A = \int_A \delta\mathbf{u}^T\, \mathbf{q}\, \mathrm{d}A + \int_S \delta\mathbf{u}^T\, \mathbf{f}_s\, \mathrm{d}S \tag{3.50}$$

where $\mathbf{q}$ is the body load, $\mathbf{f}_s$ is the boundary load, A is the mid-surface area and S, the mid-surface boundary of the plate. Here, $\overline{\mathbf{M}}$ is given by:

$$\overline{\mathbf{M}} = \int_{-h/2}^{h/2} \rho(z)\, \mathbf{Z}^T\, \mathbf{Z}\, \mathrm{d}z = \begin{bmatrix} I_0 & 0 & 0 & 0 & I_1 \\ 0 & I_0 & 0 & -I_1 & 0 \\ 0 & 0 & I_0 & 0 & 0 \\ 0 & -I_1 & 0 & I_2 & 0 \\ I_1 & 0 & 0 & 0 & I_2 \end{bmatrix} \tag{3.51}$$

where:

$$[I_0, I_1, I_2] = \int_{-h/2}^{h/2} \rho(z)[1, z, z^2]\,\mathrm{d}z \tag{3.52}$$

This integration can be performed numerically, using the Gaussian quadrature. To avoid shear locking and slightly improve the computational efficiency, computation of the stiffness matrices should also be evaluated using an appropriate reduced integration scheme[82].

Finally, definition of material and geometric stiffness matrix can be performed in the same way as for 3D elements. Solution of eigenvalue problems also follow the same procedure.

3.4 Analysis validation

In this section, we present some numerical examples for functionally graded and laminate composites, comparing structural responses obtained to those from other researchers. Structural analyses will be performed using the in-house software Finite element AnalysiS Tool (FAST). Formulation for unidirectional FGMs and laminates using the FSDT was implemented in FAST on recent works[108, 3]. In this work, we implemented the analysis of tridirectional FG structures considering the FSDT, as well as the analysis of unidirectional and tridirectional FG structures using solid elements.

3.4.1 Functionally graded materials

In this section, analysis of different FGMs will be performed. Table 1 presents the material properties considered here. Thermal properties were evaluated using the Touloukian's equation for $T = 300$ K[85].

Table 1 – Isotropic material properties.

Material	E (GPa)	ν	ρ (kg/m^3)	α (10^{-6}/°C)	κ (W/mK)
Al	70.000	0.3	2707	23.00	229.0
SUS304	201.04	0.3	8166	15.32	12.14
Al_2O_3	380.00	0.3	3800	7.400	64.99
Si_3N_4	348.43	0.3	2370	5.872	83.87

First, the evaluation of the critical buckling load factor of a clamped FG square plate will be performed. The plate has $a/h = 100$, where a is the plate length and h is the plate thickness. A 8×8 cubic NURBS mesh is employed for structural analysis. Material properties are $E_m = 207.79$ GPa, $\nu_m = 0.28$, $E_c = 322.27$ GPa, and $\nu_c = 0.28$. Effective properties are

estimated via the rule of mixtures, while the volume fraction is described by the power-law function shown in Eq. (2.10), with $V_{c,b} = 0.0$ and $V_{c,t} = 1.0$. The non-dimensional buckling load:

$$\lambda_{norm} = \frac{N_{cr}\, a^2}{\pi^2 D_c} \text{ where } D_c = \frac{E_c h^3}{12\,(1-\nu_c^2)} \tag{3.53}$$

is compared with the results found by Bateni, Kiani and Eslami[110] and Ribeiro et al.[2] in Table 2. The obtained results are in excellent agreement with the ones from the reference.

Table 2 – Non-dimensional buckling load of a clamped FG square plate.

p	Bateni, Kiani and Eslami[110]	Ribeiro et al.[2]	This work
0.0	10.057	10.069	10.075
0.5	8.6538	8.6817	8.6876
1.0	8.1424	8.1798	8.1854
2.0	7.7745	7.8136	7.8188
3.0	7.6173	7.6514	7.6564
5.0	7.4275	7.4527	7.4573

The second example deals with a simply supported FG square plate. This time, the plate has $a/h = 10$, where $a = 10$ m, and the non-dimensional buckling load is evaluated as:

$$\lambda_{norm} = \frac{N_{cr}\, a}{D_m} \tag{3.54}$$

Now, the gradation is given by a unidirectional B-Spline with 9 control points, symmetric to the mid-plane. The material properties are $E_m = 70$ GPa, $E_c = 380$ GPa, and $\nu_m = \nu_c = 0.30$. Table 3 shows the results found for different control points sequences. This plate was analyzed by Do, Lee and Lee[1], which used a Higher-order Shear Deformation Theory (HSDT), and by Ribeiro et al.[2], which used the First-order Shear Deformation Theory (FSDT). In this work, the latter was also employed. Moreover, the results for the 3D model are also shown. For the 2D model, a 8×8 cubic NURBS mesh is employed, while the 3D model uses a $10 \times 10 \times 2$ cubic NURBS mesh.

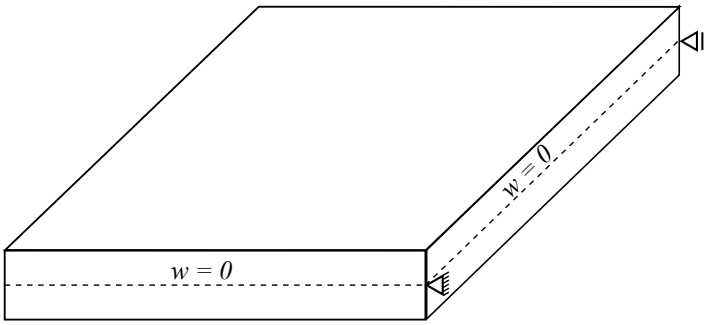
The results from the 3D model differ from those from the 2D models. While this is expected, this difference might be too large. The reason for this lies in the boundary conditions used. Figure 9 shows how the boundary conditions are applied in the 3D model to represent a simply supported plate. Out-of-plane displacements are constrained in the mid-plane, and in-plane displacements are constrained in two corners of the plate.

Now, Figure 10 shows how the boundary conditions can be applied for the 2D model. While most researchers adopt the Simply Supported 1 condition, which is here denominated as

Table 3 – Non-dimensional buckling load of a simply supported FG square plate.

Control point sequence									HSDT[1]	FSDT[2]	This work	
											FSDT	3D
[0	0	0	0	0	0	0	0	0]	-	-	3.7371	3.4042
[1	1	0.4	0	0	0	0.4	1	1]	11.185	11.634	11.633	9.8927
[1	1	1	0	0	0	1	1	1]	14.603	15.357	15.357	12.577
[1	1	1	0.45	0	0.45	1	1	1]	16.193	16.705	16.704	14.091
[1	1	1	1	1	1	1	1	1]	-	-	20.287	18.480

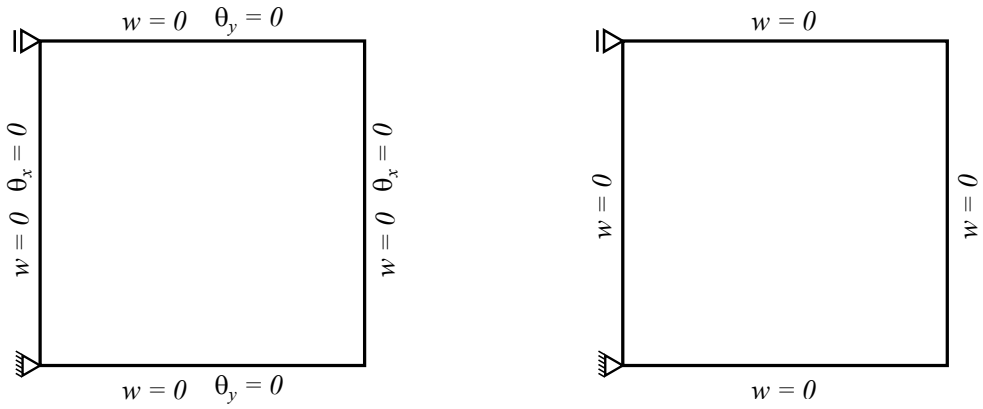
Figure 9 – Boundary conditions for the 3D model.



Source: the author

SS1, there are some constraints that can not be easily represented in a 3D model, namely those related to a constraint in the node rotation. Thus, the Simply Supported 2 condition, denominated as SS2, is better fit to represent the same conditions as the 3D model.

Figure 10 – Boundary conditions for the 2D model



(a) Simply Supported 1 (SS1)

(b) Simply Supported 2 (SS2)

Source: the author

Table 4 shows a comparison between the results from the 2D model (now using the SS2 boundary condition) and the 3D model. This time, the results are closer, as expected. However, due to the different theories employed, there is still an average 8.0% relative difference between these approaches.

Table 4 – Non-dimensional buckling load of a simply supported FG square plate using different boundary conditions.

Control point sequence	This work		
	FSDT (SS2)	3D	Diff. (%)
[0 0 0 0 0 0 0 0 0]	3.4526	3.4042	1.42
[1 1 0.4 0 0 0 0.4 1 1]	10.567	9.8927	6.82
[1 1 1 0 0 0 1 1 1]	13.964	12.577	11.03
[1 1 1 0.45 0 0.45 1 1 1]	15.235	14.091	8.12
[1 1 1 1 1 1 1 1 1]	18.743	18.480	1.42

Now, results for the validation of thermal buckling analyses will be shown. This example deals with a clamped Al/Al_2O_3 FG square plate with varying a/h. Effective properties are estimated using the rule of mixtures, and the volume fraction is described by a power-law function with $V_{c,b} = 0.0$ and $V_{c,t} = 1.0$. This problem is solved by different researchers. Zhao, Lee and Liew[111] use a mesh-free method based on the FSDT, where a different approach is used to evaluate the shear correction factor. Kiani, Bagherizadeh and Eslami[112] use closed-form approximate solutions based on the CPT. Nguyen-Thanh et al.[113] use an IGA formulation based on the Higher-order Shear Deformation Theory (HSDT). Bateni, Kiani and Eslami[110] use a multi-term Galerkin solution which considers a parabolic distribution of shear strains over the structure thickness. Then, Silva[114] uses an IGA formulation based on the FSDT, very similar to the one used in this work.

Table 5 compares the results for this problem. In this work, a 8×8 cubic NURBS mesh was employed. Our results are very close to those presented by Silva[114]. However, as p increases, our responses become more different from those found in other sources. This is likely due to the plate theory considered, the FSDT, where variation of shear strains are considered in a simplified manner.

Finally, a validation for the natural frequency analysis is presented. To this end, a simply-supported $SUS304/Si_3N_4$ FG plate with $a/h = 10$ is studied. Effective material properties are evaluated using the Mori-Tanaka scheme, and volume fraction distribution is defined using the power-law rule. Here, the non-dimensional natural frequency is evaluated as:

$$\omega_{norm} = \omega h \sqrt{\frac{\rho_c}{G_c}} \tag{3.55}$$

Structural analysis is performed using a 8×8 NURBS cubic mesh. Table 6 shows the results found for this example. Here, results are compared to those shown in Nguyen et al.[115], using a generalized HSDT, and to those shown in Ribeiro et al.[2], found using the FSDT. The results

Table 5 – Buckling temperature (°C) of a clamped FG square plate using different a/h.

a/h	Source	p 0.0	0.5	1.0	2.0	5.0
100	Zhao, Lee and Liew[111]	44.17	24.90	20.77	18.48	19.15
	Kiani, Bagherizadeh and Eslami[112]	45.51	25.79	21.15	18.75	19.34
	Nguyen-Thanh et al.[113]	47.50	26.54	21.70	19.18	19.80
	Bateni, Kiani and Eslami[110]	45.28	25.65	21.04	18.65	19.23
	Silva[114]	45.28	26.51	22.76	21.18	21.48
	Present work	45.30	26.64	23.02	21.56	21.81
50	Zhao, Lee and Liew[111]	175.8	99.16	82.35	71.01	74.59
	Kiani, Bagherizadeh and Eslami[112]	182.1	103.2	84.58	74.99	77.36
	Nguyen-Thanh et al.[113]	188.3	105.3	86.07	76.07	78.06
	Bateni, Kiani and Eslami[110]	180.3	102.2	83.84	74.30	76.50
	Silva[114]	180.1	105.5	90.56	84.23	85.34
	Present work	180.2	106.0	91.57	85.73	86.67

are in very good agreement with those found from the literature.

Table 6 – Normalized natural frequency of a FG square plate using different p-exponents.

p	Source Nguyen et al.[115]	Ribeiro et al.[2]	Present work
1	0.0542	0.0545	0.0545
2	0.0485	0.0487	0.0487
5	0.0438	0.0439	0.0439
10	0.0416	0.0416	0.0416

3.4.2 Laminate composite materials

In this section, analysis of laminate composites through FAST will be validated by comparisons with results found from other researchers. First, the assessment of the critical buckling load factor of a 2-ply simply supported laminate square plate will be performed. Here, the SS1 condition is employed, as shown in Figure 10. In this case, the plate has $a/h = 50$, where a is the plate length and h is the plate thickness. Orthotropic materials are considered for the analysis of laminates, and their properties are taken from Nguyen et al.[116], as shown in Table 7.

For validation purposes, a 8×8 cubic NURBS mesh is employed for structural analysis. The non-dimensional buckling load:

$$\lambda_{norm} = \frac{N_{cr}\, a^2}{E_2\, h^3} \tag{3.56}$$

Table 7 – Material properties for the validation of laminate analysis.

Material	E_1 (GPa)	E_2 (GPa)	G_{12} (GPa)	G_{13} (GPa)	G_{13} (GPa)	ν_{12}
I	120	3.00	1.80	1.80	1.50	0.25
II	120	4.80	2.40	2.40	0.96	0.25

is compared with the results found by Nguyen et al.[116] and Maia[3] in Table 8. In both references, analysis was also performed using the FSDT. Once again, the results are in excellent agreement with the ones from the reference.

Table 8 – Normalized buckling load of a simply-supported laminate square plate.

Layup	Material	Nguyen et al.[116]	Maia[3]	Present work
[0°/90°]	I	12.906	12.896	12.895
[±45°]	II	15.374	15.373	15.372

Now, the buckling load of a 10-ply clamped squared will be assessed. Here, results will be compared to those found by Maia[3]. The author also compares with results found by the commercial software ABAQUS[117]. These are shown in Table 9. Again, results found are in excellent agreement with those from the reference.

Table 9 – Normalized buckling load of a clamped laminate square plate.

Layup	Material	ABAQUS[117]	Maia[3]	Present work
$[(0°/90°)_2/0°]_s$	I	125.49	125.58	125.56
$[\pm 45°_2/45°]_s$	I	109.88	110.19	110.12

4 OPTIMIZATION OF COMPOSITE STRUCTURES

For a long time, engineers were only concerned with conceiving safe and functional designs. However, with advances in computational processing and analysis, the optimization of those structures became a requirement for efficient designs. The optimization process aims at finding the best outcome for a given operation while satisfying certain constraints[25]. In structural optimization, these constraints are usually very straightforward, since they are often defined by limit state conditions taken from design codes[118]. For instance, a structural element should not exceed its critical load or a given maximum displacement.

In general, a continuous nonlinear optimization problem can be defined as[25, 119]:

$$\begin{cases} \text{find} & \mathbf{x} = \{x_1, x_2, \ldots, x_m\} \\ \text{that minimizes} & f(\mathbf{x}) \\ \text{subject to} & g_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, n_c \\ \text{with} & \mathbf{x}_L \leq \mathbf{x} \leq \mathbf{x}_U \end{cases} \tag{4.1}$$

where $f(\mathbf{x})$ is the objective function, $g_i(\mathbf{x})$ is the i-th constraint function, n_c is the number of constraints, and $\mathbf{x}$ is the vector of design variables, where $\mathbf{x}_L$ and $\mathbf{x}_U$ define a lower and an upper bound to each variable, respectively. On a maximization problem, one may simply perform the minimization of $-f(\mathbf{x})$. For a discrete problem, $x_i \in [x_{i,1}, x_{i,2}, \ldots, x_{i,s}]$, where s is the number of possible values for the i-th design variable.

One should not think that optimization is a simple task. For example, on a problem with 5 design variables where each may assume 20 different values, an exhaustive search would require over three million evaluations to cover the entire design space. If the time required for the objective and constraint functions evaluation is only 1 ms, the algorithm would take an hour to test all possible designs. However, if the analysis time is 10 s, it would take a whole year. This combinatorial problem is considered to be NP-hard[120], and different algorithms have been proposed to deal with these problems more efficiently. An optimization algorithm should be reliable, efficient, and easy to use[121]. In that sense, Arora[119] describes different types of algorithms, which are discussed in the following.

In gradient-based methods, the objective and constraint functions are assumed to be continuous in the design space and at least twice continuously differentiable. The evaluation of these derivatives should be available to the algorithm, as the optimization step is defined almost entirely by the gradient information. The process continues until a well-established stopping

criterion is met[119]. In some cases, the derivative can be approximated by a numerical method, e.g. using a finite-difference approach[122]. It is important to note that, in these methods, the most time-consuming part of the process is precisely the derivative evaluation (or estimation)[27]. Also, such methods are usually dependant on the initial solution and, even though they guarantee convergence, they may find a local optimum instead of the global one[123, 119].

Nature-inspired search methods are zero-order optimization algorithms, which require only that the objective and constraint functions can be evaluated in the entire design space[119]. This also means that, with minor adaptations, they can handle discrete or even mixed problems[119]. Since no gradient information is used, there is no guarantee of convergence[124]. However, as their name suggests, they try to imitate nature aspects and, in many cases, mimic the gradient information with operators such as the particle's velocity. Furthermore, these algorithms work with a set (or population) of designs and, thus, there is a better exploration of the design space. This way, it is very easy to employ parallel computing in the optimization process[27]. Most of these methods are stochastic, as they employ randomized operators in their search for the optimum[27, 119], further improving their exploration capabilities. These aspects are very important for the optimization of nonlinear multi-modal complex functions. When compared to gradient-based algorithms, nature inspired search methods are much less prone to be trapped in a local optima.

Due to these very own reasons, bio-inspired algorithms are the most popular algorithms for the optimization of composite structures[25, 5, 14]. Lagaros, Papadrakakis and Kokossalakis[27] presented a comparison between variations of Genetic Algorithms (GAs), Evolution Strategies (ES), and the mathematical programming technique known as Sequential Quadratic Programming (SQP). The authors state that meta-heuristics are often more efficient, but that the combination between algorithms seems promising. Kou, Parks and Tan[26] also showed that the PSO outperforms classical mathematical programming optimizers in FG problems due to their global search capabilities.

Genetic Algorithms (GAs)[125, 126], Particle Swarm Optimization (PSO)[127], and Differential Evolution (DE)[128] are examples of meta-heuristics commonly employed in the optimization of composite structures[26, 15, 30, 16, 31]. Multi-objective optimization is often performed using variations of these algorithms[129, 130, 17, 131, 132]. The choice for a given approach often depends on aspects such as variable type, modality, and dimensionality[119, 127, 121, 133]. Kitayama, Arakawa and Yamazaki[134] presented a comparison between a variety of meta-heuristics for

continuous structural optimization, showing that the best approaches were the PSO and the DE, where the latter is sometimes able to find the optimum with a lower number of function calls, therefore being more efficient. Similar results can be found in Ribeiro et al.[34].

In the optimization of composite structures, the most common objective functions are related to the structure critical buckling load, fundamental frequency, deflection, stresses, weight, and cost[5, 14]. One should note that the evaluation of most of these responses requires a computational simulation. On laminate problems, variables are usually discrete, such as the number of plies and their orientation, thickness, or even material. On the other hand, on FG problems, variables are often continuous, mostly being related to the material gradation through the structure. Besides, multiobjective optimization of composite structures is receiving increasing interest in the last few years, and new methods are being constantly developed and employed[135, 136, 85, 17, 96, 137]. A review of the optimization of laminate and FG structures will be presented in the following.

Regarding laminated composites, Genetic Algorithms (GAs) are often preferred due to their discrete behavior. Karakaya and Soykasap[138] employed GA for buckling optimization of a balanced-symmetric composite panel. The authors performed the analysis considering an analytical approach, based on the Navier solution. A GA was also employed by Irisarri et al.[139] to find the Pareto front for the optimal stacking sequence of a laminated plate considering two objectives: mass minimization and buckling load factor maximization. Here, the authors performed the analysis considering the CLT. Le-Manh and Lee[140] used a GA to maximize the strength of laminated plates. The authors derive an IGA formulation for the nonlinear analysis of plates considering the FSDT.

Different researchers also successfully applied other algorithms for optimization of laminated composites. A comparison between GA and Simulated Annealing (SA) was presented by Sciuva, Gherlone and Lomario[141] for optimization of laminated and sandwich plates, considering higher-order plate theories. The authors show that both algorithms perform very similarly, but the SA is less time-consuming. Chang et al.[142] argue that a variant of the PSO algorithm was the most efficient method for the maximization of the buckling load factor of laminated plates, which was evaluated via analytical solutions.

Roque and Martins[143] employed DE for maximization of the natural frequency of laminated plates using a finite element formulation based on the FSDT. Vo-Duy et al.[144] used a similar procedure for frequency maximization of carbon nanotube reinforced laminates

using DE. Bargh and Sadr[145] also tried to maximize the natural frequency of laminated plates, this time by the use of a PSO algorithm. Numerical analyses were performed using a Finite Strip Method (FSM) formulation based on the CLT. Jing et al.[146] employed a mechanics-based optimization algorithm for the maximization of the fundamental frequency of doubly curved laminated composites. Analyses were performed using the Rayleigh-Ritz method, and the authors found the optimal stacking sequence for such structures.

A methodology for weight minimization considering ply drops can be found in Fan, Wang and Chen[147] by the use of a modified GA. The authors present modifications to existing operators to facilitate the optimization process and show that optimal designs found outperform those from conventional laminates. Deka et al.[148] considered the weighted average between two objectives, weight and cost, to perform the multi-objective optimization of laminated composites using a GA. The authors made use of the Tsai-Hill failure criterion, and the design variables were the fiber orientations.

Regarding failure criteria, some authors tested different approaches to understand how these may affect the optimization process. Lopez, Luersen and Cursi[149] performed the weight and cost minimization considering different failure criteria: maximum stress, Tsai-Wu, and Puck failure criterion. The authors show that optimal designs depend on the failure criteria, but no criterion was always the most or least conservative. In a similar fashion, Satheesh, Naik and Ganguli[150] used GA for weight minimization considering three different failure criteria, this time the maximum stress, the Tsai-Wu, and the mechanism-based. The authors propose the consideration of a failure envelope in such a way that the most conservative criterion is always considered.

Rocha, Parente and Melo[15] presented a modified GA where a special encoding for laminates allowed for the incorporation of three specific operators. The authors then perform the optimization of laminated plates and panels, finding the optimal stacking sequence, material, and thickness of each ply. Barroso, Parente and Melo[16] used a similar encoding, but optimization was performed using a hybrid PSO-GA, where a mutation operator is incorporated to the PSO to improve exploration.

Even with the shortcomings that laminated structures present when subject to thermal stresses, Kamarian, Shakeri and Yas[151] used the Firefly Algorithm (FA) to find the optimum stacking sequence for maximization of the thermal buckling load factor considering the FSDT. Using the same theory, Singha, Ramachandra and Bandyopadhyay[152] performed the maximiza-

tion of the buckling temperature of laminated plates using GA considering various aspect ratios and boundary conditions. The authors found the optimal stacking sequence, also considering each ply thickness as a design variable. The total plate thickness was considered as a design constraint. Spallino and Thierauf[153] performed a similar optimization using GA and a CLT-based finite element formulation. The authors study different constraint-handling methods, showing that all of them performed similarly. In the same vein, Vijayachandran et al.[154] showed that GA is able to find optimal fiber curve paths for the maximization of the thermal buckling load of a laminated plate with a cutout.

Finally, Kalita, Haldar and Chakraborty[155] provided an extensive and comprehensive review on the optimization of composite laminates using conventional and Surrogate-Based Optimization (SBO) approaches. In short, the authors state that FEM is the most popular method for analysis of composite plates, where the FSDT is the most popular theory due to its simplicity and accuracy. Also, the authors comment that, since the last decade, metamodeling is becoming extremely popular, and most works focus on uncertainty quantification or optimization.

Now, a short review over the optimization of FG structures will be presented. These materials are often preferred when thermal effects are considered. Regarding unidirectional Functionally Graded Plates (FGPs), where the gradation is usually given in the thickness direction, Ding and Wu[156] performed the minimization of peak thermal stresses to find the optimal material distribution using a GA, considering a coupled thermo-mechanical analysis. A similar procedure can also be found in Chiba and Sugano[157]. Na and Kim[158] performed a multi-objective optimization considering a weighted average between the peak stress and the critical buckling temperature, finding the optimal volume fraction gradation defined by the power-law equation. Effective material properties were evaluated using the Voigt model.

A comparison between PSO and GA was performed by Ashjari and Khoshravan[91] for the mass minimization of FG beams, considering deflection and stress constraints. Analyses were performed considering the HSDT. The authors used a Hermite polynomial to represent the volume fraction distribution and the effective properties are evaluated by the rule of mixtures. PSO is shown to be superior both in terms of convergence speed and accuracy.

For FGPs where gradation changes in two directions, Hussein and Mulani[159] performed the minimization of cost while respecting buckling constraints, considering uniaxial and shear loadings. Here, the finite element formulation considered the CPT and, again, effective properties are evaluated by the rule of mixtures. Later, the same authors[160] performed a similar

study, now considering constraints on maximum deflection and stresses.

Trying to mimic the bone structure, Huang and Rapoff[161] presented the optimization of FGPs with circular cutouts. The authors minimized the failure index, and the FG structure is shown to outperform homogeneous plates in terms of strength to weight ratios. Silva and Loja[162] used DE to perform the optimization of sandwich FG panels. Design variables were related to the material gradation and the thickness of the plate.

Also regarding sandwich structures, Lieu et al.[106] performed shape and size optimization of FGPs using a Firefly Algorithm. The authors derived an IGA formulation based on the HSDT, and effective mechanical properties were assessed by the Mori-Tanaka model. The authors show that the use of B-Spline functions to control the material gradation may lead to designs that are easier to be applied to practical manufacturing. In many cases, one-parameter rules seem to be fairly restrictive. In the design of wings and tail fins of aircrafts, Torabi and Afshari[163] used a PSO to maximize critical aerodynamic pressure. Design variables are related to the plate geometry, power-law index, and angle of attack. Analyses were performed considering the FSDT.

Material distribution and size optimization of FG plates and shallow-shells are performed by Moita et al.[164]. Material gradation is given by the power-law function, and equivalent properties are assessed using the Voigt model. Regarding multi-objective optimization, Correia et al.[130] maximized the buckling load and the natural frequency while also minimizing the cost and weight of FGPs described by the power-law function. Effective properties are evaluated via the Mori-Tanaka scheme. The authors apply the Tsai-Hill failure criteria to ceramic, metal, and graded phases, considering the strengths of each material separately. Later, the same authors[85] used a similar procedure but considered thermo-mechanical loading. Correia et al.[96] employed a SA algorithm in the optimization of FGPs under thermo-mechanical loadings. In these papers, all numerical analyses are performed considering high-order theories.

Minimization of stresses, displacements, and mass was also performed by Moleiro et al.[137] in the optimization of FGPs under thermo-mechanical loading. In the optimization problem, material failure is assessed considering the Tsai-Hill failure criteria. The authors show that the methodology is able to achieve optimal designs considering different homogenization methods. The same authors[165] later performed the multi-objective optimization of symmetric and non-symmetric sandwich FGPs and studied the optimal material distribution for different loading cases. Wang et al.[17] used a modified PSO algorithm for the multi-objective size

optimization of FGPs. The authors employ an IGA formulation based on a recently developed so-called "simple FSDT". Material gradation is described by different functions, and effective properties are evaluated via the Voigt model.

Regarding shell designs, Vel and Pelletier[166] employed a GA-based method for the multi-objective optimization of thick shells under thermal loadings. The authors were concerned with minimizing stresses and weight. Material gradation is defined by Hermite polynomials, and effective properties are evaluated by the Self-Consistent model. A similar approach was adopted for multi-objective optimization of hollow cylinders by Nabian and Ahmadian[28], where authors performed the minimization of mass and maximization of the fundamental frequency.

PSO is employed by Shabana et al.[167] to minimize stresses in multi-layered FG cylinders subject to pressure. The authors evaluate the structural response by an analytical approach and show that peak stresses can be reduced by nearly 18%. Comparing different algorithms, Tornabene and Ceruti[168] found the optimal material gradation in FG panels and shells. The authors claim that PSO and GA were the best-performing methods for the minimization of deflection and maximization of the fundamental frequency.

Recently, due to its major gain in efficiency, the Surrogate-Based Optimization (SBO) of composite structures has been gaining interest. Ootao et al.[169] used an Artifical Neural Network (ANN) in the minimization of thermal stress to find the optimal material gradation for FGPs. The authors show that, after the evaluation of the sampling points and model building, the ANN is able to provide a very fast estimate for the true function. To determine the optimal reinforcing pattern of aluminum composite plates, Akbari et al.[170] used experimental results to train an ANN and then performed a GA-based multi-objective optimization over the approximate response surface. Using this methodology, the authors were able to maximize the hardness and ultimate tensile strength of composite plates.

Similarly, Garmsiri and Jalal[171] employed an ANN to assist in the multi-objective optimization of cylindrical shells, maximizing strength and natural frequency. In the optimization of composite cylinders, Blom, Stickler and Gürdal[172] approximated the buckling load factor by a surrogate model to reduce the number of high-fidelity function evaluations. The authors show that the model is able to facilitate convergence, while also increasing the efficiency of the process. Also employing a surrogate model, Njim, Bakhy and Al-Waily[173] performed the optimization of porous FGPs. Numerical analyses are performed considering the CPT.

Pitton, Ricci and Bisagni[174] perform the buckling optimization of Variable Stiffness

Composites (VSCs) using a PSO. These composites use multiple layers, similar to laminates, but, due to its variable fiber paths, the optimization problem presents continuous design variables. The optimization process is assisted by an Artifical Neural Network (ANN) built with 135 input-output pairs. While the ANN is able to provide an approximate solution at a very minor cost, the authors state that, considering the dataset creation and training costs, the time required to compute the optimal configuration is about 14 days.

Do, Lee and Lee[1] performed the maximization of the buckling load factor and fundamental frequency of unidirectional FGPs considering a HSDT. After evaluating 10,000 sampling points, a Deep Neural Network (DNN) was trained for each approximated objective function. The authors show that the optimization of the approximate response surface by a symbiotic organisms search algorithm leads to minor errors when compared to the true global optima. The same authors[18] later employed a similar approach considering multi-directional FGPs. Here, the authors show that the time required to create the dataset is the main bottleneck, as it corresponds to almost 98% of the total optimization cost. A DNN was also employed by Truong, Lee and Nguyen-Thoi[42], this time in the multi-objective optimization of multi-directional FG beams. The authors show that the DNN presents an Mean Squared Error (MSE) lower than 0.002% in both training and test sets.

Yin et al.[175] used an ensemble of different meta-models (polynomial regression, RBF, Kriging, SVR, and MARS) to assist in the optimization of FG foam-filled tubes. The authors optimize the response surface using a modified PSO and show that the ensemble is able to perform better than each of the models separately. Another approach to improve upon Surrogate-Based Optimization (SBO) is the use of an adaptive sampling scheme, where the sample is updated in each iteration. This is an important idea for this work and will be discussed in future sections, along with a proper review.

A more extensive review of the optimization of laminated and FG composites can be found on Nikbakt, Kamarian and Shakeri[5] and Nikbakht, Kamarian and Shakeri[14], respectively. This work will also employ three different bio-inspired optimization algorithms: the GA, the PSO, and the DE. These meta-heuristics will be further described in the following sections.

4.1 Optimization model

A review over recent papers on the optimization of laminated and functionally graded composites shows that many different optimization problems can be formulated. In this work, a

major focus will be given to the optimization of plates and shallow shells considering the 3D solid continuum theory and a First-order Shear Deformation Theory (FSDT), both described in Chapter 3.

Objective functions considered will be related to eigenvalue analyses, such as maximization of the mechanical or thermal buckling load, or optimization of natural frequencies.

For FGMs, to allow for complex and efficient structures, gradation will be defined by unidirectional or tridirectional B-Spline functions. Thus, design variables will be related to the control points of these B-Spline functions. Constraints will be related to a maximum ceramic percentage $\overline{V}_c$, structure cost C_{tot}, or, in some cases, structural responses such as natural frequencies. The ceramic percentage may be evaluated by:

$$\overline{V}_c = \frac{1}{V}\int_V V_c \, dv \tag{4.2}$$

where V_c is the ceramic percentage in a given point of the structure. The structure cost, on the other hand, may be evaluated via:

$$C_{tot} = C_c \int_V V_c dv + C_m \int_V V_m dv \tag{4.3}$$

where C_c and C_m are the costs by unit volume of the ceramic and metal, respectively.

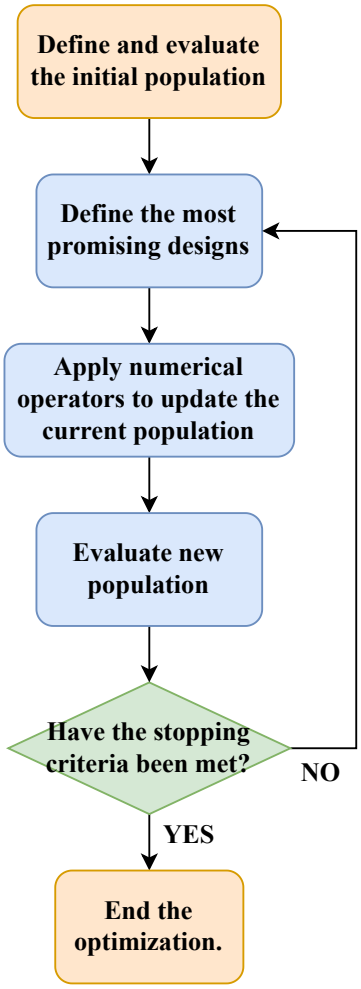
For laminates, design variables will be related to the orientation and, in some cases, the thickness of each ply. Constraints, on the other hand, will be related to a maximum number of contiguous plies or the thickness for the structure. For the latter, a repair algorithm can be employed, to guarantee that all designs present the same total thickness. Before evaluating an individual, the thickness of the innermost layer can be set in order to guarantee that $\sum h_i = h_{max}$. Then, if this value is outside the design space, the thickness of the other plies are adjusted, prioritizing the outermost layers.

Figure 11 presents a general flowchart for optimization via conventional nature-inspired algorithms. In the following sections, three different algorithms will be describe: Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Differential Evolution (DE).

4.2 Genetic Algorithms

Genetic Algorithms (GAs) were among the first bio-inspired algorithms employed. The initial version was first proposed in 1975 by Holland[176], being inspired by Darwin's theory of natural selection[119]. Its implementations try to mimic genetic operators.

Figure 11 – General procedure for nature-based optimization algorithms.



Source: the author

The first GA was based on a binary coding, further simulating the genetic encoding[177]. The individual was defined by its genotypical representation, while fitness assessment was performed based on their phenotypical behavior. The binary coding also further emphasized the focus on discrete optimization. Indeed, GAs are still very strong optimizers for discrete optimization. Nowadays, however, integer coding is often employed due to the easier implementation, or even real-number coding, which allows for continuous optimization. In this work, the version described in Rocha, Parente and Melo[15] is employed, which is presented in the following.

A GA is initialized by generating an initial population of N_p random individuals, each represented by a vector of design variables $\mathbf{x}_j$. Each individual is then evaluated, assessing its objective function f_j. In each generation, this population will be continuously improved. To guide the optimization process, a fitness function is often considered in order to select the best

designs. For minimization problems, the fitness can be evaluated by[15]:

$$Fit_j = \max\left(|f_{min}|, |f_{max}|\right) - f_j \tag{4.4}$$

where f_{min} and f_{max} are the minimum and maximum objective function values in the population.

Then, the mating pool should be selected from the old population. The number of selected designs depends on the crossover rate C_r, and the selection should be biased towards designs with a high fitness value. The probability of selection can be evaluated by[80, 119]:

$$p = \frac{Fit_j}{\sum_j^{N_p} Fit_j} \tag{4.5}$$

The classic roulette method can be employed to define which designs will be drawn to the mating pool[15, 119].

Once the mating pool has been defined, the crossover starts, and, at each iteration, two parents, $\mathbf{x}_1^{(p)}$ and $\mathbf{x}_2^{(p)}$, are selected from the mating pool. Then, the new trial designs can be defined using linear combination:

$$\begin{aligned} \mathbf{x}_i &= r\,\mathbf{x}_1^{(p)} + (1-r)\,\mathbf{x}_2^{(p)} \\ \mathbf{x}_{i+1} &= (1-r)\,\mathbf{x}_1^{(p)} + r\,\mathbf{x}_2^{(p)} \end{aligned} \tag{4.6}$$

where r is a random number between 0 and 1.

Finally, to prevent the algorithm from getting trapped in a local minimum, the mutation operator is employed. Each variable has a small probability p_{mut} of mutating to a random value in the design space.

These operators define a base GA, which is able to deal with binary, discrete, and continuous optimization. The algorithm will be carried out until a stopping criterion is met. In this work, two criteria will be considered: one related to the maximum number of generations (N_{gen}), and the other related to the maximum number of stall generations (Gen_{stall}), that is, the number of generations with no relevant improvement over the objective function.

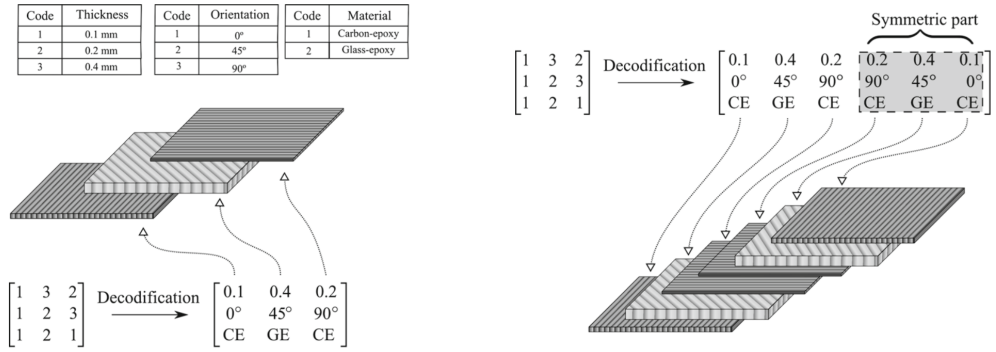
The version proposed by Rocha, Parente and Melo[15] focused on the optimization of laminate structures, presenting a special encoding and specific operators for these problems. These will be explained in the following.

4.2.1 Laminate problems

Laminate problems are usually considered using discrete variables, such as the number of plies and each ply thickness, orientation, and material. Here, each individual is represented

by an integer matrix, where each row represents a variable and each column represents a ply. The decoding process is performed as depicted in Figure 12.

Figure 12 – Laminate decoding process.



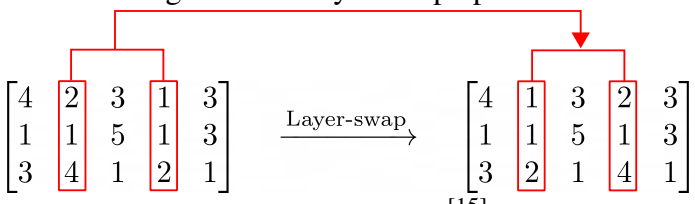
(a) General laminates

(b) Symmetric laminates

Source: Barroso, Parente and Melo[16]

This special encoding allows for easier implementation of three specific operators: the layer swap, the layer addition, and the layer deletion. In the layer-swap, there is a small chance p_{swap} that the position of two layers (columns in the matrix) are swapped, as shown in Figure 13. This operator is especially interesting for problems with deflection, buckling, or natural frequency constraints[16].

Figure 13 – Layer-swap operator.



Source: Rocha, Parente and Melo[15]

The layer deletion can be employed when the number of plies is a design variable, where there is a probability p_{del} to set the thickness of a layer to zero. Likewise, the layer addition operator has a probability p_{add} that a given layer will be reactivated, by setting its thickness to a value other than zero[15, 16].

4.3 Particle Swarm Optimization

The Particle Swarm Optimization (PSO) is a very popular meta-heuristic, which aims at optimizing nonlinear continuous functions[127, 91]. Kennedy and Eberhart[127] proposed

the technique based on the behavior of a swarm of particles in its search for food. The method is computationally inexpensive (in terms of speed and memory requirements), as only primitive mathematical operators are employed[127]. Since the proposition of the method, different researchers have expanded upon the initial PSO concept[178]. This work will employ the version described in Barroso, Parente and Melo[16].

The PSO algorithm starts by generating an initial population of N_p random particles. Each particle j is assigned a position $\mathbf{x}_j^0$ and a velocity $\mathbf{v}_j^0$. At each iteration (i), the particles move in the design space according to their velocity:

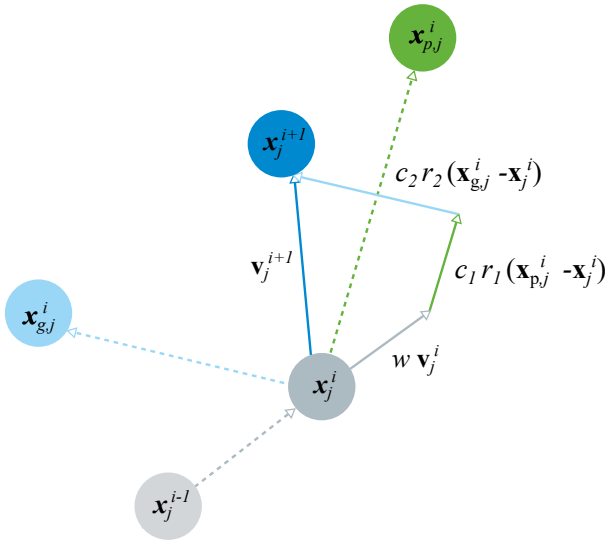
$$\mathbf{x}_j^{i+1} = \mathbf{x}_j^i + \mathbf{v}_j^{i+1} \tag{4.7}$$

where:

$$\mathbf{v}_j^{i+1} = w\,\mathbf{v}_j^i + c_1\,r_1\,(\mathbf{x}_{p,j}^i - \mathbf{x}_j^i) + c_2\,r_2\,(\mathbf{x}_{g,j}^i - \mathbf{x}_j^i) \tag{4.8}$$

Here, w is the inertia, c_1 is the cognitive factor, c_2 is the social factor, $\mathbf{x}_{p,j}^i$ is the best position the particle j obtained during the optimization, and $\mathbf{x}_{g,j}^i$ is the best position the particles on the neighborhood of particle j obtained during the optimization. The parameters r_1 and r_2 are uniformly distributed random numbers between 0 and 1. Thus, both the cognitive and social experiences affect the particle velocity[91], as depicted in Figure 14. This iterative process goes on until a stopping criterion is met.

Figure 14 – Particle movement at each iteration.



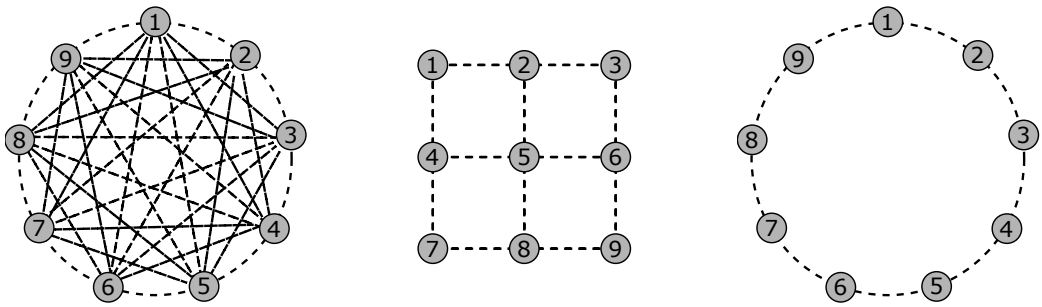
Source: Maia[3]

The definition of $\mathbf{x}_{g,j}^i$ depends on the neighborhood of particle j, which is defined by the swarm topology. In early versions, Kennedy and Eberhart[127] would consider that all

particles are connected and, thus, $\mathbf{x}^i_{g,j}$ would be the best position found among all particles. Even though this formulation leads to a very fast convergence, the algorithm may get stuck in a local optimum[179].

Later, the Square and the Ring topologies were also developed. In the former, the swarm is arranged as a matrix and the particle has a neighbor in each direction (above, below, right, and left). In the Ring topology, each particle may only be influenced by the two closest ones. The use of these alternatives delay the sharing of social experience between the particles, thus improving the exploration of the design space but slowing down the algorithm convergence[178, 16]. Figure 15 depicts each of these alternatives. In the so-called standard PSO, Bratton and Kennedy[178] employ the Ring topology.

Figure 15 – Swarm topologies.

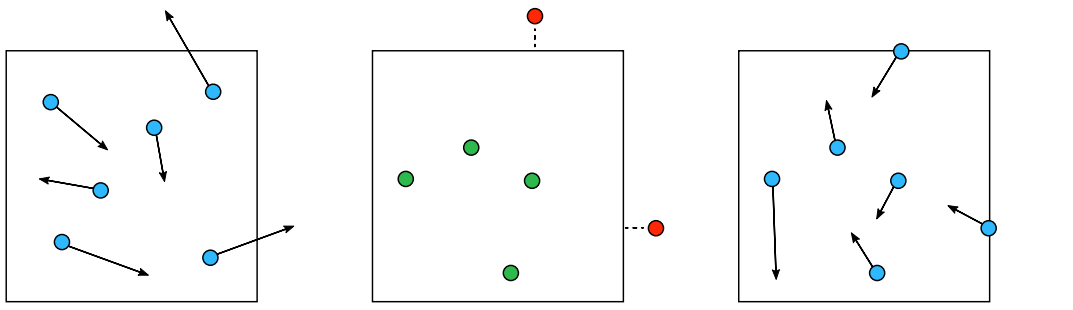


(a) Global (b) Square (c) Ring

Source: Ribeiro et al.[2]

During the optimization process, a particle may violate an upper or lower variable bound. To prevent particles from leaving the design space, a simple procedure is performed as depicted in Figure 16, where the particle is set to the violated bound and the violated velocity component is modified to the opposite direction[16, 2].

Figure 16 – Boundary violation handling.



(a) Define the velocity (b) Handle the violations (c) New positions and velocities

Source: Ribeiro et al.[2]

Even with the appropriate topology, the PSO may still suffer from premature convergence. Also, depending on the size of the population, it may be very hard for the algorithm to explore the entire design space. In this work, a mutation operator will be employed to improve this aspect[16]. The operator is applied to the particle position, where each variable has a small probability p_{mut} of mutating to a random value, between the lower and upper variable bounds.

Barroso, Parente and Melo[16] also extended the PSO formulation to laminate designs, based on the special encoding proposed by Rocha, Parente and Melo[15], described in Section 4.2.1.

4.4 Differential Evolution

The Differential Evolution (DE) was initially developed by Storn and Price[128] as a robust and simple method with good convergence properties. The algorithm was inspired by the well-known GA, but the use of floating-point instead of bit-string encoding along with arithmetic operators turned it into a strong numerical optimizer[180, 128], more appropriate to deal with continuous optimization problems[31, 32]. Besides, the existence of only a few user-defined parameters makes it a simple and easy-to-use method[121].

Similar to other population-based algorithms, the DE starts by randomly generating a population of N_p individuals in the design domain. These individuals, represented by vectors of design variables $\mathbf{x}$, will continuously evolve, simulating a developing population[121, 34]. Once each individual is evaluated, the process of conceiving a new child population starts. To each individual, taken as the base vector, three operators are applied. First, the differential mutation works by adding a scaled difference vector to the base vector as:

$$\mathbf{v}_j^i = \mathbf{x}_{r0}^i + F\left(\mathbf{x}_{r1}^i - \mathbf{x}_{r2}^i\right) \tag{4.9}$$

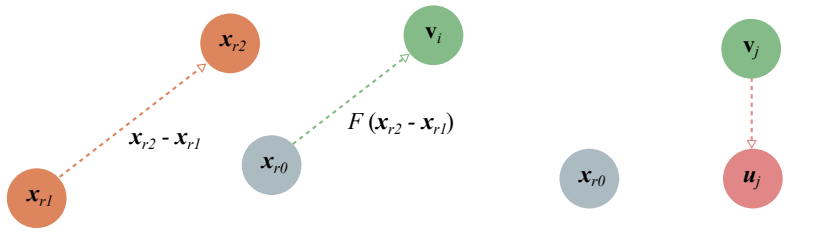
where F is the user-defined scale factor. In the initial DE formulation, the indexes $r0$, $r1$ and $r2$ are chosen at random, and they represent arbitrary individuals of the population, while $\mathbf{x}_{r0}^i$ is a base vector in the i-th iteration. A binomial crossover is then carried out by combining the mutated vector and the base vector:

$$u_{j,k}^i = \begin{cases} v_{j,k}^i & \text{if } r_{c,k}^i \leq C_r \\ x_{j,k}^i & \text{otherwise} \end{cases} \tag{4.10}$$

Here, the subscript k refers to the k-th variable in the design vector, C_r is the crossover probability, and $r_{c,k}^i$ is a random value taken from an uniform distribution. This value is compared to C_r to

decide whether the child vector inherits from the base vector or the mutated vector. Figure 17 illustrates how these two operators work.

Figure 17 – Differential Evolution operators.



(a) Differential mutation

(b) Crossover

Source: the author

The final operator, the selection, compares the child vector $\mathbf{u}_{j,i}$ with the base vector $\mathbf{x}_{j,r0}$. Thus, the best design between those two is carried over to the next population. The base vector gets replaced if the child vector presents an improvement upon it[121]:

$$\mathbf{x}_j^{i+1} = \begin{cases} \mathbf{u}_j^i & \text{if } f(\mathbf{x}_j^i) \leq f(\mathbf{u}_j^i) \\ \mathbf{x}_j^i & \text{otherwise.} \end{cases} \tag{4.11}$$

This initial version of the DE is known as DE/Rand/1/Bin[121], where Rand refers to the base vector employed in the differential mutation operator, 1 to the number of differentiation vectors, and Bin to the type of crossover used, the binomial crossover. In the following years, a multitude of versions was proposed[181]. In particular, many researchers show that carrying over information regarding the best individual in the population can be beneficial to the convergence speed[182, 183]. Two of the most promising approaches are the Best/1/Bin:

$$\mathbf{v}_j^i = \mathbf{x}_{best}^i + F\left(\mathbf{x}_{r1}^i - \mathbf{x}_{r2}^i\right) \tag{4.12}$$

and the Current-to-best/1/Bin:

$$\mathbf{v}_j^i = \mathbf{x}_j^i + F\left(\mathbf{x}_{best}^i - \mathbf{x}_j^i\right) + F\left(\mathbf{x}_{r1}^i - \mathbf{x}_{r2}^i\right) \tag{4.13}$$

where $\mathbf{x}_{best}^i$ is the best design found during the optimization process until the i-th iteration[121, 182, 184].

4.5 Constraint-handling

In nature-inspired search methods, explicit constraints (e.g. design variable bounds) are dealt with directly by preventing these from occurring (e.g. not allowing a bound constraint

to be violated). However, implicit constraint functions require special procedures to prevent an infeasible design from being chosen.

In these cases, constrained optimization is often solved by employing a penalty function that undermines infeasible designs, turning the problem into an unconstrained one. For instance, a simple static penalty can be employed[185, 186, 27]:

$$f_p(\mathbf{x}) = f(\mathbf{x}) + K \sum_{i=0}^{n_c} c_i(\mathbf{x}) \tag{4.14}$$

where K is the static penalty parameter and $c_i(\mathbf{x}) = g_i(\mathbf{x})$ if the i-th constraint is violated, and 0 otherwise.

In this work, however, the adaptive penalty proposed by Lemonge and Barbosa[186] will be employed, where penalization is performed as:

$$f_p(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{, if } \max g(\mathbf{x}) \leq 0 \\ \overline{f}(\mathbf{x}) + \sum_{i=0}^{n_c} K_i\, g_i(\mathbf{x}) & \text{, otherwise} \end{cases} \tag{4.15}$$

where:

$$\overline{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{, if } f(\mathbf{x}) > f_m \\ f_m & \text{, otherwise} \end{cases} \tag{4.16}$$

where f_m is the average objective function value in the current population. The penalty factor K_i is then evaluated by:

$$K_i = |f_m| \frac{v_{i,m}}{\sum_{j=1}^{n_c} (v_{j,m})^2} \tag{4.17}$$

where $v_{l,m}$ is the violation of the l-th constraint averaged over the current population.

5 SEQUENTIAL APPROXIMATE OPTIMIZATION

Simulation-based structural analysis, e.g. by the use of the Finite Element Method (FEM) or the Isogeometric Analysis (IGA), may incur high computational costs. For instance, Gu[24] reported that Ford takes about 36 to 160 hours to run one crash simulation. In such cases, bio-inspired search methods are not adequate since they usually require hundreds or even thousands of structural analyses to find the global optimum[35, 36, 37]. Therefore, it is not feasible to perform a conventional optimization with such a high analysis time[38]. Meanwhile, even though the last few decades presented major advances in computing processing power, the required fidelity and model complexity on engineering analyses also increased. Thus, the computational cost and time constraints have not been eliminated[39]. While parallel computing is one very effective alternative for this issue, this procedure requires powerful machines which are not always available to the user.

The Sequential Approximate Optimization (SAO) was first proposed by Schmit and Farshi[187] as a method to deal with computationally expensive optimization problems. In this approach, a costly function is approximated by a response surface, which is then used to guide the optimization process. The approximation is iteratively improved by focusing on promising regions in the design space. In its initial formulation, SAO was performed using the trust region approach, where a move limit strategy was employed to define regions in which sub-problems should be solved[187, 188]. The design space would be sequentially funneled down until a stopping criterion is met. Different frameworks which aim at performing this procedure can be found in the literature[189, 188]. Venkataraman and Haftka[190] state that, here, optimization-related computational costs depend on three indices of complexity: modeling, analysis, and optimization complexity.

The approximate response surface employed is often denominated as a metamodel or a surrogate model. Using the true function values evaluated at a set of sampling points, it is possible to fit a given surrogate model to serve as a cheaper, but approximated, way of evaluating the true function. The accuracy of this approximation is dependent on the choice of sampling points, the model used, and the true function complexity.

Wang and Shan[191] comment that there are many advantages of applying metamodeling in optimization. First, efficiency is greatly improved, when compared to conventional optimization approaches. Second, parallel computation is supported for evaluation of the sampling points, when these may be obtained independently. Third, modeling can usually help

to understand the importance of each design variable. Finally, with proper adaptations, both continuous and discrete optimization can be handled.

According to Forrester, Sobester and Keane[40], surrogate models are "educated guesses as to what an engineering function might look like, based on a few points in space where we can afford to measure the function values". The authors state that there are three main stages in the construction of an approximation by a surrogate model[40]. First, the sampling points are selected and evaluated.

In this work, these points are evaluated via computational methods for structural analysis (e.g. IGA or FEM), with a sufficiently refined mesh. For single-fidelity models, these methods are assumed to provide the exact response for the true function. The user should seek an appropriate method to define these points so that fewer costly evaluations are required[192].

In the second stage, the user should set the model hyper-parameters and fit the surrogate model. The use of reliable methods, capable of fitting linear and nonlinear spaces, is very important to guarantee a robust prediction[192]. Finally, in the third stage, the model should be tested to ensure that it provides sufficient accuracy. This can be done by the use of a validation sample[40].

These steps define the conventional surrogate modeling approach, denominated here as static single-fidelity surrogate models[52]. Even though this approach has been widely used in engineering problems[46, 50, 1, 18, 48, 193, 33, 194], the optimization of a constant surrogate model heavily relies on the global approximation accuracy of this model[195], thus requiring a large number of sampling points to guarantee an accurate solution.

Along with single-fidelity models, Song et al.[52] points out three other modeling approaches. First, in hybrid surrogate models[196, 197, 66, 198], an ensemble of models is employed in a way to offer a more robust approximation, while also removing the need for selecting the model type *a priori*.

On the other hand, multi-fidelity surrogates[70, 52, 45] may consider data from both a High-Fidelity (HF) and a Low-Fidelity (LF) source, allowing for the use of a much larger sampling space. These sources may be, for instance, related more refined and a coarser analysis meshes, which may provide responses with different levels of accuracy and computational efficiency. If sources are well-correlated, multi-fidelity surrogates are able to provide very accurate approximations of the true function at a lower cost than single-fidelity models.

Finally, the third approach is the adaptive sampling-based, where the model is

iteratively used to locate promising regions in the design space. Then, new data can be selected from these regions, improving model accuracy in its surroundings. This is the most interesting approach for optimization since the user is only concerned with accuracy in the neighborhood of the optimum which, ideally, should be identified as a promising region. Forrester, Sobester and Keane[40] even state that, in this case, the user is not required to test the model, since global accuracy is not as important.

The adaptive sampling effectively minimizes the number of sampling points required for the model to provide accurate predictions in promising regions[199, 200, 201, 202]. It is worth noting that different approaches can be employed together, e.g. using adaptive sampling on multi-fidelity[78, 79, 77, 73] or hybrid surrogates[203, 75, 204, 205]. In this work, we will try out the former, by the use of multi-fidelity surrogates to improve model accuracy and, thus, better assist in the selection of new sampling points, also known as *infill points*. These will be further described in Chapter 6, while, here, we focus on single-fidelity models.

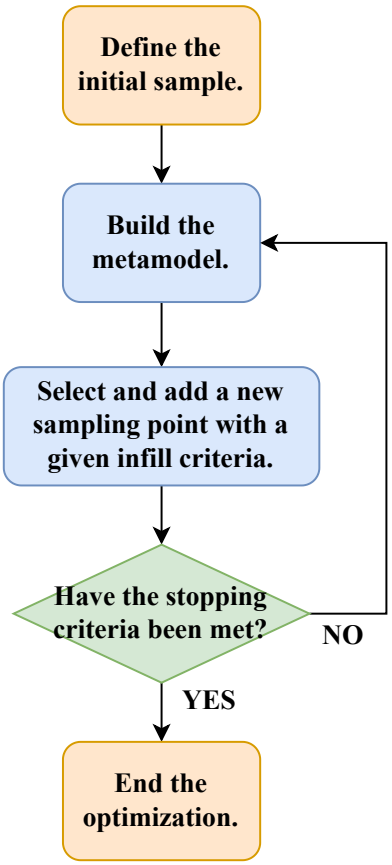
The use of adaptive sampling can be seen as a type of Sequential Approximate Optimization (SAO), as the model is sequentially improved throughout the optimization process[53, 55, 40]. However, unlike the trust region approach, adaptive sampling-based SAO takes advantage of the global response surface in the whole process. This way, no sampling points are ever lost during the optimization. Furthermore, the method is able to better explore the entire design space, if necessary[35].

Figure 18 presents the general flowchart for SAO based on adaptive sampling. The set of methods employed in each stage defines a SAO approach. For instance, the Efficient Global Optimization (EGO) algorithm proposed by Jones, Schonlau and Welch[53] uses a Latin Hypercube Sampling (LHS) to select the initial sampling points, then fit a Kriging model to the data, and select a new point by maximizing the Expected Improvement (EI). The stopping criterion is the value for the EI itself, and the process is terminated if it is lower than 1% of the best current value. Riche and Picheny[206] state that EGO and its variants behave similarly (or even better) to state-of-the-art algorithms for low-dimensionality multi-modal functions.

It is worth pointing out that the EI is an error-based infill criterion[40], as it considers the uncertainty of the process in the selection of a new candidate point. Such techniques are very promising and can be employed because the Kriging model is a Gaussian Process (GP), allowing for an analytical estimation of the error (or uncertainty).

In the last decade, SAO has become very popular for expensive engineering optimiza-

Figure 18 – General SAO flowchart.



Source: the author

tion. A variety of researchers have employed the method in structural optimization, achieving good results in many areas[56, 2, 207, 51, 208, 209, 78]. However, SAO still has to overcome certain issues so that it can be reliably employed in most optimization problems[210, 58, 206, 211]. For instance, the handling of expensive constraints or discrete design spaces is still an open issue.

In this work, different SAO approaches will be tried out. The following sections will further discuss each of its main stages, describing techniques to perform the definition of the initial sampling, the surrogate model building, and the further selection of infill points. A review of important papers on the matter is also presented.

5.1 Initial sampling

The first stage of building a surrogate model is data preparation. Here, a small number of observations is gathered, which will be later employed to fit a given model. These observations should assist in identifying important patterns in the data, so that promising regions may be located. Care must be taken in this phase since, if important information is discarded, the

model overall accuracy can be compromised[41]. In that matter, Design of Experiments (DoE) techniques are usually employed to pick an appropriate set of sampling points.

In real-world experiments, data tends to be noisy due to the inherent randomness of the process[192, 212]. Thus, it is often a good choice to sample multiple data points on small sub-regions of the design space so that the user may identify noisy patterns[192]. However, computer experiments tend to be deterministic and are not affected by some common constraints in real-world experiments[213, 214]. Thus, sample sets that ensure the coverage of the design space are often the most appropriate choice if the user has no prior information about the behavior of a given function[212, 40].

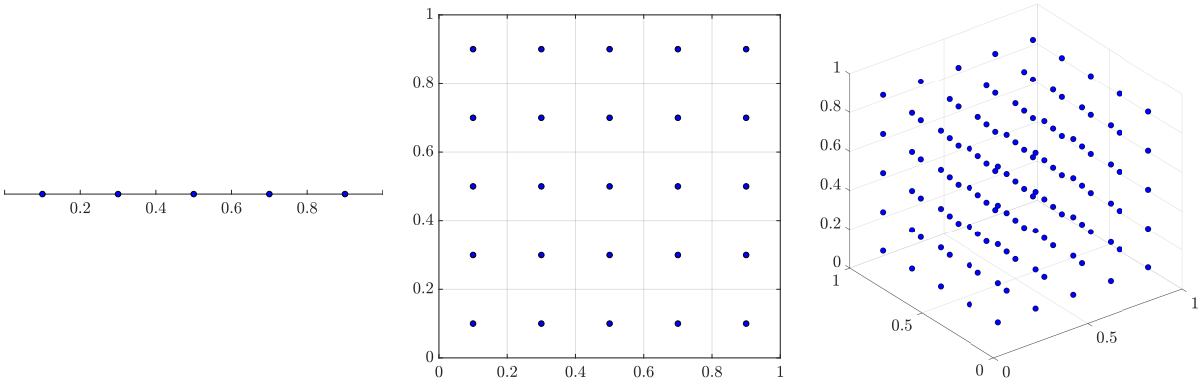
That being said, there are two important requirements for a good experimental design[215, 214]. First, the sampling space should be space-filling, with design points uniformly spread over the design space. Second, if the user knows which dimensions are important, two design points should not share any coordinate value. This requirement leads to a better understanding of each variable alone and is often denominated as projective property[215].

The accuracy of the metamodel improves as the sample size increases, which is consistent with intuition[192, 216]: the more information is given to the model, the better it may represent the true function. However, there is a clear trade-off here between accuracy and efficiency, as the use of a larger data set leads to a higher number of costly high-fidelity evaluations.

The number of sampling points required to provide sufficient accuracy to the model increases exponentially with the problem dimensionality[217]. This is a major concern in surrogate modeling and is usually referred to as the Curse of Dimensionality[40, 41]. For instance, take the full-factorial example, where each dimension may have five equally spaced points. Figure 19 depicts the sampling space for one, two, and three dimensions. The number of sampling points grows from 5 to 25, and then to 125. This aspect makes dimensionality reduction methods very important, especially in black-box functions with many design variables[218, 219, 220]. Examples of such methods are the screening, where the most important variables are identified, and the decomposition, where a high dimensionality problem is decomposed into multiple low dimensionality ones.

Full-factorial designs are not flexible in the number of sampling points employed. Also, even though these designs are space-filling, they have a very poor projective property[214]. This way, these experimental designs are not very efficient. A variety of Design of Experiments

Figure 19 – Full-factorial designs.



(a) One dimension (b) Two dimensions (c) Three dimensions

Source: the author

(DoE) techniques have been proposed in the literature to better tackle these issues.

In most of these techniques, the user is actually able to choose how many data points will be considered. Many researchers have proposed different methods to suggest an adequate number of points required to fit a sufficiently accurate model, which is often a function of the number of design variables m. For reference, one may consider the number of points required to fit an interpolating second-order polynomial[221, 198]:

$$n = \frac{(m+1)(m+2)}{2} \tag{5.1}$$

However, the number of sampling points might be too large in high-dimensional problems. Other authors suggest that a lower number of sampling points is enough to provide sufficient accuracy, such as $n = 10m$, $n = 5m$, or even $n = 2m$[222, 223, 215, 224, 206, 225]. For constrained optimization, it might be wise to increase the initial number of sampling points to make sure that the feasible region will be found. Also, it is important to remember that the evaluation of the initial data can be easily run in parallel[212]. Thus, if parallelism is available, it is interesting to set the number of initial sampling points as a multiple of the number of cores on the machine to maximize efficiency.

It is very important to use a robust sampling technique[221] since it is unreasonable to re-evaluate an entirely new initial sample if a model has a poor performance. Examples of commonly employed techniques are the Hammersley Sequence Sampling (HSS), the Sobol Sequence Sampling (SSS), the Latin Hypercube Sampling (LHS), or even the Random Sampling (RS). Some researchers have compared multiple sampling criteria in the context of surrogate modeling[192, 36, 215]. For the deterministic sampling, Hammersley Sequence Sampling (HSS) tends to be better for a small number of variables, but it is surpassed by the Sobol Sequence Sam-

pling (SSS) for high dimensionalities. However, most researchers employ stochastic sampling, where the LHS is the preferred choice.

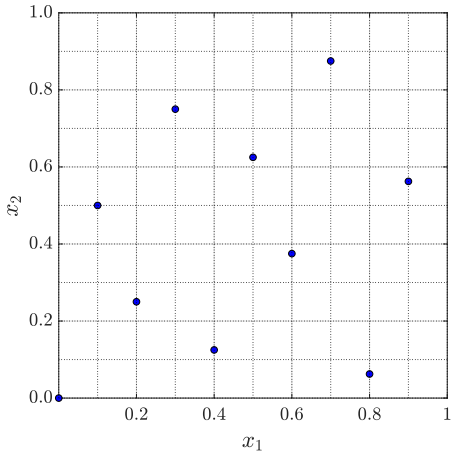
This work will employ either the HSS or the LHS in the initial designs. These techniques will be further described in the following.

5.1.1 Hammersley Sequence Sampling

The Hammersley Sequence Sampling (HSS) is a low-discrepancy sampling based on the Hammersley points[226], a quasi-random sequence that uses successive primes as bases[221]. With the same input parameters, the technique is deterministic, always generating the same design.

Even though the approach is based solely on the projective property[215], these designs are very uniform and space-filling for a low-dimensional space[192, 221], even though it is surpassed by the Sobol Sequence Sampling (SSS) for $m > 6$[215, 36]. Figure 20 shows an example of a 10-point HSS design on a two-dimensional space. The design is very uniform, as the sampling points are scattered in almost the entire design space.

Figure 20 – Example of the Hammersley Sequence Sampling (HSS).



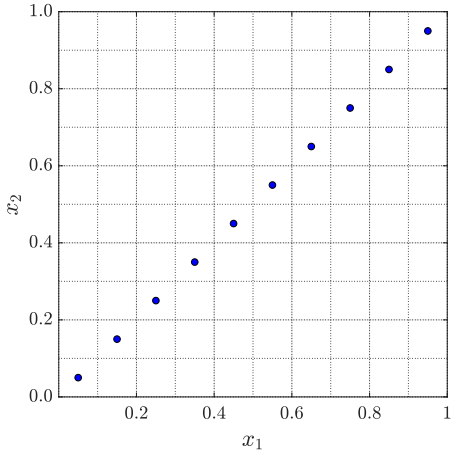
Source: the author

5.1.2 Latin Hypercube Sampling

The Latin Hypercube Sampling (LHS) is a stratified stochastic sampling method[36]. To generate n data points, each dimension is divided into n equally spaced intervals (or bins). Then, one point is selected at random in each interval, and the following data points can not be selected in the same interval. The procedure is repeated until the entire sample is generated.

Once again, this design is based on the projective property alone, while the randomness often takes care of the space-filling property. However, that is no guarantee that the LHS will cover the entire design space[215]. Figure 21 depicts a very poor design, which may indeed be generated by the LHS technique.

Figure 21 – Example of a poor LHS.



Source: the author

Even so, the fact is that the LHS is the most popular sampling technique for modeling purposes[53, 56, 47, 70]. To lower the chance of using poor designs, the user may employ an Optimized Latin Hypercube Sampling (OLHS), where the maximization of the minimum distance between two points is performed[40]:

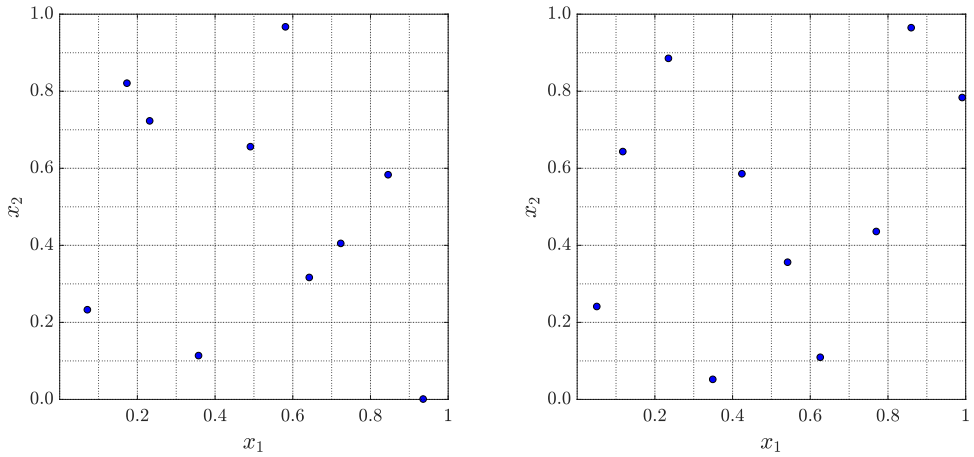
$$d_p(\mathbf{x}^{(i)},\mathbf{x}^{(j)}) = \left(\sum_{k=1}^{m} |x_k^{(i)} - x_k^{(j)}|^p\right)^{\frac{1}{p}} \tag{5.2}$$

where, for $p = 2$, d_p is the Euclidean distance. However, this procedure can be very computationally intensive[40, 212], especially for larger m. Alternatively, one may simply test N different data sets, and choose the one with the highest d_p[221, 2]. In this work, this strategy will be denominated as LHS$_N$. Figure 22 depicts a usual LHS and a LHS$_{20}$. While the former has $d_p = 0.114$, the latter has $d_p = 0.224$. For reference, the HSS depicted in Figure 20 has $d_p = 0.236$. Thus, this particular LHS$_{20}$ was almost capable to surpass the HSS by this performance criterion.

5.1.3 Handling of discrete variables

In problems with discrete variables, the sampling points should also be located in places where a design can be drawn. The easiest way to perform this is to use a mapping approach by setting each point to the closest discrete design[214, 58]. Figure 23 shows an example

Figure 22 – Example of Latin Hypercube Sampling (LHS).

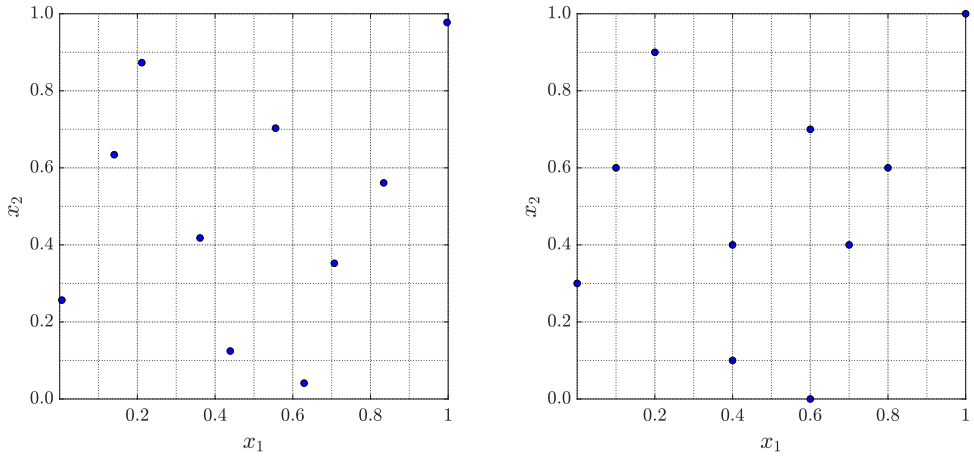


(a) Usual LHS (b) LHS_{20}

Source: the author

of an LHS_{20} re-mapped to a discrete space, where each variable can be set to 11 different equally spaced values.

Figure 23 – Mapping of sampling points to the discrete space.



(a) Continuous design (b) Discrete design

Source: the author

When performing this approach, two data points may be mapped to the same discrete position. When this occurs, the sample can be redrawn until a valid sample is built. It should be noted that this can be checked before the evaluation of the data points. However, some properties of the DoE techniques may be lost with mapping[227]. For instance, a mapped LHS may have two different data points with the same coordinate values.

5.2 Surrogate modeling

After data is prepared, the user may fit an adequate response surface to make future predictions. Given a training set of input-output pairs $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{n}$, the model may provide an approximate surface that can replace the true function for a cheaper alternative. One should note that the input space should be normalized so that the size of the design space does not affect the model hyper-parameters. This can be performed by a transformation of the design space to bring all designs to the domain $[0,1]^m$.

Traditional response surface methodology is often related to low-order polynomial regression[35]. For a smooth introduction on the matter, a quick explanation of those basic models will be made. Besides, most of the issues that affect those also remain, on some sort, in more robust techniques The generalized form of a univariate polynomial regression is given by[41]:

$$\hat{y}(x) = \sum_{i=1}^{n} w_i \, \psi_i(x), \text{ where } \psi_i(x) = x^{i-1} \tag{5.3}$$

where w_i are the coefficients of the polynomial and $\hat{y}(x)$ is the model prediction for a given x. The function $\psi_i(x)$ is known as the basis function, and, here, it is represented by a monomial of order $i-1$. The summation of these monomials produces a polynomial regression. The basis functions can be understood as the building blocks that will assist in creating robust approximations, and more complex models often employ different types of basis functions. In matrix form, Eq. (5.3) can be written as:

$$\hat{\mathbf{y}} = \mathbf{w}^T \, \boldsymbol{\psi} \tag{5.4}$$

It is important to note that, while $\hat{\mathbf{y}}$ can be a nonlinear function of $\mathbf{x}$, it is a linear function of $\mathbf{w}$. This aspect makes the fitting and analysis of such models much easier[41]. The prediction in the sampling points can be given by:

$$\begin{aligned}
&w_1 \, \psi_1(x_1) + w_2 \, \psi_2(x_1) + \ldots + w_m \, \psi_m(x_1) = \hat{y}(x_1) \\
&w_1 \, \psi_1(x_2) + w_2 \, \psi_2(x_2) + \ldots + w_m \, \psi_m(x_2) = \hat{y}(x_2) \\
&\qquad\qquad\qquad\qquad \vdots \\
&w_1 \, \psi_1(x_n) + w_2 \, \psi_2(x_n) + \ldots + w_m \, \psi_m(x_n) = \hat{y}(x_n)
\end{aligned} \tag{5.5}$$

where x_i is the i-th sampling points. In matrix form:

$$\boldsymbol{\Psi} \mathbf{w} = \hat{\mathbf{y}} \tag{5.6}$$

The building of the model is related to the fitting of the model parameters. In this case, this is the vector of coefficients $\mathbf{w}$, also known as the weight vector. The optimal $\mathbf{w}$ can be found by minimizing an error function:

$$E(\mathbf{w}) = \sum_{i=1}^{n} [y_i - \hat{y}(x_i, \mathbf{w})]^2 = (\mathbf{y} - \boldsymbol{\Psi}\mathbf{w})^T (\mathbf{y} - \boldsymbol{\Psi}\mathbf{w}) \tag{5.7}$$

which accounts for the squared error between the model predictions at the sampling points $\mathbf{x}$ and their respective true responses $\mathbf{y}$, for a given $\mathbf{w}$. This approach is very straightforward, and it also maximizes the likelihood that the model represents the data well[41]. Since $\hat{\mathbf{y}}$ is linear with respect to the weight vector, $E(\mathbf{w})$ is quadratic. Thus, it is easy to determine the extreme of the function by:

$$\frac{\partial E(\mathbf{w})}{\partial \mathbf{w}} = 0 \tag{5.8}$$

which means that the error is minimized for:

$$\mathbf{w} = \left(\boldsymbol{\Psi}^T \boldsymbol{\Psi}\right)^{-1} \boldsymbol{\Psi}^T \mathbf{y} \tag{5.9}$$

Figure 24 depicts examples of different order polynomial regressions in the fitting of nine data points afflicted by very small noise. It can be seen that, as the order of the approximation grows from one to four, the model fits the true function much better. However, the eighth-order model is not suitable to approximate the true function since it overfits the data.

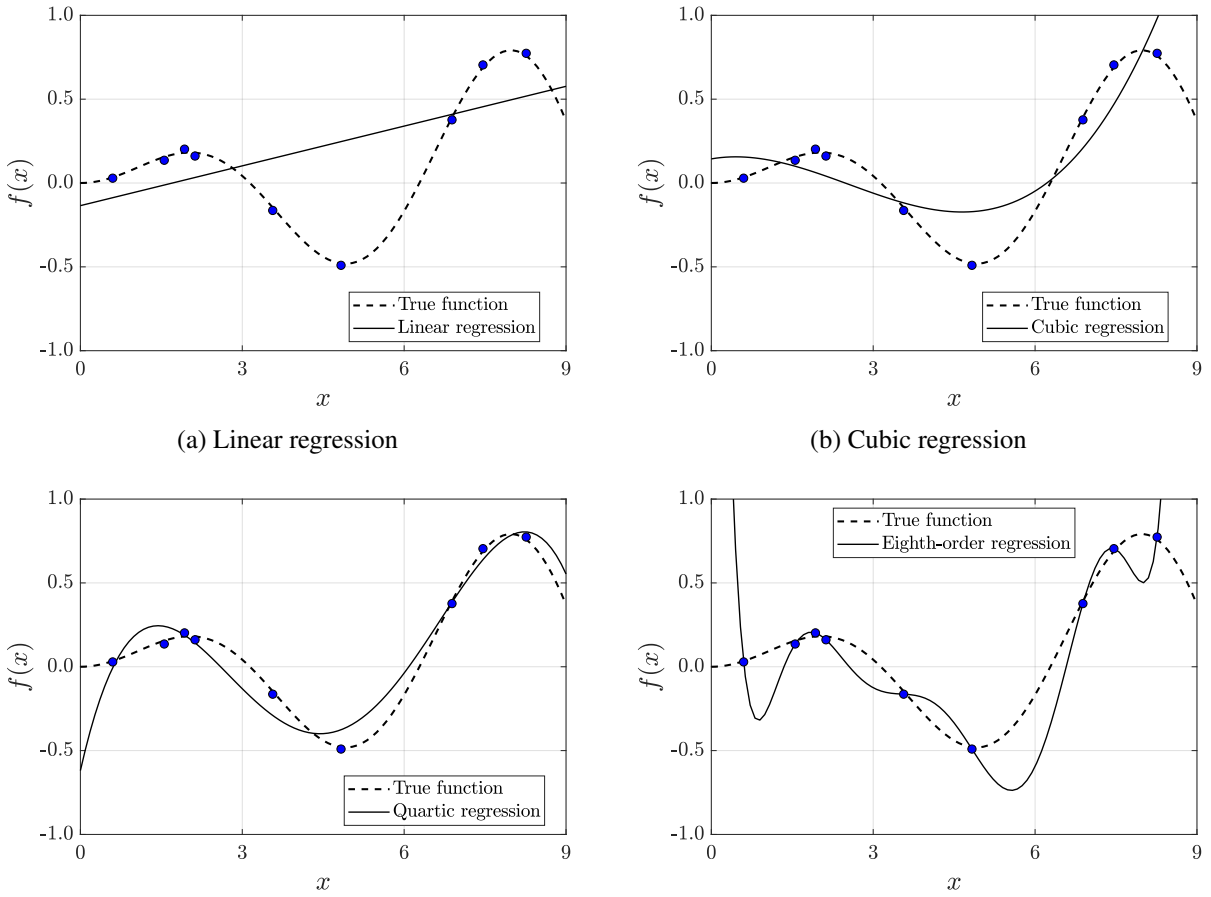
In the eighth-order model, the set of parameters $\mathbf{w}$ that minimizes the error is the one that interpolates the data, where $E(\mathbf{w}) = 0$. Thus, by realizing that $\hat{y}(x_i) = y(x_i)$ at all sampling points x_i, one may solve the minimization problem simply by solving the linear system:

$$\boldsymbol{\Psi}\mathbf{w} = \mathbf{y} \quad \Rightarrow \quad \mathbf{w} = \boldsymbol{\Psi}^{-1}\mathbf{y} \tag{5.10}$$

In this case, since $\boldsymbol{\Psi}$ is a square matrix, the system is determined and has a single unique solution. It will be shown that, for the models employed in this work, this system can always be determined, since the effective number of parameters adapts automatically to the size of the sample[41].

However, for the eighth-order model, this solution is not adequate since it is prone to over-fitting, as seen in Figure 24. There are regularization approaches that could attempt to fix this approximation. Also, over-fitting becomes less of a problem as the size of data increases[41]. However, it is important to realize that a more complex model does not always present a better approximation of a given function.

Figure 24 – Different order polynomial regressions.



(a) Linear regression

(b) Cubic regression

(c) Quartic regression

(d) Eighth-order regression

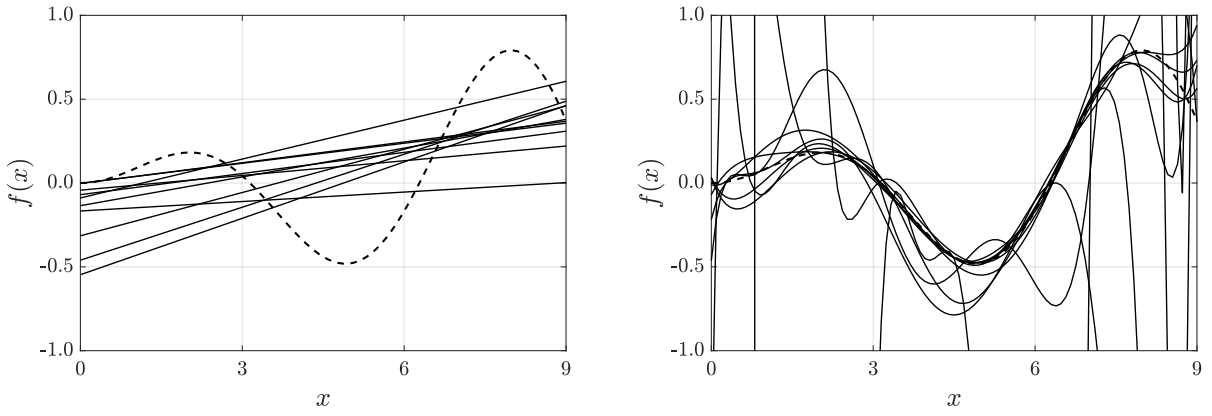
Source: the author

Another important aspect can be seen by analyzing how different models behave when the data is changed. Figure 25 shows 10 different linear regression models, with 10 different data sets $\mathcal{D}$ from the same output function. A similar plot is shown for the eighth-order approximation.

In the linear regression case, the 10 models created are similar to each other and, even though they fit the true function very poorly, most of them seem to at least capture its global trend. It is said that linear regression has a very high bias. Models with more bias make very strong assumptions about the nature of the data[228]. For instance, the linear regression implicitly assumes that the data is given by a linear function, which is definitely not the case for most engineering functions. Usually, models with a lower bias are preferred, since the estimated surface should have some freedom to better fit any unknown black-box function.

In the eighth-order regression, however, the predictions are all over the place. This means that model prediction is greatly influenced by which data is used for model building, and

Figure 25 – Models created using different data sets.



(a) Linear regression

(b) Eighth-order regression

Source: the author

it is said that these models have a very high variance. The variance is related to the number of model parameters and to the basis function ψ. Models with high variance are often discarded since they are more prone to over-fitting the data, thus being less robust.

The ideal algorithm should present low bias and low variance. However, in many cases, balancing both aspects is not trivial. Besides, a set of assumptions that works well in a domain may do very poorly in another[228]. This saying can be summarized by the very well-known "no free-lunch" theorem[229]. Ultimately, this means that it is not easy to select the best model for a given task *a priori*, without any knowledge about the true function. This example demonstrates that the choice of sampling points may greatly influence the model prediction, as already pointed out in the last section.

For engineering applications, polynomial regression models are very weak for a robust approximation since they are unable to provide a smooth global fit for most nonlinear functions[230, 231]. Also, it is not as simple to derive a formulation for multivariate spaces. Thus, commonly employed models for these cases are Artifical Neural Networks (ANNs), Support Vector Regression (SVR), Radial Basis Functions (RBFs), and Kriging[191, 40].

Jin, Chen and Simpson[222] presented a comparison study over four modeling techniques: polynomial regression, Kriging, RBF, and Multivariate Adaptive Regression Splines (MARS). The authors consider 13 mathematical benchmark problems and one engineering design problem in an attempt to represent different features found in engineering design problems. The RBF presented the best overall results. The Kriging model outperformed the RBF for low-order problems, but it was also the most time-consuming method. On a study focused on the same modeling techniques, Simpson, Lin and Chen[192] also advocate for the use of Kriging and RBF,

which presented good approximations for a wide range of DoE techniques and sample sizes. On the other hand, Kim, Lee and Choi[216] performed a comparison between Moving Least Squares (MLS), Kriging, RBF, and SVR, claiming that Kriging and Moving Least Squares (MLS) showed the most accurate approximations.

Williams and Cremaschi[225] performed a comparison between ANNs, RBF, MARS, random forests, and SVR in 127 different analytical test functions. The results show that, for all models, there is a clear relation between dimensionality and model accuracy. Even though the best modeling approach seems to depend on the data characteristics, the authors found that, in general, MARS and RBF were able to find the most accurate results. For Surrogate-Based Optimization (SBO), SVR, random forests, and RBF were able to find designs closer to the global optimum.

Díaz-Manríquez, Toscano-Pulido and Gómez-Flores[64] compared the polynomial approximation, RBF, Kriging, and SVR for the Surrogate-Based Optimization (SBO) using evolutionary algorithms. Six well-known mathematical benchmark functions were tested, with varying modality and dimensionality. The authors showed that the Kriging model presented the best results for low-dimensionality problems ($m < 15$), while the RBF outperforms Kriging in problems with high-dimensionality. Modality seems to have less effect on the definition of the best technique. With the same focus on evolutionary algorithms, Valadão and Batista[38] compared the RBF with different Kriging models, namely Ordinary Kriging (OK), Universal Kriging (UK), and Blind Kriging (BK). The authors show that the most accurate approach is the Ordinary Kriging (OK) and, even though the RBF does present the lowest computational cost, it also presents the worst performance, by a small margin.

Nik et al.[65] performed a comparison between different modeling techniques for the optimization of Variable Stiffness Composites (VSCs). Kriging and RBF presented the highest accuracy, where Kriging was slightly better in terms of robustness. However, the authors, once again, show that the RBF outperforms Kriging for high-dimensionality problems.

In short, most of these works show that the RBF and the Kriging are the most accurate models for a variety of applications. Usually, the RBF presents the lowest computational cost and seems to behave better for high-dimensional problems, while the Kriging presents a better approximation for low-dimensional functions. However, one should note that there are exceptions to this trend, as the development of general guidelines for model selection might not be practical. Model performance depends on many characteristics of the problem, such as dimensionality and

degree of nonlinearity[198, 232, 233, 225].

In this work, a major focus will be given to Gaussian Process (GP) models, since they have very nice mechanisms to assist in the selection of infill points. This aspect will be better discussed later, but, for now, it is worth noting that both the Kriging and, to some extent, the RBF can be viewed as forms of Gaussian Process (GP) models[41]. Also, since their prediction is based on the distance between $\mathbf{x}$ and the data points, there is a natural approach for modeling discrete spaces[234, 235, 58]. The following sections further describe each of these models.

5.2.1 Radial Basis Functions

The Radial Basis Function (RBF) model was first proposed by Hardy[236] as a way to interpolate topography and other irregular surfaces. Recently, it was shown to present a great predictive capability for structural problems related to composite structures[237]. In this model, the basis ψ_i is a function of the radial distance between $\mathbf{x}$ and the i-th basis center[41] so that:

$$\hat{y}(\mathbf{x}) = \sum_{i=1}^{n} w_i \, \psi_i(r), \text{ where } r = ||\mathbf{x} - \mathbf{c}_i|| \tag{5.11}$$

In this work, the basis centers $\mathbf{c}_i$ are considered to be the data points $\mathbf{x}_i$. This equation can be written in matrix form as:

$$\hat{\mathbf{y}} = \mathbf{w}^T \, \boldsymbol{\psi} \tag{5.12}$$

which is the same as Eq. (5.4). In fact, the major difference between these models is the basis functions employed. For the RBF, common basis functions include linear and cubic polynomials, multiquadric and Gaussian basis functions[40, 238]. These are shown on Table 10.

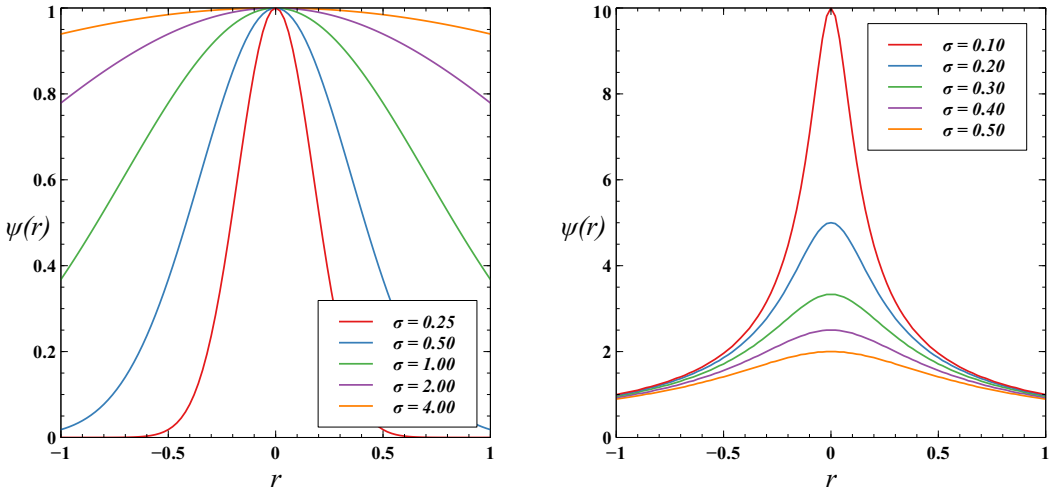
Table 10 – Common basis functions for Radial Basis Functions (RBFs).

Name	$\psi(r)$
Linear	r
Cubic	r^3
Multiquadric	$(r^2+\sigma^2)^{1/2}$
Inverse multiquadric	$(r^2+\sigma^2)^{-1/2}$
Gaussian	$\exp(-r^2/\sigma^2)$

In some of these functions, in addition to the distance r, ψ also depends on a parameter σ known as the width, spread, or shape parameter[40, 221]. On parametric basis

functions, σ should be set before fitting the model. Figure 26 shows a comparison between basis functions with different σ for the Gaussian and inverse multiquadric basis functions.

Figure 26 – Influence of the width parameter on different basis functions.



(a) Gaussian (b) Inverse multiquadric

Source: the author

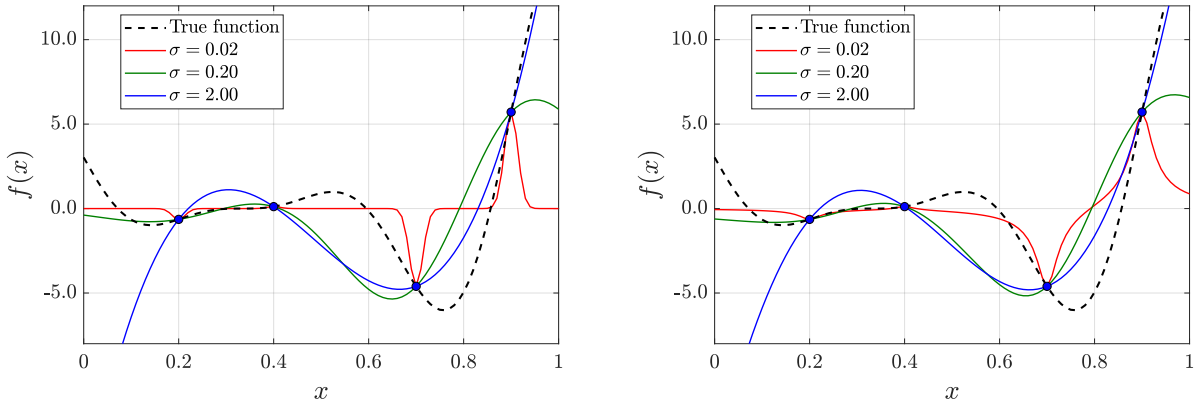
For these basis functions, the width parameter is related to the amplitude of the basis function and, for higher widths, the influence exerted by a given basis' center covers a larger region of the design space[2]. However, a higher width may cause instabilities, since the interpolation matrix may numerically cease to be positive definite[46, 239]. In any model, hyper-parameters are very important for correct model prediction. Figure 27 presents the prediction of the RBF model using different σ. For small values, the prediction resembles a "needles in a haystack" function, where only regions near sampling points present accurate predictions. On the other hand, greater values make the predicted surface smoother than it should be[2].

There are many different ways to define σ. While some researchers attempt to optimize the parameter[240, 46, 241, 242, 232], this procedure can be non-trivial and too costly, depending on the size of the data set. On the opposite spectrum, Haykin[243] presents a simple technique to define the width for Gaussian RBFs by:

$$\sigma = \frac{d_{max}}{2n} \tag{5.13}$$

where d_{max} is the maximum distance between two points in the data set. Later, other researchers proposed similar analytical formulations, considering that the width should also be dependant on the problem dimensionaty m[244, 47]. However, the disadvantage of these techniques is that they are *ad-hoc* in nature, as they make no consideration over the true function response.

Figure 27 – Influence of the width parameter on the prediction using different basis functions.



(a) Gaussian

(b) Inverse multiquadric

Source: the author

Alternatively, one may employ a cross-validation technique, such as the k-Fold Cross Validation (k-FCV)[203]. Here, the data set is divided into k groups, and, at each iteration, one of the groups is used as a validation set. Some different values of σ are tested, and the error is computed from each validation group. The width value with the lower mean squared error is then selected.

The number of groups k is usually set to 5 or 10, or as a function of the number of sampling points n. In the particular case where $k = n$, the approach becomes the so-called Leave-One-Out Cross Validation (LOOCV). Usually, cross-validation approaches have a very low bias. Since the training is performed with a considerable part of the whole data set, the difference between the calculated error and the real error is very small[245]. That being said, the lowest bias is seen on the Leave-One-Out Cross Validation (LOOCV). However, the variance is also much higher in this approach[246]. In fact, instability may cause the results to be worse than approaches with a lower k[247, 2]. Besides, the LOOCV may be very time-consuming for a large number of sampling points[203, 2]. Regarding the trial widths, for the Gaussian basis function, Sobester, Leary and Keane[55] tested 20 values of σ at each iteration, all in the domain $[10^{-2}, 10]$[1]. The same approach will be adopted in this work.

Finally, after defining the σ parameter, the user may fit the model. Similar to the polynomial regression case, the weight vector can be evaluated by realizing that, on the sampling

[1] Sobester, Leary and Keane[55] consider a slightly different Gaussian function in their work. To maintain the same effective width values, all values will be multiplied by $\sqrt{2}$.

points, $\hat{y}(x_i) = y_i$. By interpolation:

$$\begin{bmatrix} \psi_{11} & \psi_{12} & \cdots & \psi_{1n} \\ \psi_{21} & \psi_{22} & \cdots & \psi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{n1} & \psi_{n2} & \cdots & \psi_{nn} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} y(x_1) \\ y(x_2) \\ \vdots \\ y(x_n) \end{bmatrix} \tag{5.14}$$

Thus, the weight vector is given by:

$$\mathbf{w} = \mathbf{\Psi}^{-1}\,\mathbf{y} \tag{5.15}$$

Since this work deals with deterministic functions, interpolation should do fine[248]. However, alternative approaches can be employed if the user is concerned with noise in the data[41, 47]. In the RBF model, the interpolation matrix $\mathbf{\Psi}$ is known as the Gram matrix. To prevent instabilities, the Gram matrix should not be singular, which is true for a variety of basis functions (including the Gaussian and the inverse multiquadric) if all data points are distinct[249, 250, 40].

5.2.2 Kriging

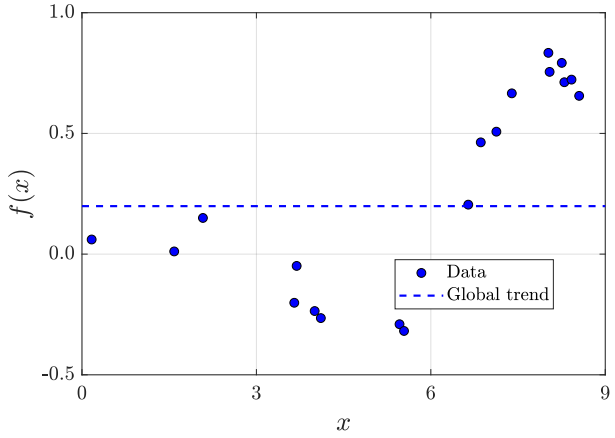
The Kriging model was first proposed by Krige[251], but it was only brought to engineering design by Sacks et al.[252]. The model is known by its almost optimal prediction if sampling points are taken at nearby locations[253]. In its general form, Kriging can be seen as the sum between a global trend $g(\mathbf{x})$ and its autocorrelated localized deviations $Z(\mathbf{x})$:

$$\hat{y}(\mathbf{x}) = g(\mathbf{x}) + Z(\mathbf{x}) \tag{5.16}$$

where $\hat{y}(\mathbf{x})$ is the Kriging prediction. In the Universal Kriging (UK), the global trend is given by polynomial functions, usually linear or quadratic[49, 198, 38]. However, the most common approach is the Ordinary Kriging (OK), which considers that the global trend is given by a constant term. Figure 28 depicts this case. This is the most popular Kriging algorithm by far, and is the version that is going to be used in this work.

In a Kriging model, the observed responses y_i are assumed to come from a stochastic process with mean μ (even if they come from a deterministic computer code)[40]. These responses are related to each respective m-dimensional input design vector $\mathbf{x}_i$. For a given data set $\mathfrak{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{n}$, we may assume that the correlation between the responses comes from a measure of similarity between their input data. For example, if we have $\mathfrak{D} = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2)\}$ and want to predict the response $\hat{y}_p$ for the input $\mathbf{x}_p$, it is fair to assume that, if $\mathbf{x}_p$ is closer to $\mathbf{x}_1$,

Figure 28 – Ordinary Kriging (OK) global trend.



Source: the author

$\hat{y}_p$ should also be closer to y_1 than to y_2. This assumption also means that the response function should be smooth and continuous.

The measure of similarity can be evaluated by a wide range of arbitrarily picked functions[198]. However, the most popular correlation function is, by far, the Gaussian kernel, where the correlation between the two responses y_i and y_j is given by:

$$\text{cor}\left[y_i, y_j\right] = \exp\left(-\sum_{k=1}^{m} \theta_k \left|x_{i,k} - x_{j,k}\right|^{p_k}\right) \tag{5.17}$$

Here, the parameters θ_k and p_k should always be positive. Using this kernel function, if two points are very close to each other, the correlation between them is closer to 1. However, if these points are infinitely far away from each other, their correlation is zero.

This way, the correlation matrix can be built by computing the correlation between each sampled response:

$$\boldsymbol{\Psi} = \begin{bmatrix} \text{cor}\left[y_1, y_1\right] & \text{cor}\left[y_1, y_2\right] & \dots & \text{cor}\left[y_1, y_n\right] \\ \text{cor}\left[y_2, y_1\right] & \text{cor}\left[y_2, y_2\right] & \dots & \text{cor}\left[y_2, y_n\right] \\ \vdots & \vdots & \ddots & \vdots \\ \text{cor}\left[y_n, y_1\right] & \text{cor}\left[y_n, y_2\right] & \dots & \text{cor}\left[y_n, y_n\right] \end{bmatrix} \tag{5.18}$$

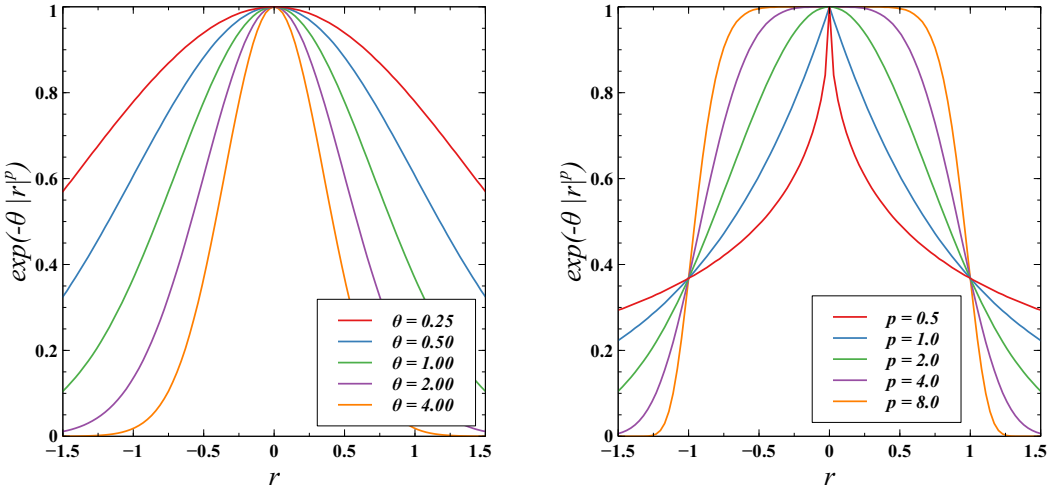
From Eq. (5.16), $Z(\mathbf{x})$ can be assumed to be a realization of a stochastic process with mean zero and covariance given by:

$$\text{cov}(\mathbf{y}, \mathbf{y}) = \sigma^2 \boldsymbol{\Psi} \tag{5.19}$$

where σ^2 is the process variance. The computation of $\boldsymbol{\Psi}$ involves the determination of the hyper-parameters $\mathbf{p}$ and $\boldsymbol{\theta}$. Just like the RBF model, these parameters are very important for the

model prediction. Figure 29 presents how the correlation function behaves for different values of θ_k and p_k. The θ_k parameter is similar to the width parameter σ from the Gaussian RBF model, as it regulates the influence exerted by a sampling point in the design space[40]. It is worth noting that, on the Kriging model, the parameter is related to each design variable k. Thus, the optimal value of θ_k can be a good estimator for how important the k-th variable is to the problem[40, 254]. The parameter p, however, controls how smooth the function is: for a lower p, the correlation between two data points drops very quickly. A very low value for p means that there is no immediate correlation between two points[40].

Figure 29 – Behavior of the Gaussian kernel with different hyper-parameters.



(a) Variation of θ, $p = 2$

(b) Variation of p, $\theta = 1$

Source: the author

Different methods have been proposed for defining the model hyper-parameters, such as the Maximum Likelihood Estimator (MLE) and Cross Validation (CV) techniques. Usually, the former is more reliable in defining the optimal value of these parameters[248, 255]. In terms of the sample data, the likelihood function is given by[40]:

$$L(\mathbf{y}|\mu,\sigma) = \frac{1}{(2\pi\sigma^2)^{n/2}|\mathbf{\Psi}|^{1/2}} \exp\left[-\frac{(\mathbf{y}-\mathbf{1}\mu)^T\mathbf{\Psi}^{-1}(\mathbf{y}-\mathbf{1}\mu)}{2\sigma^2}\right] \tag{5.20}$$

which expresses how well the model represents its sampling points. This can be seen as the likelihood of $\mathbf{y}$ for a process with mean μ and variance σ^2. The maximization of the likelihood estimator achieves a bias-variance trade-off by penalizing both low-data likelihood and model complexity[256]. It is very common to consider the natural logarithm of this function before performing the maximization[40]:

$$\ln L = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2}\ln|\mathbf{\Psi}| - \frac{(\mathbf{y}-\mathbf{1}\mu)^T\mathbf{\Psi}^{-1}(\mathbf{y}-\mathbf{1}\mu)}{2\sigma^2} \tag{5.21}$$

This procedure does not change the maximum likelihood estimators, since the natural logarithm is a monotonic function. Finally, by differentiating in terms of μ and σ^2 and equating to zero, we find the Maximum Likelihood Estimators (MLEs) for the mean and the variance[40]:

$$\begin{aligned} \hat{\mu} &= \frac{\mathbf{1}^T \boldsymbol{\Psi}^{-1} \mathbf{y}}{\mathbf{1}^T \boldsymbol{\Psi}^{-1} \mathbf{1}} \\ \hat{\sigma}^2 &= \frac{(\mathbf{y} - \mathbf{1}\hat{\mu})^T \boldsymbol{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu})}{n} \end{aligned} \tag{5.22}$$

Now, we still need to assign values to the hyper-parameters $\boldsymbol{\theta}$ and $\mathbf{p}$. Substituting the estimators for the mean and variance in Eq. (5.21) and removing the constant terms we end up with the so-called concentrated ln-likelihood function[40]:

$$\ln L \approx -\frac{n}{2} \ln(\hat{\sigma}^2) - \frac{1}{2} \ln |\boldsymbol{\Psi}| \tag{5.23}$$

A straightforward approach for finding $\boldsymbol{\theta}$ and $\mathbf{p}$ would be to simply differentiate this function, similar to the estimators for the mean and the variance. However, this function is nonlinear and multi-modal, which means that there is no unique solution. Thus, we should use an optimization technique to define the optimal value for the hyper-parameters. Since Eq. (5.23) is multi-modal in nature and the assessment of its gradient is time-demanding[217], these optimal values will be found, in this work, using bio-inspired search methods. Furthermore, to reduce complexity, a usual approach is to consider that $p_k = 2.0$[248, 231, 257], which guarantees that the basis functions are smooth, as shown in Figure 29, and makes the correlation function very similar to the Gaussian kernel of the RBF.

This way, the assessment of the MLE for $\boldsymbol{\theta}$ is an unconstrained optimization problem which can be defined as:

$$\begin{cases} \text{find} & \boldsymbol{\theta} = \{\theta_1, \theta_2, \ldots, \theta_m\} \\ \text{that minimizes} & -\ln L(\boldsymbol{\theta}) \\ \text{with} & \boldsymbol{\theta}_L \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_U \end{cases} \tag{5.24}$$

Since the design variables are always continuous, PSO or DE algorithms can find appropriate values for the parameter [217, 2]. Besides, due to how the parameter θ_k is considered in Eq. (5.17), the variable bounds $\boldsymbol{\theta}_L$ and $\boldsymbol{\theta}_U$ are usually given in logarithmic scale. After the optimal $\boldsymbol{\theta}$ is found, $\boldsymbol{\Psi}$, $\hat{\mu}$, and $\hat{\sigma}$ should be computed one last time to finish model building.

Similar to the RBF, the correlation matrix $\boldsymbol{\Psi}$ should be non-singular to prevent instabilities from arising. However, as the sample size increases, the condition number may increase very rapidly, thus resulting in ill-conditioned matrices. To prevent such instabilities, the

user may add a very small value τ to the diagonal of $\boldsymbol{\Psi}$, which should not affect the model itself or its interpolation properties. This value depends on the scale of the problem, but a value of $\tau = 10^{-8}$ is usually sufficient if the input design vectors are normalized[258, 255].

It should be noted that the biggest practical limitation of the Kriging model (and also the RBF, to a lesser extent) is precisely the need of repeatedly inverting large, dense, and frequently ill-conditioned matrices[259]. Emmerich, Giannakoglou and Naujoks[260] states that the time complexity for the entire process of training the Kriging model is $O(N n^3 m + n^2 m)$, or $O(N n^3 m)$ for higher n. Here, N is the number of times Eq. (5.23) is evaluated. It is important to note that the time spent to determine the Kriging hyper-parameters can be significant, especially when dealing with high-dimensional problems with many sampling points[261, 40].

Finally, the definition of a new predicted $\hat{y}(\mathbf{x}_p)$ should be consistent with the data and with the correlation function. The new prediction located at $\mathbf{x}_p$ can be found by augmenting the correlation matrix:

$$\tilde{\boldsymbol{\Psi}} = \begin{bmatrix} \boldsymbol{\Psi} & \boldsymbol{\psi} \\ \boldsymbol{\psi}^T & \operatorname{cor}[\hat{y}(\mathbf{x}_p), \hat{y}(\mathbf{x}_p)] \end{bmatrix} \tag{5.25}$$

where $\boldsymbol{\psi}$ is a column vector with $\psi_i = \operatorname{cor}[y_i, \hat{y}(\mathbf{x}_p)]$, for $i = 1, 2, \ldots, n$. Once again, $\hat{y}(\mathbf{x}_p)$ can be found by differentiating the concentrated ln-likelihood function in terms of the new prediction and equating to zero, which results in the estimator[40]:

$$\hat{y}(\mathbf{x}) = \hat{\mu} + \boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu}) \tag{5.26}$$

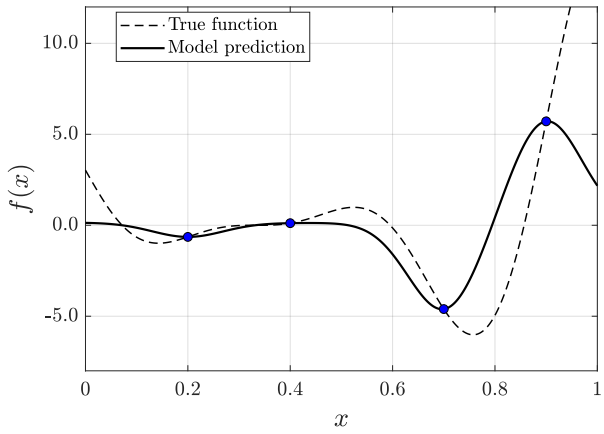
It is important to note that, after the model is built, $\boldsymbol{\Psi}^{-1}(\mathbf{y} - \mathbf{1}\hat{\mu})$ does not change. This way, if multiple model predictions must be performed, one could store the solution of this linear system beforehand, making it so that the cost of future predictions is much lower.

Returning to Eq. (5.16), $\hat{\mu}$ can be understood as the global trend of the model, while the term $\boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1}(\mathbf{y} - \mathbf{1}\hat{\mu})$ represents the localized deviations. Finally, Figure 30 depicts the Kriging prediction for a simple one-dimensional function.

Before jumping to the next topic, it is interesting to discuss a parallel between RBF and Kriging models. First, in both models, we may use similarity-based basis functions with very similar hyper-parameters. Second, the term $\boldsymbol{\Psi}^{-1}(\mathbf{y} - \mathbf{1}\hat{\mu})$ is closely related to the evaluation of the weight vector $\mathbf{w}$ for the RBF model, and we can write Eq. (5.26) as:

$$\hat{y}(\mathbf{x}) = \hat{\mu} + \boldsymbol{\psi}^T \mathbf{w} \tag{5.27}$$

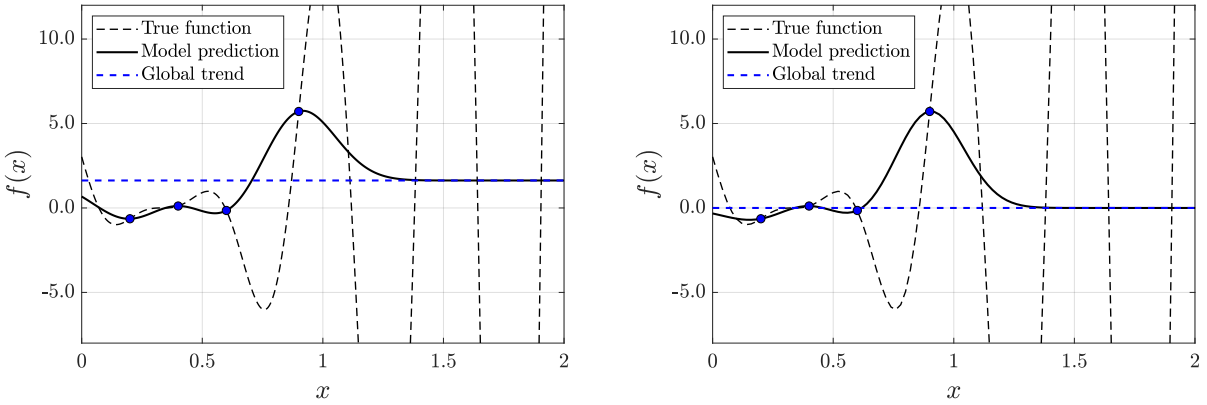
Figure 30 – Kriging prediction.



Source: the author

where $\mathbf{w} = \boldsymbol{\Psi}^{-1}(\mathbf{y} - \mathbf{1}\hat{\mu})$. The final term here closely resembles the RBF prediction but, this time, it is getting added to a constant term $\hat{\mu}$. In practice, this means that, as points get far away from the sampled data, the second term goes to zero, and the prediction $\hat{y}(\mathbf{x})$ gets closer to the global trend $\hat{\mu}$. However, on the RBF, the prediction goes to zero. This behavior is presented in Figure 31, and the implications of these aspects are very important for the selection of new data points.

Figure 31 – Behavior of the global trend far away from the sampled data.



(a) Kriging estimator

(b) RBF estimator

Source: the author

5.3 Infill criteria

After the model is built from the sampling data, it is possible to take advantage of its prediction to select new points in promising regions of the design space by the so-called adaptive sampling technique[223, 262]. This step is the core of the SAO approach, and an efficient

technique should be used to select new infill points.

Infill criteria are divided into two types[263, 264, 52]. In one-stage approaches, the definition of the new data is performed along with model building, e.g. in the Goal Seeking (GS) method, where new points are defined based on the MLE[263]. Two-stage approaches, however, are much more popular, where the model is built and, then, a given criterion is used to select new sampling points according to the model response. These techniques assume that the model is a good representation of the true function and, if that is the case, they may perform much better than the one-stage approaches. In this work, different two-stage approaches will be employed.

Thus, the user should be able to use the model to select new sampling data to be added to the data set $\mathfrak{D}$. The most straightforward approach is to simply add the point that minimizes the model prediction to the sample, assuming that it is also close to the true global optimum. In fact, many researchers adopt this simple procedure, which returns very good results when the model prediction is sufficiently accurate[207, 265, 51]. However, this is a pure exploitation method, which tends to be greedy, as it can easily get stuck on local optima[53, 266, 267, 268].

When there is insufficient information near the optimum point, exploration methods are also very important[244]. These methods favor the addition of new points in less sampled regions. Kitayama et al.[269] uses a density function that, when minimized, selects a new point in a region far from all other sampling points. It is also very common to employ techniques that start by exploring the design space and then, in later iterations, perform the exploitation[123, 270]. Chung, Park and Choi[270] considered a switching infill sampling criterion where, depending on how the SAO process is going, the algorithm decides if the following data point will exploit or explore the design space.

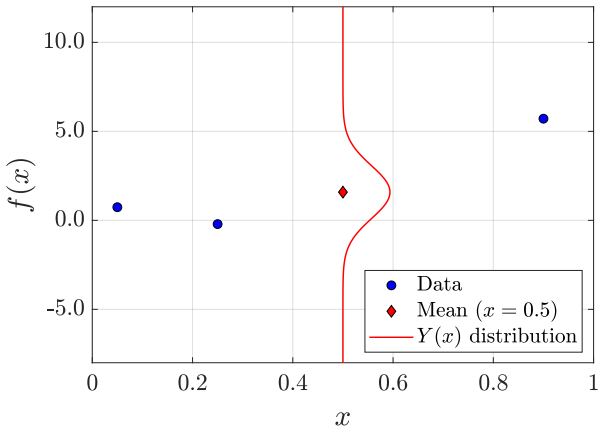
Another approach is to add multiple points in each iteration, each selected by a different criterion so that the model is able to both exploit and explore at the same time[56, 244, 271]. For example, in each iteration, Kitayama, Arakawa and Yamazaki[56] add new points related to both the minimization of the surrogate model (exploitation) and the minimization of the density function (exploration). A similar approach is explored in Balreira[272] for the minimization of laminate composites. These points can even be selected in parallel to take advantage of powerful machines with multiple cores[195, 273, 274, 275, 276]. However, it is worth noting that the parallel selection of infill points may lower the accuracy of the infill criteria since the definition of new sampling points should ideally consider all prior information available[277].

Finally, the most efficient approach is to employ a method that, itself, balances

exploitation and exploration, such as the bumpiness function proposed by Gutmann[54]. However, the balancing of these factors is not trivial[278], and the user may get stuck in a method that neither exploits nor explores very well. The major problem in these cases is that it is hard to quantify and compare the uncertainty of the model (related to the exploration) with its prediction (related to the exploitation) at a given point $\mathbf{x}$, thus making the balance between these two aspects very difficult.

The theory behind Gaussian Process (GP) models, however, provides a very good way of quantifying the uncertainty of a model[49]. GP models make the primary assumption that their responses are a realization of a stochastic process. Take, for example, the GP shown in Figure 32. Given a set of data $\mathcal{D}$, if we want to predict the response for $x = 0.5$, the GP allows us to draw a Normal distribution for this point. The mean of this Normal distribution is the most likely response, given our already sampled data. It is very important to differ the GP global trend $\hat{\mu}$ from this mean value. The former is the mean of the process itself, thus being the best prediction if we do not know x yet, while the latter is the mean for an already given x. Thus, this value can be denominated as the *posterior* mean, and is, basically, the prediction of our model $\hat{y}(x)$.

Figure 32 – Example of a Gaussian Process (GP).



Source: the author

However, there is also some uncertainty in the process, represented by the variance of the Normal distribution. If we are more confident with our prediction, the variance is lower. However, if not, the variance is larger, and the Normal distribution becomes flatter. By the theory of GPs, the posterior variance can be estimated by[228]:

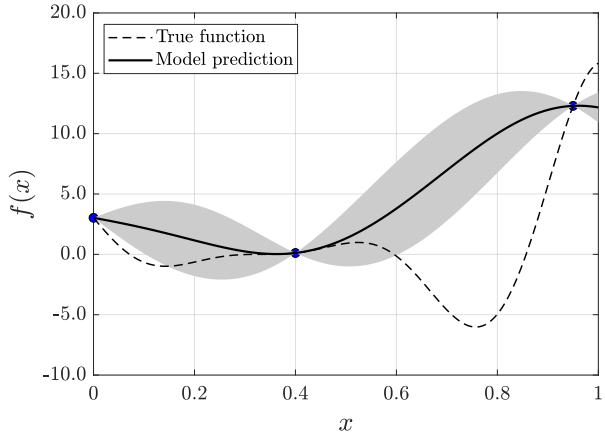
$$\hat{s}^2(\mathbf{x}) = \sigma^2 \left[1 - \boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1} \boldsymbol{\psi} + \frac{(1 - \mathbf{1}^T \boldsymbol{\Psi}^{-1} \boldsymbol{\psi})^2}{\mathbf{1}^T \boldsymbol{\Psi}^{-1} \mathbf{1}} \right] \tag{5.28}$$

where σ^2 is the variance of the process. The third term is often omitted from computations since it is often negligible[40]. This way, by having a GP model, the user has a very nice mechanism for trading off exploitation, represented by the posterior mean $\hat{y}(\mathbf{x})$, and exploration, represented by the posterior variance $s^2(\mathbf{x})$[223, 224].

Gaussian Process models are ideal when modeling smooth functions, where similar inputs should have similar outputs. This is the case for most engineering functions. Furthermore, both Kriging and RBF models can be seen as forms of GP[41]. For the Kriging model, this consideration is more clear, since most of their basic assumptions are the same. Thus, the Kriging model represents a GP with mean $\hat{\mu}$ and variance $\hat{\sigma}$. The assessment of the MLE for these parameters was discussed in Section 5.2.2. On the other hand, the RBF can be seen as a Gaussian Process (GP) with $\hat{\mu} = 0.0$ and $\hat{\sigma} = 1.0$[55, 44, 279]. This consideration can be done because the Gram matrix closely resembles the correlation matrix used in Kriging and usual GPs, also being able to represent the similarity between two data points[2].

Figure 33 shows the predictor of a Kriging model and its confidence interval, given by $\hat{y}(\mathbf{x}) \pm s(\mathbf{x})$. Note how the uncertainty increases as x gets far from the sampling points, and goes to 0 in the existing data. This occurs due to the consideration that the model should interpolate the data. Thus, there is no uncertainty in the already observed data. This can also be seen by analyzing Eq. (5.28). In the sampling points, $\boldsymbol{\psi}$ is a column of $\boldsymbol{\Psi}$, which means that $\boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1} \boldsymbol{\psi} = 1$ and, thus, $\hat{s}^2(\mathbf{x}) = 0$.

Figure 33 – Confidence interval of a Gaussian Process (GP).



Source: the author

[2] For the Gram matrix to act as a correlation matrix, its basis functions should have the same properties of a correlation function, i.e. they should result in higher values for points closer to basis' centers. Good examples are the Gaussian and the inverse multiquadric basis functions. Other researchers tried to propose more generalized methods to assess the uncertainty[280, 279], but these do not retain the same mathematical and statistical justification[37].

Usually, the selection of new data points is based on the optimization of a given function related to the model. Such a function is usually denominated an *acquisition* function[199]. Since the approximation often induces highly multi-modal landscapes, bio-inspired optimizers perform well in locating these promising regions in the design space[55, 64, 281, 34].

When employing a surrogate model to fit a set of data points, one often desires to learn about the function in the whole domain. However, in an optimization context, the user is interested solely in the point that minimizes a given objective function. Thus, the acquisition function should be biased to locate points that have a higher probability to minimize the objective function, while attempting to balance both the exploitation and the exploration aspects[267]. In unconstrained minimization problems, the user should try to select new points with a low $\hat{y}(\mathbf{x})$, where the model expects a lower value for the objective function, while also making some considerations over points with a high $\hat{s}(\mathbf{x})$. In theory, new points should be located where the model is more likely to improve upon the actual optimum. Such procedures are also known, in engineering design, as *error-based exploration*[53, 40] or, in statistics, as *Bayesian optimization*[199, 266, 224].

De Ath et al.[224] comment that this problem can be seen as a multi-objective optimization, where one tries to minimize $\hat{y}(\mathbf{x})$ while also maximizing $\hat{s}(\mathbf{x})$. The authors state that, ideally, a point from the Pareto optimal set of this multi-objective problem should be added to the sample. While the solution of this problem may over-complicate the procedure, alternative single-objective methods will be discussed which always select points from the Pareto set[224].

Thus, in this work, a focus will be given to the error-based methods. After Jones, Schonlau and Welch[53] popularized these techniques by the use of the Expected Improvement (EI) criterion, different papers have employed error-based exploration with different acquisition functions in structural optimization. Liu et al.[195] performed the shape optimization of transonic wings using three different error-based approaches: the minimization of the Lower Confidence Bound (LCB), the maximization of the Probability of Improvement (PI), and the maximization of the Expected Improvement (EI). The error found on the surrogate model is very minor compared to the expensive CFD analysis.

Regarding the optimization of composite structures, Ribeiro et al.[2] performed the maximization of the buckling load and the maximization of the fundamental frequency of FG plates using a SAO method based on the RBF. The authors perform the selection of new data points by two different methods, the Expected Improvement (EI) and the Weighted Expected

Improvement (WEI)[55], and show that the SAO approach can be up to 50 times faster than the conventional bio-inspired algorithms for the optimization. Maia, Parente and Melo[282] use a Kriging-based SAO in the optimization of FG plates and shallow shells. Here, the WEI performed slightly better than the EI, and the authors find a similar gain in efficiency. The authors state that, for more expensive high-fidelity functions, SAO can be even more advantageous.

The present work will employ four different infill criteria: the Lower Confidence Bound (LCB), the Probability of Improvement (PI), the Expected Improvement (EI), and the Weighted Expected Improvement (WEI)[55, 44, 2, 283]. All of these approaches will be further discussed in the following. Finally, methods to handle constrained and discrete problems will also be discussed.

5.3.1 Lower Confidence Bound

The most intuitive acquisition function for Bayesian optimization is the Lower Confidence Bound (LCB) criterion[228, 199, 266, 198]. The method was first derived for maximization problems, where it aimed to maximize the Upper Confidence Bound (UCB). For minimization problems, however, the method should instead minimize the Lower Confidence Bound (LCB) given by[195]:

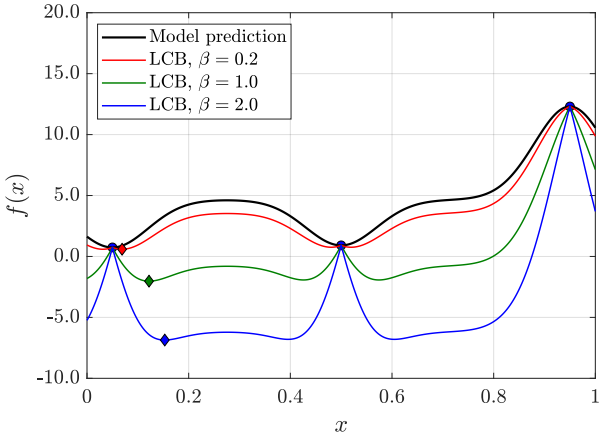
$$\mathrm{LCB}(\mathbf{x}) = \hat{y}(\mathbf{x}) - \beta\,\hat{s}(\mathbf{x}) \tag{5.29}$$

where $\beta \geq 0$ is a parameter which regulates the balance between exploitation and exploration.

This criterion, although simple and straightforward, always selects solutions in the Pareto front of the multi-objective optimization of $\hat{y}(\mathbf{x})$ and $\hat{s}(\mathbf{x})$[224]. However, it is not easy to select an appropriate value for β *a priori*[228, 284], as the relationship between $\hat{y}(\mathbf{x})$ and $\hat{s}(\mathbf{x})$ might depend on the number of sampling points, the true function response, and even the model hyper-parameters. For a given arbitrary function, Figure 34 shows the LCB for different β, and which point would be picked in each one. As β increases, so does the importance of exploration, and the new sampling point is farther from the already sampled data.

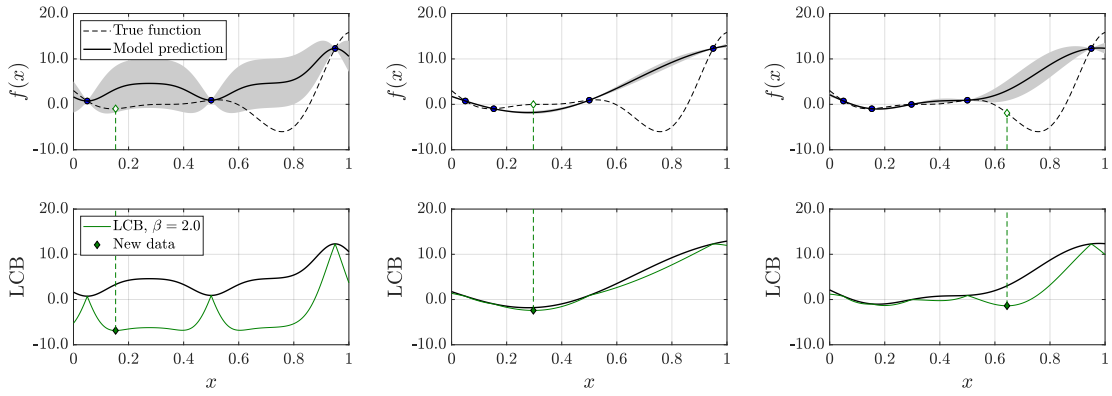
Figure 35 shows how the criterion performs throughout multiple iterations for $\beta = 2.0$, and the change in the model prediction with the addition of new data. Due to how the initial data is arranged, the criterion wastes a few initial iterations exploiting the model. However, once more data is gathered in this region, the criterion starts exploring the design space, finding points very close to the optimum.

Figure 34 – Lower Confidence Bound (LCB) criterion for different β.

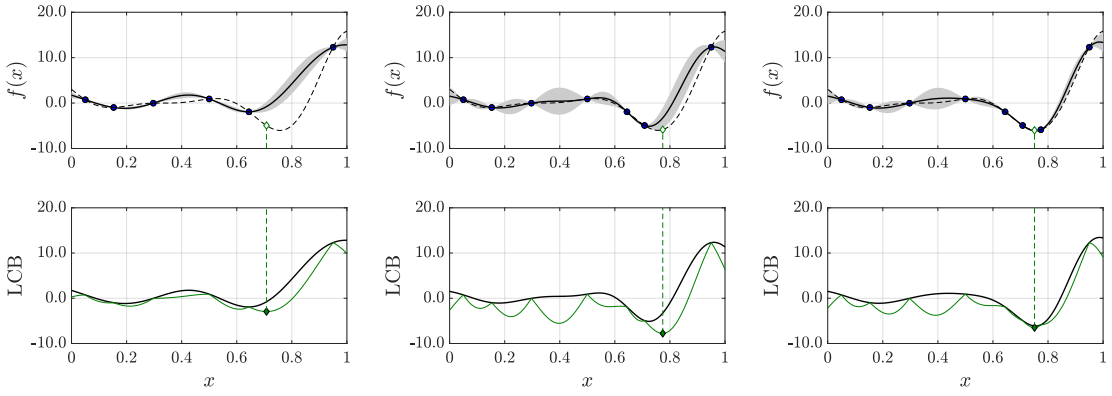


Source: the author

Figure 35 – Lower Confidence Bound (LCB) criterion for multiple iterations with $\beta = 2.0$.



(a) Iteration 1 (b) Iteration 2 (c) Iteration 3



(d) Iteration 4 (e) Iteration 5 (f) Iteration 6

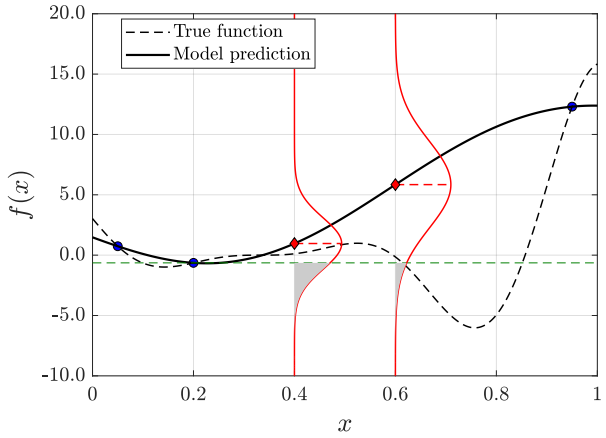
Source: the author

5.3.2 Probability of Improvement

For those with a background in statistics, another seemingly intuitive approach is to select the point where the Probability of Improvement (PI) presents the highest value[40, 199].

This method was first proposed by Kushner[285] and is given by the probability that the addition of an arbitrary point improves the current best point in the sample. Figure 36 illustrates the Probability of Improvement (PI) for two different points, represented by the gray area below the green dashed line, which represents the current best objective function. For $x = 0.4$, the prediction is closer to a data point, and the variance is lower. Since the model itself is predicting that this point has a low $\hat{y}(\mathbf{x})$, it still has a high PI. On the other hand, for $x = 0.6$, the model prediction is not very promising. However, since uncertainty is high in this region, there is still some Probability of Improvement (PI) due to the high variance.

Figure 36 – Illustration of the Probability of Improvement (PI) criterion.



Source: the author

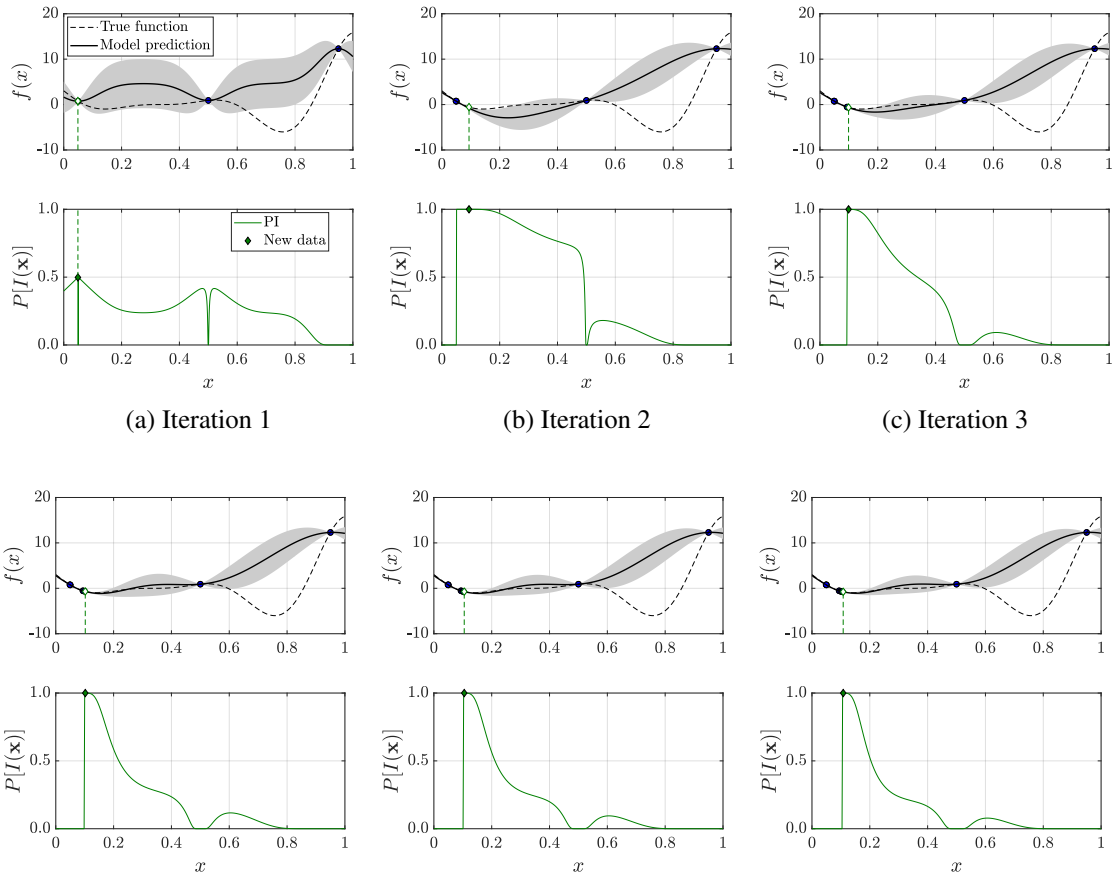
The area behind the curve is given by the Cumulative Distribution Function (CDF) $\Phi(Z)$. For the Normal distribution, this is given by[40]:

$$P[I(\mathbf{x})] = \Phi\left(\frac{y_{min} - \hat{y}(\mathbf{x})}{\hat{s}(\mathbf{x})}\right) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{y_{min} - \hat{y}(\mathbf{x})}{\hat{s}(\mathbf{x})\sqrt{2}}\right)\right] \tag{5.30}$$

Figure 37 shows how the PI criterion selects additional points for subsequent iterations. Note that the initial data set $\mathfrak{D}$ was deliberately chosen due to its very poor initial prediction, and it is the same as the one employed for the demonstration of the LCB criterion. However, the PI behaves much worse.

This occurs because the PI greatly favors exploitation over exploration[199]. Also, unlike other more promising methods, it does not select solutions in the Pareto front of the multi-objective optimization of $\hat{y}(\mathbf{x})$ and $\hat{s}(\mathbf{x})$[224]. Due to these limitations, the PI is not a method employed very often in this context[263]. However, if the user has the prior knowledge of an important threshold for the function y, or if the optimum y_{opt} is known beforehand, the user could replace y_{min} with those values in Eq. (5.30). In these situations, the PI could behave much better.

Figure 37 – Probability of Improvement (PI) criterion for multiple iterations.



(a) Iteration 1 (b) Iteration 2 (c) Iteration 3

(d) Iteration 4 (e) Iteration 5 (f) Iteration 6

Source: the author

By analyzing Figure 37, one may notice that the maximum PI is constant for many x, where the Probability of Improvement is close to 100%. An alternative approach would be to not only consider the probability that the response is improved, but also the magnitude of this improvement so that the single best design is picked. This is a core idea for the next criterion.

5.3.3 Expected Improvement

Based on the theory of expected utility[286], Močkus[287] proposed the Expected Improvement (EI) criterion for the selection of new data points, where the point which maximizes the EI is added to the sample. The method was then popularized by Jones, Schonlau and Welch[53] after the EI was employed on their widely known EGO algorithm, and, nowadays, it is perhaps the most popular infill criteria for Bayesian optimization[224, 283, 37].

The Expected Improvement (EI) of a given point $\mathbf{x}$ is given by[40]:

$$E[I(\mathbf{x})] = \begin{cases} (y_{min} - \hat{y}(\mathbf{x}))\,\Phi\left(\dfrac{y_{min} - \hat{y}(\mathbf{x})}{\hat{s}(\mathbf{x})}\right) + \hat{s}(\mathbf{x})\,\phi\left(\dfrac{y_{min} - \hat{y}(\mathbf{x})}{\hat{s}(\mathbf{x})}\right) & \text{, if } \hat{s}(\mathbf{x}) > 0 \\ 0 & \text{, if } \hat{s}(\mathbf{x}) = 0 \end{cases} \tag{5.31}$$

where $\phi(Z)$ and $\Phi(Z)$ are respectively the Probability Density Function (PDF) and the Cumulative Distribution Function (CDF) for the Normal distribution. The second condition for $\hat{s}(\mathbf{x}) = 0$ is considered just to prevent instabilities from arising when trying to evaluate $\Phi(Z)$ or $\phi(Z)$ on the already sampled data, which would result in a zero variance distribution. It is worth noting that, on usual computers, due to numerical problems, $\hat{s}(\mathbf{x})$ will never be 0, but rather a value very close to it. Thus, it is a good common practice to consider a small tolerance in this computation. For $\hat{s}(\mathbf{x}) > 0$, the EI can also be written as:

$$E[I(\mathbf{x})] = (y_{min} - \hat{y}(\mathbf{x}))\,\frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{y_{min} - \hat{y}(\mathbf{x})}{\hat{s}\sqrt{2}}\right)\right] + \hat{s}\,\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y_{min} - \hat{y}(\mathbf{x})}{\hat{s}}\right)^2\right] \tag{5.32}$$

Figure 38 presents how the criterion behaves for subsequent iterations, using the same initial data set as before. Once again, similar to the LCB criterion, the first couple of iterations are spent by exploiting the region close to the model minimum prediction. However, as the uncertainty on this region lowers, the algorithm starts to add new points on regions with a higher $\hat{s}(\mathbf{x})$. By the end of six cycles, the algorithm has already found the optima region, and will likely find the actual optima in the next few iterations.

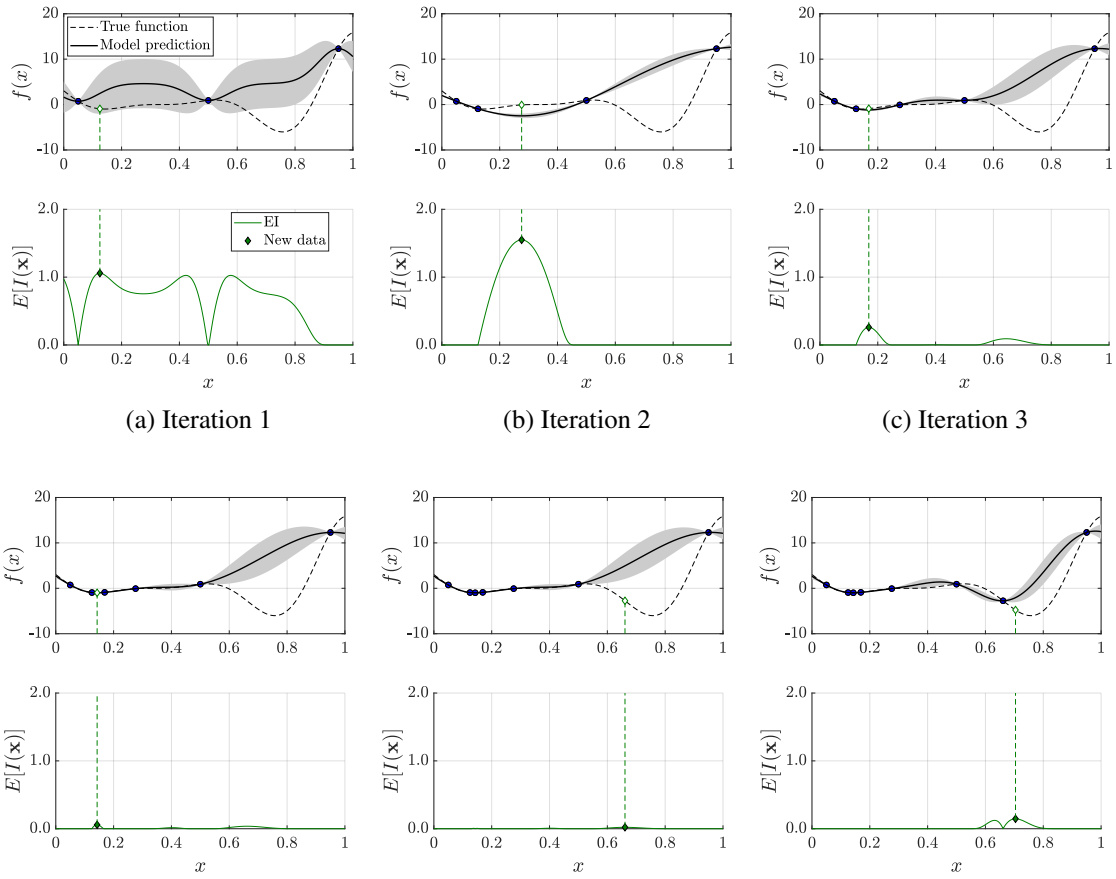
In fact, the EI seems to be a criterion much better the PI. De Ath et al.[224] comment that, just like the LCB, the EI always selects solutions from the multi-objective optimization of $\hat{y}(\mathbf{x})$ and $\hat{s}(\mathbf{x})$. It is worth pointing out that this function can be highly multi-modal and, similar to the PI, it may exhibit vast regions where the objective function is constant ($E[I(\mathbf{x})] = 0$) thus making gradient-based methods ineffective.

5.3.4 Weighted Expected Improvement

The final criterion to be discussed in this work is the Weighted Expected Improvement (WEI), proposed by Sobester, Leary and Keane[55] as an alternative to the EI. In Eq. (5.32), the first term is related to exploitation and the second is related to exploration. Thus, the authors introduced a parameter w which allows the user to control the importance given to either aspect. For $\hat{s}(\mathbf{x}) > 0$, the Weighted Expected Improvement (WEI) is given by:

$$WE[I(\mathbf{x})] = w\,(y_{min} - \hat{y}(\mathbf{x}))\,\frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{y_{min} - \hat{y}(\mathbf{x})}{\hat{s}\sqrt{2}}\right)\right] + (1-w)\,\frac{\hat{s}}{\sqrt{2\pi}}\exp\left[\frac{-(y_{min} - \hat{y}(\mathbf{x}))^2}{2\hat{s}^2}\right]$$

Figure 38 – Expected Improvement (EI) criterion for multiple iterations.



(a) Iteration 1 (b) Iteration 2 (c) Iteration 3

(d) Iteration 4 (e) Iteration 5 (f) Iteration 6

Source: the author

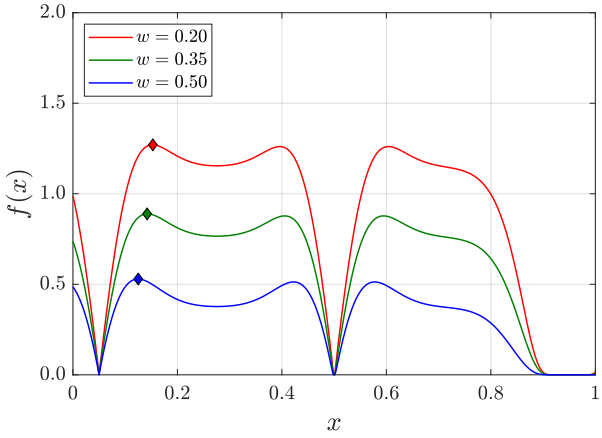
(5.33)

For lower w the algorithm favors exploration, while for higher values the algorithm favors exploitation. De Ath et al.[224] states that, while that is a valid procedure, the WEI only selects solutions from the Pareto optimal front for $w \in [0.185, 0.50]$. Figure 39 shows which points are selected for different w. In this example, while these points are close to each other, the new data gets farther from the sampling points as w lowers. It is worth noting that, for $w = 0.50$, this criterion is equivalent to the EI.

Figure 40 shows how the method behaves for subsequent iterations with $w = 0.20$, which favors the exploration aspect. This time, the optimal region is found much earlier, since a higher priority is given to the exploration.

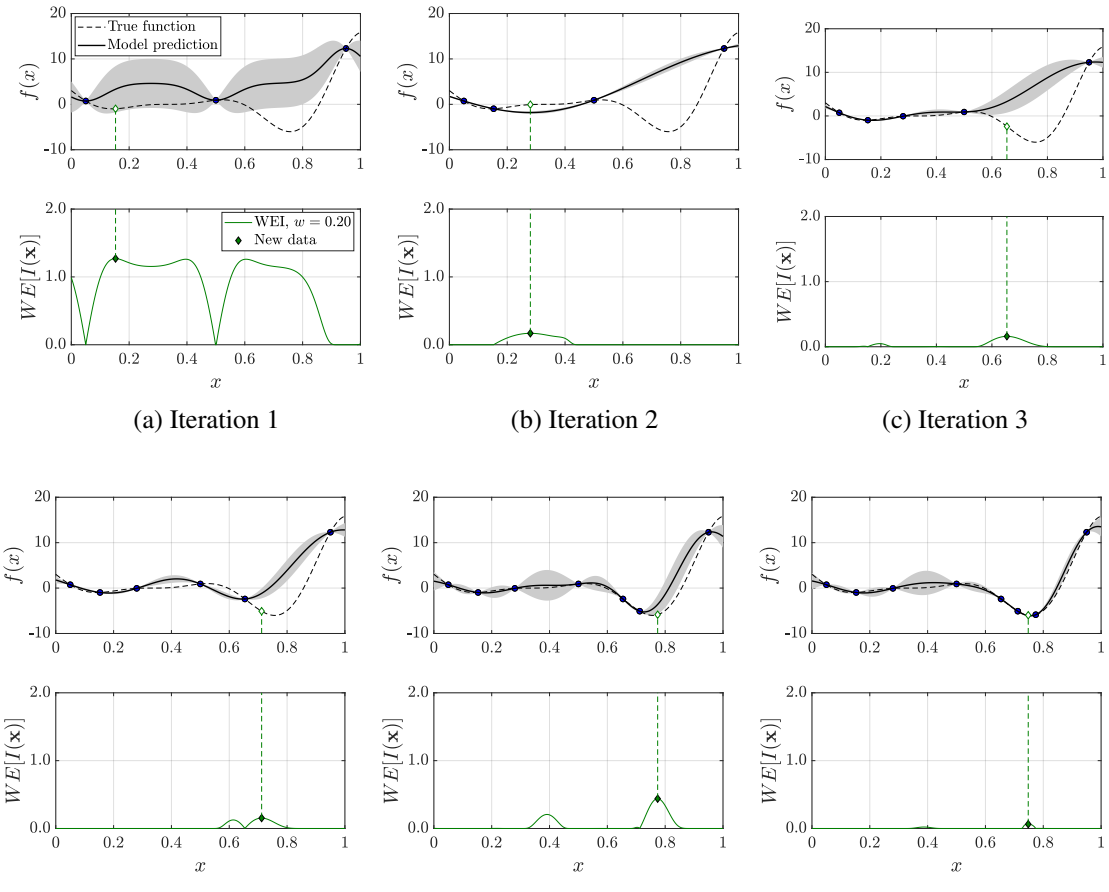
Different w can also be employed in different iterations. Fuhg, Fau and Nackenhorst[202] discuss different approaches of varying this factor, e.g. a decreasing strategy, a greedy strategy (where only extreme values, 0 and 1, are used), and a switching strategy (where values

Figure 39 – Weighted Expected Improvement (WEI) criteria for different w.



Source: the author

Figure 40 – Weighted Expected Improvement (WEI) for multiple iterations with $w = 0.20$.



(a) Iteration 1 (b) Iteration 2 (c) Iteration 3

(d) Iteration 4 (e) Iteration 5 (f) Iteration 6

Source: the author

change smoothly through the iterations). Such an approach can also be used to select different points in the same iteration so that these points can be evaluated in parallel[274]. This practice can increase the robustness and accuracy of the algorithm. A similar approach can also be employed for the LCB, described in Section 5.3.1.

5.3.5 Constraint handling

For constrained optimization problems, the constraint functions should be considered in the definition of the acquisition function. Therefore, an appropriate method should be chosen to handle constrained optimization problems. In some cases, when the feasible region is relatively small, even finding feasible sampling points is not a trivial task. A higher number of initial sampling points could be used to assist in finding feasible designs. Otherwise, a few initial iterations could be devoted to first finding a feasible sampling point[288, 289, 290]. Once a feasible point is found, the algorithm may try to improve upon it[273, 291].

Usually, to ease model building and evaluation, the sampling points for all models (related to the expensive objective and constraint functions) are the same throughout the process. This approach will also be adopted in this work, even though some researchers have been working on different criteria in recent years[284]. For the LCB criterion, constraints can be considered as in bio-inspired algorithms: by the penalization of infeasible designs, where the design's feasibility can be estimated by the constraint approximated model. However, in probabilistic approaches, namely the PI, the EI, and the WEI, it is common to employ appropriate procedures, since usual penalty methods may detract from their probabilistic essence. For instance, it makes no sense to have a negative Probability of Improvement (PI).

First, for these cases, since the improvement should be performed upon the best feasible point $y_{min,feas}$, it should replace y_{min} on the evaluation of the acquisition function[292, 293]. That also means that, ideally, an infeasible design should not be picked as a new sampling point, since those may not improve upon the current optimum design. When the optimization problem has a costly objective function but the constraint functions are cheap to evaluate, one may perform the exact evaluation of the constraint. If the constraint is violated, the expected improvement of a point can be set to 0:

$$CE[I(\mathbf{x})] = \begin{cases} E[I(\mathbf{x})] & \text{, if } g_j(\mathbf{x}) \leq 0 \\ 0 & \text{, if } g_j(\mathbf{x}) > 0 \end{cases} \tag{5.34}$$

This approach often returns very good results. That way, it is rather simple to extend usual SAO to problems where constraints are cheap and, therefore, can be evaluated by their exact function.

However, when the constraints are costly to evaluate, they should also be approximated by a surrogate model to maintain efficiency. This introduces uncertainty to the process, and the task of handling these constraints becomes much more complicated. A variety of methods have been proposed, but there is no silver bullet to handle constrained optimization[294, 58]. It

is not trivial to derive a method that is able to minimize the objective function while satisfying approximated constraints with a reasonable number of evaluations[295].

The most basic approach, denominated in this work as the direct approach, is to evaluate the Expected Improvement (EI) by[55, 290, 268]:

$$CE[I(\mathbf{x})] = \begin{cases} E[I(\mathbf{x})] & \text{, if } \hat{g}_j(\mathbf{x}) \leq 0 \\ 0 & \text{, if } \hat{g}_j(\mathbf{x}) > 0 \end{cases} \tag{5.35}$$

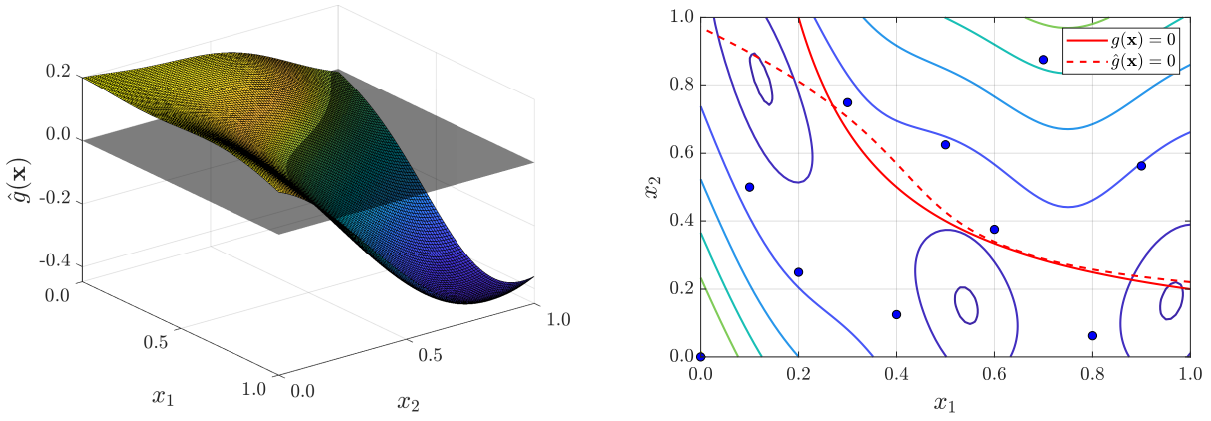
which is very similar to Eq. (5.34), but the exact constraint $g_j(\mathbf{x})$ is replaced here by the approximated constraint $\hat{g}_j(\mathbf{x})$. Other authors proposed similar methods based on the penalization of the acquisition function according to the estimated constraint. Instead of setting the $CE[I(\mathbf{x})]$ to 0 for an infeasible individual, Sasena, Papalambros and Goovaerts[61] suggested the penalization of the EI by a constant penalty factor P. The authors denominate this method as the Penalty-adjusted Expected Improvement (PaEI). Parr et al.[292], on the other hand, penalize the prediction of the model $\hat{y}$ if $\hat{g}_i$ is violated, which in turn affects the EI.

Figure 41 depicts the prediction of an RBF for the exact constraint:

$$g(\mathbf{x}) = 0.2 - x_1 x_2 \leq 0 \tag{5.36}$$

Here, any point below the gray plane is considered to be feasible, according to the predicted surface. However, when we compare it to the exact constraint, we can see that the prediction fails to estimate the real constraint threshold at some points, especially in regions with a lower sampling density. Thus, it is not advisable to rely entirely on the approximation of the constraint.

Figure 41 – Approximate constraint response.



(a) Predicted constraint surface

(b) Comparison with the exact constraint

Source: the author

In fact, it is important to consider the uncertainty in the process[37]. This way, Schonlau[296] proposed that one could use the Probability of Feasibility (PF) to penalize the EI at

a given point **x**. This approach is often denominated as the Constrained Expected Improvement (CEI)[40, 37]. The assessment of the PF resembles the evaluation of the PI and is given by:

$$P_j[F(\mathbf{x})] = \Phi\left(-\frac{\hat{g}_j(\mathbf{x})}{\hat{s}(\mathbf{x})}\right) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{-\hat{g}_j(\mathbf{x})}{\hat{s}(\mathbf{x})\sqrt{2}}\right)\right] \tag{5.37}$$

Thus, considering that all expensive constraints are independent from each other, the CEI is given by:

$$CE[I(\mathbf{x})] = E[I(\mathbf{x})] \prod_{j=1}^{n_c} P_j[F(\mathbf{x})] \tag{5.38}$$

where n_c is the number of expensive constraints. This approach is a good baseline for the constraint-handling problem, since it considers the uncertainty of the process[296, 60, 40, 291]. Thus, if a region has a lower sampling density, the method knows that the prediction $\hat{g}(\mathbf{x})$ is less reliable.

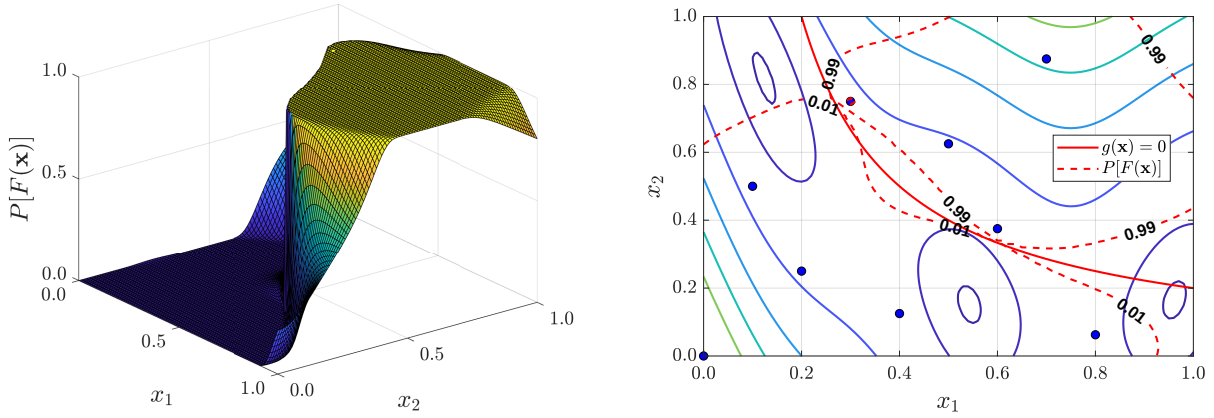
Basudhar et al.[295] suggest that the PF may also be assessed using a classification model, such as the Support Vector Machine (SVM). The idea behind these techniques is the same, and their result seems to be fairly close. The use of a classification model is more adequate for problems where the number of expensive constraint functions is very high since there is no need to build n_c models. That being said, Sohst, Afonso and Suleman[297] follow the same idea of building only one model for all expensive constraint functions but use a Ordinary Kriging (OK) model. While this is not adequate, as Kriging is more suitable for approximating smooth functions, the authors show that the results are also fairly accurate for low-dimensional problems.

Figure 42 shows how the function $P[F(\mathbf{x})]$ behaves. In the contour plot, it is shown the contours for some values of $P[F(\mathbf{x})]$. In regions closer to the sampling points, the contour lines are closer to each other, and the PF goes very quickly from 0% to 100%. This means that there is more certainty about these boundaries. However, for low sampling density regions, $P[F(\mathbf{x})]$ changes more slowly throughout the design space.

One may note that the PF actually penalizes a large region located in the feasible space. This is especially important because, for engineering problems, the global optimum is often located in the boundary of the feasible space[292]. Thus, the PF may take many iterations to locate this point since it penalizes designs that are actually closer to the true feasibility threshold, especially in an unexplored region. Parr et al.[292] suggest that further improvement of the PF could be made by directly searching for solutions laying on constraint boundaries.

To overcome this issue, some researchers presented modified formulations, where a feasibility function $F(\mathbf{x})$ is employed. These are considered just as the PF, and penalize the EI

Figure 42 – Behavior of the Probability of Feasibility (PF) function.



(a) PF surface

(b) Comparison with the exact constraint

Source: the author

by:

$$CE[I(\mathbf{x})] = E[I(\mathbf{x})] \prod_{j=1}^{n_c} F_j(\mathbf{x}) \tag{5.39}$$

Examples of different feasibility functions can be found in the literature. Tutum, Deb and Baran[62] proposed the consideration of:

$$F_j^{(T)}(\mathbf{x}) = \begin{cases} 2 - \text{erf}\left(-\dfrac{\hat{g}(\mathbf{x})}{\hat{s}(\mathbf{x})}\right) & \text{, if } 0 < \text{erf}\left(-\dfrac{\hat{g}(\mathbf{x})}{\hat{s}(\mathbf{x})}\right) \leq 1 \\ 0 & \text{, otherwise} \end{cases} \tag{5.40}$$

A very similar function is also employed by Wang, Hu and Li[289]. On the other hand, Bagheri et al.[63] use:

$$F_j^{(B)}(\mathbf{x}) = \min\left(2\,\Phi\left(-\frac{\hat{g}(\mathbf{x})}{\hat{s}(\mathbf{x})}\right), 1\right) \tag{5.41}$$

Finally, Sohst, Afonso and Suleman[297] propose that a sine function can be used to handle the uncertainty more generously, as in:

$$F_j^{(S)}(\mathbf{x}) = \sin\left(P_j[F(x)] \times \frac{\pi}{2}\right)^n \tag{5.42}$$

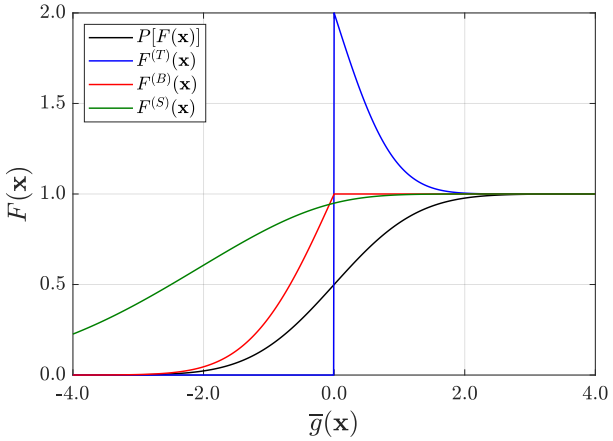
The authors state that the best results were found for $n = 0.15$.

Figure 43 shows how these different feasibility functions behave for different $\overline{g}(\mathbf{x})$, where:

$$\overline{g}(\mathbf{x}) = -\frac{\hat{g}(\mathbf{x})}{\hat{s}(\mathbf{x})} \tag{5.43}$$

All of these approaches favor points closer to the feasibility boundary, in comparison with the method proposed by Schonlau[296]. Also, one may argue that completely penalizing infeasible designs may limit the update of the threshold approximation[288], as does Tutum, Deb and Baran[62] as soon as $\hat{g} < 0.0$. The approach proposed by Sohst, Afonso and Suleman[297], however, applies a very smooth penalty for unfeasible designs. Further research is still needed to define the most appropriate constraint handling method[37]. In this work, these different constraint handling approaches will be compared.

Figure 43 – Different feasibility functions.



Source: the author

Furthermore, a couple of other methods can be employed to handle constraints in SAO. Sasena, Papalambros and Goovaerts[61] proposed the SuperEGO algorithm, which applies a penalty factor when the PF in a given point is lower than some fixed threshold τ, which is defined as 0.95 by default. Other constraint handling methods can be derived based on augmented Lagrange multipliers[298, 299], or even by performing the multi-objective optimization of the EI and the PF[300]. Also, note that these probabilistic criteria can also be used for the LCB, even though it might be necessary to adopt adequate modifications[297]. These approaches will not receive a major focus in this work.

In engineering optimization, sometimes, the user desires to optimize a cheap objective function (e.g. the mass or cost of a structure) while respecting a couple of costly constraints (e.g. displacement, critical load, fundamental frequency). To handle these problems, one may replace the Expected Improvement (EI) for the actual improvement[118]:

$$I(\mathbf{x}) = \max(y_{min} - y(\mathbf{x}), 0) \tag{5.44}$$

Then, the actual improvement may be penalized by the feasibility functions in the same way as the EI, as in Eqs. (5.38) and (5.39).

5.3.6 Handling of discrete variables

Usually, the surrogate modeling techniques described in this work are derived considering that variables are continuous in the design space. However, on engineering problems, this is often not the case. For example, optimization of laminate structures is often performed using discrete design variables[5]. Appropriate methods need to be employed so that surrogate models can be used in discrete problems[58, 37, 301].

On Surrogate-Based Optimization (SBO), the discrete nature of the data should be taken into account during the initial sampling (see Section 5.1.3) and in the infilling phase, as only feasible data points should be added. Stork et al.[58] comment that this aspect is still an open issue, and they describe different approaches that may be employed in these cases. First, the basic approach is to simply round the result of a continuous optimization problem to the nearest discrete value[37]. The authors denominate this as the naive approach, and, even though it is very straightforward, it may not find the true optima for the discrete optimization problem.

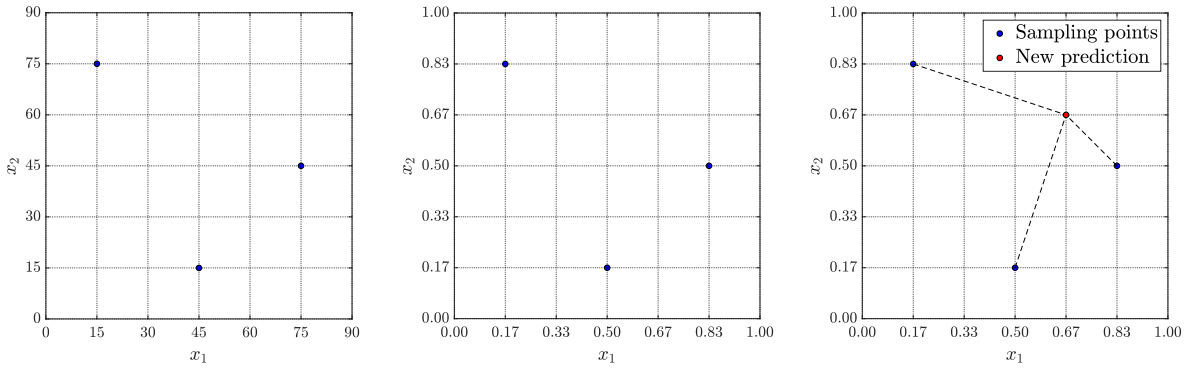
On the other hand, mapping approaches may solve this issue. By the use of dummy variables or auto-encoders, the optimization process and the definition of sampling points may consider the discrete nature of the data[302]. In that matter, the special encoding presented in Section 4.2.1 may assist in the use of such mapping approaches.

Figure 44 shows how a discrete space can be normalized. For RBF and Kriging models, predicted response is a function of the distance between sampling and prediction points. Here, distances can be evaluated using, for instance, the Euclidean distance. Thus, it is very straightforward to generalize RBF and Kriging to dealing with discrete spaces. For combinatorial problems, Zaefferer et al.[234] suggest that an isotropic Gaussian correlation function may be used to evaluate the correlation between two data points. The authors also evaluate the correlation between binary strings using appropriate distance functions, such as the Hamming distance, instead of the usual Euclidean distance.

Furthermore, discrete variables can also be handled by using inherently discrete models, such as regression trees or random forests; custom-fit solutions[303], which only work for specific problems; and extraction of its numeric features[304]. However, for the optimization of composite structures, these approaches do not seem as robust or promising as the other ones described.

Finally, some authors have proposed different algorithms to deal with these kinds of problem. On the topic of mixed-variable optimization, Müller, Shoemaker and Piché[305]

Figure 44 – Normalization of the discrete space for surrogate modeling.



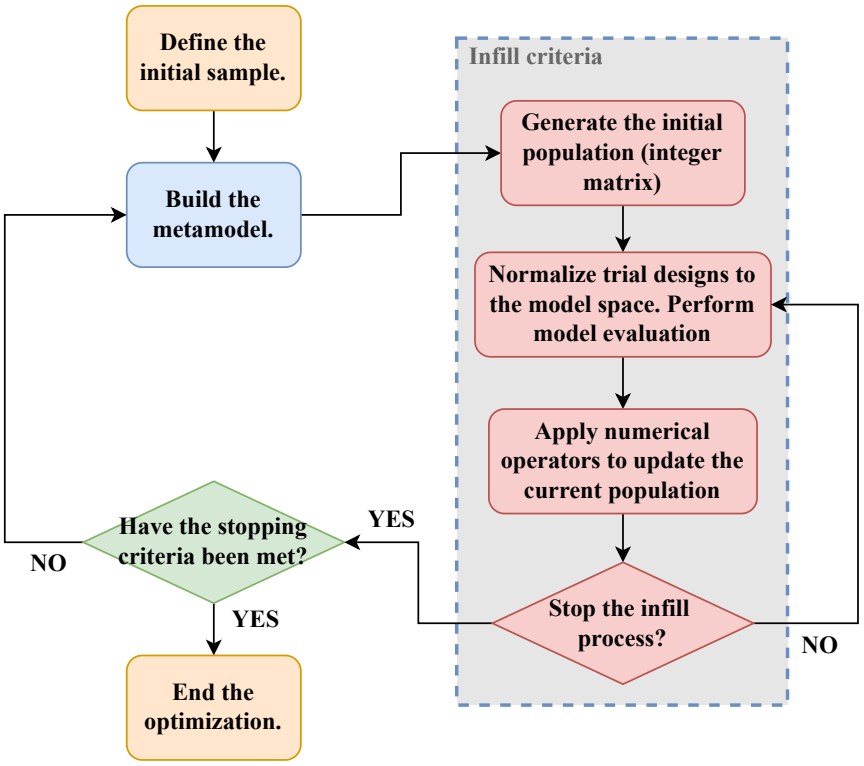
(a) Fenotypical representation (b) Normalization (c) Assessment of distances

Source: the author

presented the Surrogate Optimization-Mixed Integer (SO-MI), a computational algorithm for expensive mixed-integer optimization problems. In this algorithm, the initial sample is evaluated using a LHS where points are rounded to the nearest integer, as described in Section 5.1.3, and a RBF model is fitted to these points. Infill points are then chosen according to an exploitation criterion, where candidate solutions are derived from the perturbation of the current best design. The nature of each design variable is taken into account when defining these candidate solutions. Also, it is important to remember that finding the best design requires the use of an appropriate optimization algorithm, such as the Genetic Algorithm (GA)[305]. A similar approach can be found in the MATSuMoTo toolbox for MATLAB, which was developed by the same authors[306].

Later, the same authors proposed an improvement over the SO-MI algorithm, now under the name of Mixed Integer Surrogate Optimization (MISO)[307]. MISO allows for the use of different infill criteria, such as the EI and the bumpiness function. Once again, similar to a mapping approach, candidate solutions are always drawn from proper discrete designs. In this work, a similar approach will be employed, where candidate solutions are drawn from the discrete space and different acquisition functions can be employed. Figure 45 demonstrates how the infilling phase for discrete laminate problems will be handled. We will take advantage of the special encoding for laminates described in Section 4.2.1, where designs are defined by an integer matrix. Each row is related to a design variable (thickness, orientation, or material) and each column to a specific ply. To simplify model implementation, the integer matrix is later transformed into a normalized vector and, then, model evaluation is performed.

Figure 45 – Handling of discrete variables during the infilling process.



Source: the author

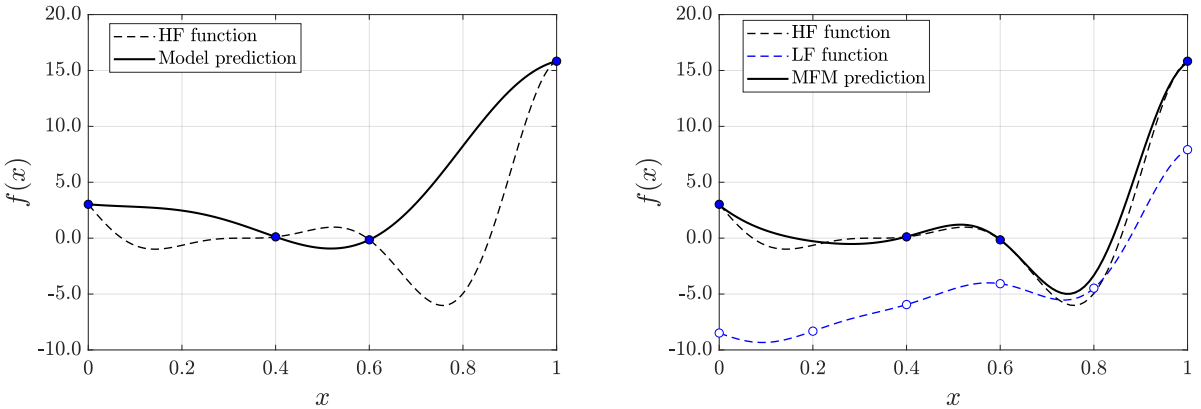
6 MULTI-FIDELITY MODELS

In engineering design, multiple models of a physical phenomenon are sometimes available with different levels of approximation. Therefore, the user is able to perform Surrogate-Based Optimization (SBO) using either lower or higher fidelity sources, and the choice for one or the other should take the accuracy and computational cost of each source into consideration. The building of a surrogate using a LF source may incur in an accumulation of errors[69], while the computational time associated with the evaluation of the sampling points can be prohibitive for a surrogate built considering a High-Fidelity (HF) source, depending on the number of sampling points required[38].

Multi-Fidelity Models (MFMs) allow for a promising way of performing the trade-off between accuracy and efficiency for SBO. By using information from both HF and LF simulations in the surrogate model building, MFMs may reduce the number of HF evaluations needed to achieve the required accuracy[308, 309]. In the last few years, some effort has also been made to further extend these to deal with multi-level fidelity surrogates, where more than two levels of fidelity are explored[259, 310, 256, 311].

Figure 46 presents the prediction of a surrogate using a regular single-fidelity model and a Multi-Fidelity Model (MFM). The consideration of information from a LF source allows for a much better exploration of the design space since more sampling points can be evaluated roughly under the same budget[312]. Thus, MFMs have the potential to reduce computational effort, while also enhancing the model accuracy[313, 52, 312].

Figure 46 – Improvement of the model accuracy by the use of a MFM.



(a) Ordinary Kriging (OK) model prediction

(b) Multi-Fidelity Model (MFM) prediction

Source: the author

To use a MFM, different levels of approximation for a particular problem should be

available. Cheng, Lu and Zhen[311] point out five main approaches that can be used to derive a LF source: simplification of a physics model, linearization, dimensionality reduction, coarser discretization, and partially converged results.

Simplification of a physics model is a very popular method in aerodynamics, where the Navier-Stokes equations can be easily replaced by the Euler equations, which disregards viscosity[70, 314]. Another example is in the use of the Kreuzer model instead of the full-linearized Navier-Stokes in photoacoustic sensor tests[309]. Similarly, in solid mechanics, 3D models could be replaced by simpler 2D models based on plate or shell theories or even 1D models based on beam theories. If the correlation between physics models is high, the approach is usually able to provide very good results. Even so, not all areas have such possibilities available. Dimensionality reduction is a more general approach that can be easily employed in optimization problems[315]. For instance, Xiao et al.[310] use Proper Orthogonal Decomposition (POD) to derive a LF source, achieving great results in the optimization of an airfoil. Nachar et al.[316], on the other hand, use Proper Generalized Decomposition (PGD) to provide a simpler way of analyzing viscoplastic structures.

For engineering problems where analysis is performed using mesh-dependent methods (e.g. FEM and IGA), the definition of a LF source by the use of a coarser discretization is very straightforward. Cai et al.[313] performed the optimization of a long cylinder pressure vessel using 3D models on ANSYS. For the HF model, a 20 mm grid mesh is employed, while a 40 mm grid mesh is used in the LF model. Guo et al.[78] performed the optimization of Variable Stiffness Composites (VSCs) using 2D models. For the assessment of the buckling load of a square plate, the HF model is built using 10,000 elements, while the LF has only 625 elements. A similar procedure is later employed by the same authors[79] in the optimization of circular and elliptical cylinders, where the HF model has 5 to 6 times the number of elements from the LF source. Rohit and Ganguli[317] performed an optimization procedure using coarse (5 elements) and refined (70 elements) meshes for uncertainty quantification of beam vibration. The authors show that the model is able to provide an average accuracy higher than 99.99% using only 2 HF and 10 LF sampling points.

Wang et al.[318] tried to minimize strains on the boom of a bucket wheel reclaimer. The authors employ an interesting approach, where different physics models are considered alongside coarser meshes. That way, the HF source was built using 374,000 Timoshenko beam elements, where analyses take about 71 s, while the LF source presented 46,750 Euler-Bernoulli

beam elements, where analyses take approximately 13 s. Similar approaches are also employed in Computational Fluid Dynamic (CFD) analyses, where mesh discretization is employed for further simplification after the Navier-Stokes are replaced by the Euler equations. Ruan et al.[74] employed a 2D model to perform the aircraft drag coefficient prediction using nearly 3 million elements. The LF model was then derived using a coarser mesh with 566,281 elements. Likewise, Song et al.[52] used a 3D model with 284,412 nodes as the HF source, while the LF source was derived by the use of a 2D model with only 9,646 nodes. The same models were also later employed by Shi et al.[45].

Finally, the LF model can also be derived by the relaxation of a numerical convergence threshold, thus resulting in partially converged results. This approach is employed by Bertram and Zimmermann[314], where the Navier-Stokes equations are solved using a model with 21,821 nodes. For the HF source, 5000 solver iterations are performed, which is able to achieve a convergence level of 10^{-6}. The LF source, however, performs only 150 iterations, which is able to achieve residuals in the range $[10^{-3}, 10^{-1}]$.

It is important to note that, if the HF and LF sources are not well correlated, MFMs may produce less accurate results. Thus, these models are not a panacea for the improvement of the model prediction[74]. Toal[319] suggests that the correlation between sources should be reasonably high, presenting an $r^2 > 0.9$. Otherwise, a single-fidelity model built with the same equivalent budget, but based solely on expensive HF simulations, may still perform better.

After defining HF and LF sources, appropriate sampling points should be defined from each source. Similar to single-fidelity models, it is not easy to define an optimal sample size beforehand. Also, data can be nested, where the location of the HF data points is the same as some of the LF data points, or non-nested, where the sampling spaces can be completely different. Nested data allows for an easier estimate of the model parameters separately, which may improve computational efficiency for larger data sets[76]. Model building under a non-nested sample is still an open issue, but most approaches try to perform an approximation of the LF response on the HF sampling locations. Thus, a non-nested sample may incur a better exploration of the design space under the same budget. In this work, only nested data will be studied.

For MFMs, the optimal ratio between HF and LF sampling points is also unknown[308]. Toal[319] suggests that the number of cheap and expensive evaluations should be related to the computational cost of each source. As a conservative approach, the author proposes that:

$$f_r > \frac{1.75}{1+1/C_r}, \text{ where } 0.10 < f_r < 0.80 \tag{6.1}$$

where f_r is the percentage of the total evaluation budget converted to cheap evaluations, and C_r is the relative expense, given by:

$$C_r = \frac{C_c}{C_e} \tag{6.2}$$

where C_c is the cost of the cheap function evaluation and C_e is the cost of the expensive function evaluation. There should always be slightly more cheap points than expensive ones. On that matter, Song et al.[52] also suggest that C_r should be lower than 0.2. Usually, analysis cost is measured by the computational time spent. One should note that the number of LF sampling points should not be extremely high, as this would make model building and model evaluation costs prohibitive.

After defining the HF and LF sources and assessing appropriate sampling points, MFMs can be built using three different approaches. The simplest one is related to the use of a scaling function to adjust the LF response. The function employed is usually a low-order polynomial[318], but more complex models, e.g. SVRs, can also be employed[43]. When input parameters are different in the HF and LF sources, one may build a MFM by using space mapping approaches[320]. These are straightforward when the LF response is defined via dimensionality reduction methods. The third method is the use of multi-fidelity Bayesian models[308, 74], which is the main focus of this work.

Multi-fidelity Bayesian models were first introduced by Kennedy and O'Hagan[71] with the use of the KOH autoregressive model, which is known today as Cooperative Kriging (Co-Kriging). The Co-Kriging is the most widely used MFM[74], but recent works have been presented introducing novel methods. For instance, Co-Kriging inspired the development of a Cooperative Support Vector Regression (Co-SVR)[45], Cooperative Radial Basis Functions (Co-RBF)[309, 52], and even Linear Regression based methods have been proposed[321]. In addition, other MFMs derived from the Kriging model were proposed, such as the Multi-Fidelity Gaussian Process (MFGP)[256, 315], the Multi-Fidelity High Dimensional Model Representation (MF-HDMR)[313], the Evofusion[322, 316], and the Hierarchical Kriging (HK)[72].

In most cases, static MFMs are employed for SBO, where the surrogate is not updated during the optimization process. In structural optimization problems, these have been employed for the optimization of engine models using Co-Kriging[319, 69] and the optimization of long cylinder pressure vessels using the Multi-Fidelity High Dimensional Model Representation (MF-HDMR)[313]. In Computational Fluid Dynamic (CFD) problems, different static MFMs

have been employed in the optimization of lift and drag coefficients or in the analysis of pressure relief valves, such as the Co-Kriging[314, 74], the Co-RBF[70, 52], the Co-SVR[45], and the HK[72].

As already stated in the previous sections, the use of adaptive sampling techniques is very interesting for optimization problems, since accuracy in the global optimum region can be iteratively improved. Some MFMs also allow for the use of error-based exploration, similar to single-fidelity GP models. Here, however, along with the definition of the new sampling point location, its fidelity should also be chosen, since there are two possible sources for the infill point[76]. In addition, very few papers deal with approximate constraint functions using MFMs, although the extension of single-fidelity constraint-handling approaches can be performed with relative ease [323].

Zhou et al.[43] performed the optimization of a long cylinder pressure vessel using an SVR-based scaling function. The authors try to balance out exploitation and exploration to select a new infill point. All new points are evaluated by the HF source. Nachar et al.[316] used a Kriging-based MFM, the Evofusion, in the maximization of the von Mises stress for viscoplastic structures. The authors use a variant of the EI as the method to select new sampling points. However, the authors claim that the method could be improved by a better choice of acquisition functions, and a comparison between different acquisition functions could be performed. Durantin et al.[309] employed Co-RBF for the optimization of the geometry of a gas sensor. The selection of new points is performed by the maximization of the bumpiness function. Again, all new points are evaluated using the HF source.

Huang et al.[324] proposed the Multiple Fidelity Sequential Kriging Optimization (MFSKO) algorithm, which employed the Co-Kriging along with an acquisition function denominated as Augmented Expected Improvement (AEI) to perform the sequential sampling. This function is derived from the EI by considering both the relative expense and the correlation between sources to select the new point's fidelity. The authors then apply the method for design improvement on a metal-forming process by minimizing the die wear. The authors derive a LF source by replacing a 27,000-element mesh for a 10,000-element mesh.

Regarding CFD analysis, Zhou et al.[308] employed a Kriging-based MFM to estimate the lift coefficient of an airfoil. The new sampling points are selected by the maximization of the mean squared error. The authors then evaluate the so-called predicted improvement level for the HF and LF sources, and the source with the highest level will be used for the evaluation of the new sampling point.

Zhang, Han and Zhang[73] presented a similar optimization problem, but the model was built using Hierarchical Kriging (HK). The HK model allows for a nice way for estimating the prediction variance for the HF and LF sources separately, which allow for the evaluation of the EI from both sources. Then, by the maximization of the EI, the authors select both the new sample location and its fidelity. This approach is denominated as VF-EI. Guo et al.[78] and Guo et al.[79] also employed the VF-EI to the optimization of Variable Stiffness Composites (VSCs).

Later, Ruan et al.[76] proposed a similar approach, denominated as Variable Fidelity Probability of Improvement (VF-PI), where, instead of using the EI to select new data points, the PI is employed. The authors apply the approach to the optimization of a micro-aerial vehicle fuselage and state that VF-PI may perform better than the VF-EI if there is a strong correlation between the sources.

In this work, MFMs will be employed along with adaptive sampling for the optimization of composite structures. The focus will be given to Cooperative Kriging (Co-Kriging) and, due to its interesting approaches to tackle SAO, to Hierarchical Kriging (HK). Similar to the single-fidelity models, the addition of new points will be performed using error-based exploration. The following sections will further describe each model and discuss how error-based exploration can be performed in these cases.

6.1 Cooperative Kriging

The Cooperative Kriging (Co-Kriging) was the first Bayesian Multi-Fidelity Model (MFM) employed, proposed by Kennedy and O'Hagan[71]. Co-Kriging is based on autoregressive models[325], which assume that the expensive response is the true one. Thus, you can not learn from the LF source if the HF evaluation at the same point is already known. This is known as the Markov assumption[40, 259].

Due to the Markov assumption, one may think that using nested data would be wasteful. However, the use of nested sampling space means that the correlation between sources can be drawn much more easily, making model building more simple. General recommendations follow that the use of a nested data is advisable for the Co-Kriging model building, although not necessary[40, 74].

Considering the Markov assumption, the MFM prediction is given by[74]:

$$\hat{y}_h = \rho\,\hat{y}_l + \hat{y}_d \tag{6.3}$$

where ρ is a scaling factor, $\hat{y}_l$ is a model based on the LF sample, and $\hat{y}_d$ is a model which represents the difference between LF and HF sources, which will be denominated here as the difference model. The Markov assumption allows different levels of fidelity to present different correlation structures and, thus, model hyper-parameters can be set independently.

Co-Kriging is able to correlate multiple sets of data, derived from different fidelity sources. Thus, the sample can be defined by input-output pairs $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}$, where:

$$\begin{aligned}
\mathbf{x} &= \begin{bmatrix} \mathbf{x}_l & \mathbf{x}_h \end{bmatrix}^T = \begin{bmatrix} \mathbf{x}_{l,1} & \mathbf{x}_{l,2} & \dots & \mathbf{x}_{l,n_l} & \mathbf{x}_{h,1} & \mathbf{x}_{h,2} & \dots & \mathbf{x}_{h,n_h} \end{bmatrix}^T \\
\mathbf{y} &= \begin{bmatrix} \mathbf{y}_l(\mathbf{x}_l) & \mathbf{y}_h(\mathbf{x}_h) \end{bmatrix}^T = \begin{bmatrix} y_l(\mathbf{x}_{l,1}) & \dots & y_l(\mathbf{x}_{l,n_l}) & y_h(\mathbf{x}_{h,1}) & \dots & y_h(\mathbf{x}_{h,n_h}) \end{bmatrix}^T
\end{aligned} \tag{6.4}$$

where subscripts h and l correspond to the HF and LF samples, respectively. For nested samples, $\mathbf{x}_h \subset \mathbf{x}_l$.

The local features of the cheap and expensive responses can be represented by the Gaussian Processes Z_l and Z_h. For an autoregressive model:

$$Z_h(\mathbf{x}) = \rho\, Z_l(\mathbf{x}) + Z_d(\mathbf{x}) \tag{6.5}$$

where Z_d represents a GP model for the difference between HF and LF responses. In regular Kriging, the covariance between data can be evaluated by Eq. (5.19). For the Co-Kriging, regarding the different sources of data, covariance matrices are given by[40]:

$$\begin{aligned}
\mathrm{cov}(\mathbf{y}_l, \mathbf{y}_l) &= \sigma_l^2\, \boldsymbol{\Psi}_l(\mathbf{x}_l, \mathbf{x}_l) \\
\mathrm{cov}(\mathbf{y}_l, \mathbf{y}_h) &= \rho\, \sigma_l^2\, \boldsymbol{\Psi}_l(\mathbf{x}_l, \mathbf{x}_h) \\
\mathrm{cov}(\mathbf{y}_h, \mathbf{y}_h) &= \rho^2\, \sigma_l^2\, \boldsymbol{\Psi}_l(\mathbf{x}_h, \mathbf{x}_h) + \sigma_d^2\, \boldsymbol{\Psi}_d(\mathbf{x}_h, \mathbf{x}_h)
\end{aligned} \tag{6.6}$$

and the Co-Kriging complete covariance matrix is then given by:

$$\mathbf{C} = \begin{bmatrix} \sigma_l^2\, \boldsymbol{\Psi}_l(\mathbf{x}_l, \mathbf{x}_l) & \rho\, \sigma_l^2\, \boldsymbol{\Psi}_l(\mathbf{x}_l, \mathbf{x}_h) \\ \rho\, \sigma_l^2\, \boldsymbol{\Psi}_l(\mathbf{x}_l, \mathbf{x}_h) & \rho^2\, \sigma_l^2\, \boldsymbol{\Psi}_l(\mathbf{x}_h, \mathbf{x}_h) + \sigma_d^2\, \boldsymbol{\Psi}_d(\mathbf{x}_h, \mathbf{x}_h) \end{bmatrix} \tag{6.7}$$

where $\boldsymbol{\Psi}_l$ and $\boldsymbol{\Psi}_d$ are the correlation matrices, which depend on the correlation function ψ. Similar to the single-fidelity Kriging, correlation functions should be able to guarantee that correlation matrices are positive definite[314]. For the Gaussian kernel:

$$\begin{aligned}
\Psi_l(\mathbf{x}_i, \mathbf{x}_j) &= \exp\left(-\sum_{k=1}^{m} \theta_{k,l}\, |x_{i,k} - x_{j,k}|^{p_{k,l}} \right) \\
\Psi_d(\mathbf{x}_i, \mathbf{x}_j) &= \exp\left(-\sum_{k=1}^{m} \theta_{k,d}\, |x_{i,k} - x_{j,k}|^{p_{k,d}} \right)
\end{aligned} \tag{6.8}$$

Here, there are two sets of parameters to be adjusted, one for each GP. Due to the Markov assumption, the building of these models can be performed separately[76]. The LF model is fitted just as the Ordinary Kriging (OK) model, presented in Section 5.2.2. Thus, its hyper-parameters can be defined by maximizing the ln-likelihood function:

$$\ln L_l = -\frac{n_l}{2}\ln(2\pi) - \frac{n_l}{2}\ln(\sigma_l^2) - \frac{1}{2}\ln|\boldsymbol{\Psi}_l(\mathbf{x}_l,\mathbf{x}_l)| - \frac{(\mathbf{y}_l - \mathbf{1}\,\mu_l)^T\,\boldsymbol{\Psi}_l(\mathbf{x}_l,\mathbf{x}_l)^{-1}\,(\mathbf{y}_l - \mathbf{1}\,\mu_l)}{2\,\sigma_l^2} \quad (6.9)$$

where the MLEs for μ_l and σ_l are given by:

$$\begin{aligned}\hat{\mu}_l &= \frac{\mathbf{1}^T\boldsymbol{\Psi}_l(\mathbf{x}_l,\mathbf{x}_l)^{-1}\mathbf{y}}{\mathbf{1}^T\boldsymbol{\Psi}_l(\mathbf{x}_l,\mathbf{x}_l)^{-1}\mathbf{1}} \\ \hat{\sigma}_l^2 &= \frac{(\mathbf{y}-\mathbf{1}\hat{\mu}_l)^T\boldsymbol{\Psi}_l(\mathbf{x}_l,\mathbf{x}_l)^{-1}(\mathbf{y}-\mathbf{1}\hat{\mu}_l)}{n_l}\end{aligned} \quad (6.10)$$

Finally, substituting these estimators in Eq. (6.9) and removing the constant terms we end up with the cheap model concentrated ln-likelihood function:

$$\ln L_l \approx -\frac{n_l}{2}\ln(\hat{\sigma}_l^2) - \frac{1}{2}\ln|\boldsymbol{\Psi}_l(\mathbf{x}_l,\mathbf{x}_l)| \quad (6.11)$$

Just as in the single-fidelity Kriging model, $p_{k,l}$ can be set to 2.0 to make the optimization process easier. In this case, the assessment of the MLE for $\theta_{k,l}$ is an unconstrained optimization problem defined by:

$$\begin{cases} \text{find} & \boldsymbol{\theta}_l = \{\theta_{1,l}, \theta_{2,l}, \ldots, \theta_{m,l}\} \\ \text{that minimizes} & -\ln L_l(\boldsymbol{\theta}_l) \\ \text{with} & \boldsymbol{\theta}_L \leq \boldsymbol{\theta}_l \leq \boldsymbol{\theta}_U \end{cases} \quad (6.12)$$

Once the LF model is built, the goal now is to build the difference model. From the sampling point responses, the difference vector can be defined by:

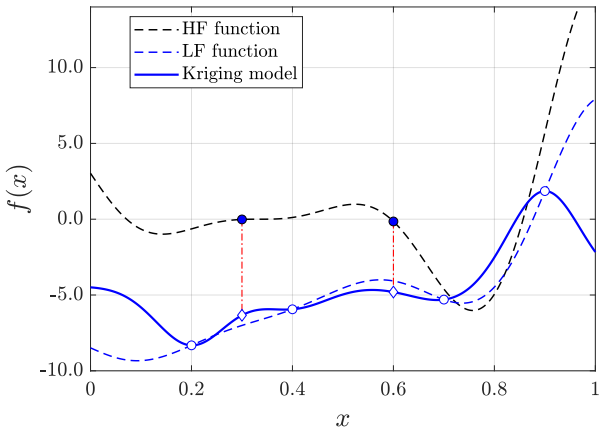
$$\mathbf{d}(\mathbf{x}_h) = \mathbf{y}_h(\mathbf{x}_h) - \rho\,\mathbf{y}_l(\mathbf{x}_h) \quad (6.13)$$

where ρ is a scaling factor, which is a hyper-parameter for the model. If data is nested, this evaluation is straightforward. If not, we may estimate $\mathbf{y}_l(\mathbf{x}_h)$ using the already fitted cheap model[40]:

$$y_l(\mathbf{x}) \approx \hat{y}_l(\mathbf{x}) = \hat{\mu}_l + \boldsymbol{\psi}_l^T\,\boldsymbol{\Psi}_l^{-1}\,(\mathbf{y}_l - \mathbf{1}\,\hat{\mu}_l) \quad (6.14)$$

Figure 47 shows how the estimate for y_l may be performed using the LF model, if data is not nested. Note how, in this case, the Kriging model is able to provide reliable estimates for the cheap function response.

Figure 47 – Approximation of y_l by the LF model prediction.



Source: the author

Then, we must fit the difference model. Now, the ln-likelihood function is given by:

$$\ln L_d = -\frac{n_h}{2}\ln(2\pi) - \frac{n_h}{2}\ln(\sigma_d^2) - \frac{1}{2}\ln|\boldsymbol{\Psi}_d(\mathbf{x}_h,\mathbf{x}_h)| - \frac{(\mathbf{d}-\mathbf{1}\,\mu_d)^T\,\boldsymbol{\Psi}_d(\mathbf{x}_h,\mathbf{x}_h)^{-1}\,(\mathbf{d}-\mathbf{1}\,\mu_d)}{2\,\sigma_d^2} \tag{6.15}$$

where the MLEs for μ_d and σ_d are:

$$\begin{aligned}\hat{\mu}_d &= \frac{\mathbf{1}^T\boldsymbol{\Psi}_d(\mathbf{x}_h,\mathbf{x}_h)^{-1}\mathbf{d}}{\mathbf{1}^T\boldsymbol{\Psi}_d(\mathbf{x}_h,\mathbf{x}_h)^{-1}\mathbf{1}}\\ \hat{\sigma}_d^2 &= \frac{(\mathbf{d}-\mathbf{1}\hat{\mu}_d)^T\boldsymbol{\Psi}_d(\mathbf{x}_h,\mathbf{x}_h)^{-1}(\mathbf{d}-\mathbf{1}\hat{\mu}_d)}{n_h}\end{aligned} \tag{6.16}$$

Then, substituting these estimators in Eq. (6.15) and removing the constant terms, the difference model concentrated ln-likelihood function is given by:

$$\ln L_d \approx -\frac{n_h}{2}\ln(\hat{\sigma}_h^2) - \frac{1}{2}\ln|\boldsymbol{\Psi}_d(\mathbf{x}_h,\mathbf{x}_h)| \tag{6.17}$$

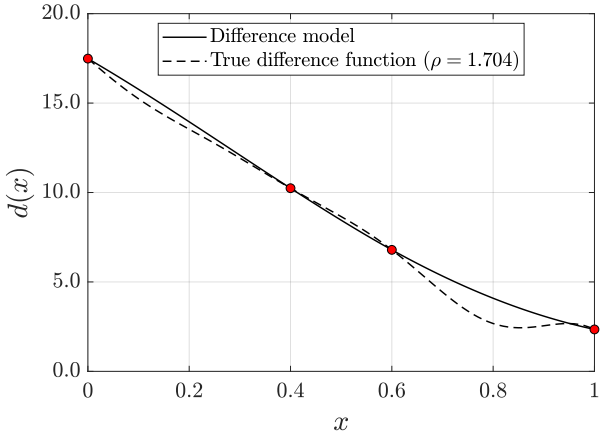
Once again, $p_{k,h}$ can be set to 2.0. Here, in addition to the correlation function hyper-parameters, we must also fit the scaling factor ρ. Thus, the evaluation of the MLE for the parameters is an unconstrained optimization problem defined by:

$$\begin{cases} \text{find} & \{\rho, \theta_{1,d}, \theta_{2,d}, \ldots, \theta_{m,d}\} \\ \text{that minimizes} & -\ln L_d(\rho, \boldsymbol{\theta}_d) \\ \text{with} & \rho_L \leq \rho \leq \rho_U \\ & \boldsymbol{\theta}_L \leq \boldsymbol{\theta}_d \leq \boldsymbol{\theta}_U \end{cases} \tag{6.18}$$

Toal[319] suggests that, when the correlation between sources is poor, ρ should not be optimized, as this procedure tends to give the LF source more importance than it should in these cases.

Figure 48 shows an example of the difference model. The model represents how the HF and LF sources are correlated, and how they differ through the design space. In this case, the difference model is able to show that, for $x \in [0,1]$, $\mathbf{d}$ is expected to decrease smoothly, and it fits almost perfectly the true difference function.

Figure 48 – Difference model prediction.



Source: the author

That way, Co-Kriging building requires two optimization processes: one for the building of a LF model, where there are n_l design variables and a $n_l \times n_l$ correlation matrix $\mathbf{\Psi}_l$, and other for the difference model, where there are $n_h + 1$ design variables and a $n_h \times n_h$ correlation matrix $\mathbf{\Psi}_h$. The correlation matrix size is important, since, in each optimization process, it should be inverted multiple times. Therefore, computational complexity for building the Co-Kriging model is $O(N_l\, n_l^3\, m + N_d\, n_h^3\, m)$, where N_l is the number of times Eq. (6.11) is evaluated and N_d is the number of times Eq. (6.17) is evaluated.

Finally, the Co-Kriging prediction can be evaluated by:

$$\hat{y}_h(\mathbf{x}) = \hat{\mu} + \mathbf{c}^T\, \mathbf{C}^{-1}\, (\mathbf{y} - \mathbf{1}\,\hat{\mu}) \tag{6.19}$$

where $\mathbf{c}$ is given by:

$$\mathbf{c} = \begin{bmatrix} \rho\, \hat{\sigma}_l^2\, \boldsymbol{\psi}_l(\mathbf{x}_l, \mathbf{x}) \\ \rho^2\, \hat{\sigma}_l^2\, \boldsymbol{\psi}_l(\mathbf{x}_h, \mathbf{x}) + \hat{\sigma}_d^2\, \boldsymbol{\psi}_d(\mathbf{x}_h, \mathbf{x}) \end{bmatrix} \tag{6.20}$$

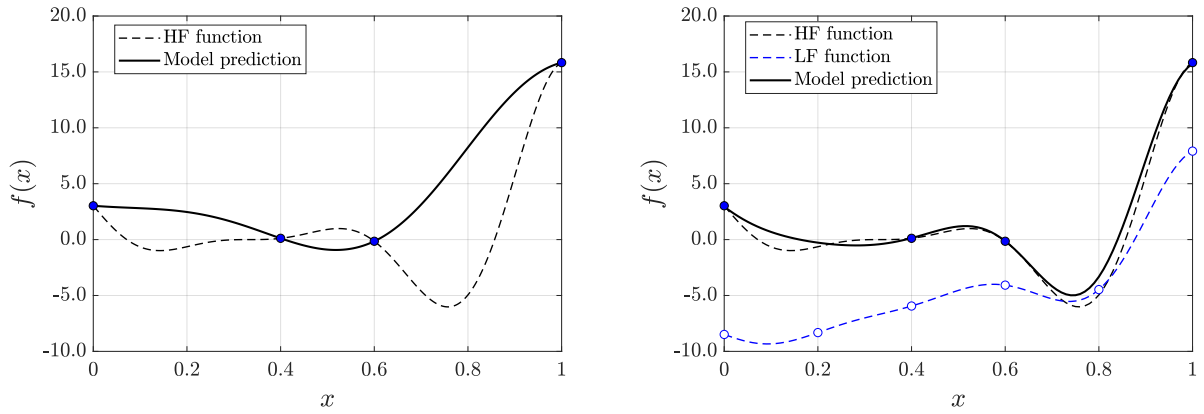
and the MLE for μ is:

$$\hat{\mu} = \frac{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{y}}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}} \tag{6.21}$$

which can be computed beforehand along with $\mathbf{C}^{-1}\, (\mathbf{y} - \mathbf{1}\,\hat{\mu})$.

Figure 49 presents an example of the Co-Kriging predictor along with a regular single-fidelity Kriging predictor. The consideration of information from a LF source allows for a much better exploration of the design space, and the Co-Kriging fits almost perfectly the HF function. Similar to the single-fidelity Kriging model, the approximation quality of Co-Kriging is a function of the number of hyper-parameters[314]. This means that prediction accuracy decreases for high-dimensional functions.

Figure 49 – Improvement of the model accuracy by the use of the Co-Kriging.



(a) Ordinary Kriging (OK) prediction

(b) Co-Kriging prediction

Source: the author

Recently, some researchers tried to improve upon Co-Kriging initial formulation to better tackle some optimization problems. In cases where data is not nested, Ruan et al.[74] proposed the Improved Cooperative Kriging (ICK) model, which presents a different approach to predict the difference vector. The authors state that the ICK performs much better than Co-Kriging when sources are not well-correlated. Xiao et al.[310] extended the Co-Kriging formulation to multi-level multi-fidelity data by attributing different weights to different levels of fidelity.

6.1.1 Adaptive sampling

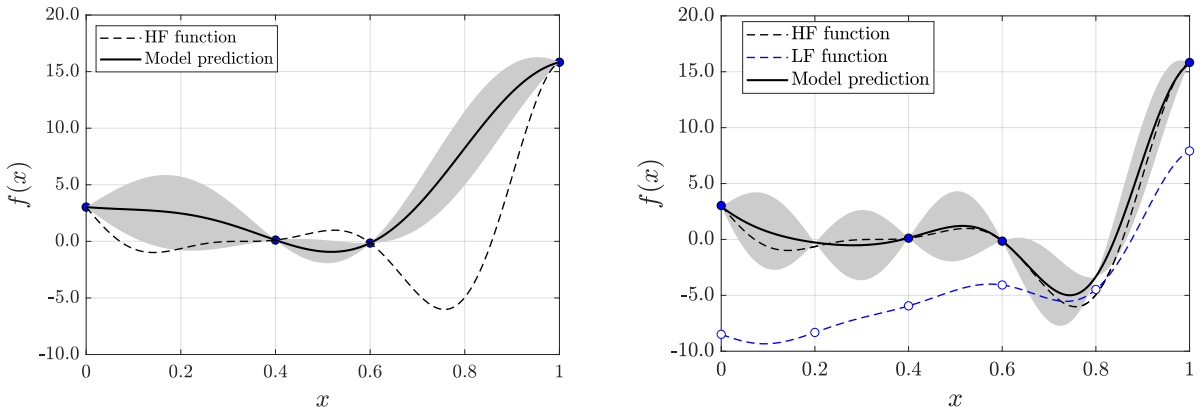
The Co-Kriging model also allows for the use of error-based exploration to assist in the choice of new sampling points[67]. This is very similar to the single-fidelity model case, discussed in greater detail in Section 5.3. On the Co-Kriging model, the posterior variance may be computed by[40]:

$$\hat{s}^2(\mathbf{x}) \approx \rho^2 \hat{\sigma}_l^2 + \hat{\sigma}_d - \mathbf{c}^T \mathbf{C}^{-1} \mathbf{c} + \frac{(1 - \mathbf{1}^T \mathbf{C}^{-1} \mathbf{c})^2}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}} \tag{6.22}$$

Similar to the Ordinary Kriging (OK), the final term is often ignored since it is frequently very small.

Figure 50 shows the confidence interval ($\hat{y} \pm \hat{s}$) for the Ordinary Kriging (OK) and the Co-Kriging prediction. For the Co-Kriging, while variance is always zero at the HF sampling points (as in the OK), it may not be zero at the location of the LF data. However, if the model is confident on the difference model prediction, variance in these points will be very low and, thus, these will not be favored in error-based exploration. This "confidence" is related to lower values of $\theta_{i,d}$[40].

Figure 50 – Confidence interval using different models.



(a) Ordinary Kriging (OK)

(b) Co-Kriging

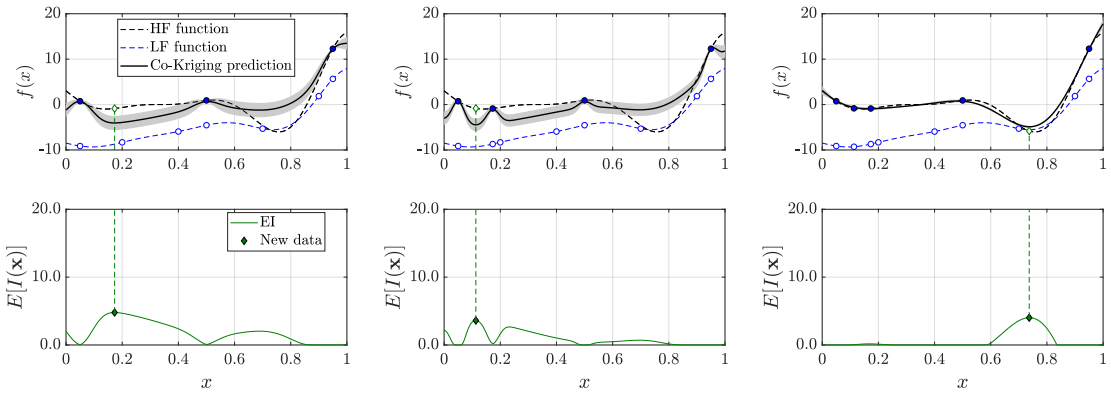
Source: the author

After assessing the posterior variance, we may employ any of the infill criteria discussed in Section 5.3, such as the LCB or the EI. Figure 51 presents three iterations of the EI using Co-Kriging. Due to a bad choice of initial data, the initial model prediction is very poor. However, the EI is still able to find a design very close to the global optima on the third iteration.

Unfortunately, for the Co-Kriging model, there is no easy method to choose the right fidelity to sample in a given iteration. Here, each new point is evaluated by both sources (HF and LF). Although simple, this approach is performed by different researchers[43, 309]. Kontogiannis and Savill[77] propose that, if the error predicted on the new data is lower than a threshold, then only the LF source is used. Even so, if q consecutive iterations are performed without a HF evaluation, then both HF and LF sources are evaluated. Here, q may be related to the relative expense:

$$q = \frac{1}{C_r} = \frac{C_e}{C_c} \tag{6.23}$$

Figure 51 – Expected Improvement (EI) criterion for Co-Kriging.



(a) Iteration 1 (b) Iteration 2 (c) Iteration 3

Source: the author

Although interesting, it is not easy to define an appropriate threshold for the method. Another interesting approach is the Augmented Expected Improvement (AEI), which was proposed by Huang et al.[324]. Here, the EI is evaluated for each by fidelity by including two multiplicative factors, related to the correlation between models for a given $\mathbf{x}$ and the relative expense[326, 37]. This allows for the definition of the source to be evaluated. However, there is no statistical or mathematical justification for the approach, and other authors have already shown that the method might present unsatisfactory results[73].

6.2 Hierarchical Kriging

The Hierarchical Kriging (HK) was proposed by Han and Görtz[72] with the main goal of being a more simple multi-fidelity approach, although as robust as other alternatives such as the Co-Kriging. The Hierarchical Kriging (HK) considers that the Kriging trend term, which is constant in Ordinary Kriging (OK), is based on a Kriging model built using the LF sample. Thus, Eq. (5.16) can be written as[72]:

$$\hat{y}(\mathbf{x}) = \beta_0 \hat{y}_l(\mathbf{x}) + Z(\mathbf{x}) \tag{6.24}$$

where β_0 is a constant scaling factor and $Z(\mathbf{x})$ represents the autocorrelated localized deviations. Similar to regular Kriging, $Z(\mathbf{x})$ can be assumed to come from a stochastic process with mean zero and covariance given by:

$$\text{cov}(\mathbf{y},\mathbf{y}) = \sigma^2 \boldsymbol{\Psi} \tag{6.25}$$

where $\mathbf{\Psi}$ is the correlation matrix, which depends on the correlation function. Again, using the Gaussian kernel:

$$\Psi_{ij} = \psi(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\sum_{k=1}^{m} \theta_k \left| x_{i,k} - x_{j,k} \right|^{p_k} \right) \tag{6.26}$$

In the HK model, just like in the Co-Kriging, the process starts by fitting a Kriging model to the LF sample. Thus, the model hyper-parameters should be set by solving the unconstrained optimization problem:

$$\begin{cases} \text{find} & \boldsymbol{\theta}_l = \{\theta_{1,l}, \theta_{2,l}, \ldots, \theta_{m,l}\} \\ \text{that minimizes} & -\ln L_l(\boldsymbol{\theta}_l) \\ \text{with} & \boldsymbol{\theta}_L \leq \boldsymbol{\theta}_l \leq \boldsymbol{\theta}_U \end{cases} \tag{6.27}$$

where $\ln L_l$ is given by Eq. (6.11).

This procedure allows us to perform an estimate for the LF Kriging model by:

$$\hat{y}_l(\mathbf{x}) = \hat{\mu}_l + \boldsymbol{\psi}_l^T \, \mathbf{\Psi}_l^{-1} \left(\mathbf{y}_l - \mathbf{1} \, \hat{\mu}_l \right) \tag{6.28}$$

To understand the correlation between the HF and LF sources, the LF model prediction is evaluated at the HF sampling points:

$$\mathbf{F} = \hat{y}_l(\mathbf{x}_h) \tag{6.29}$$

where, if data is nested, $\mathbf{F} \subset \mathbf{y}_l$. Then, the building of the HK can be performed by maximizing its corresponding ln-likelihood function:

$$\ln L = -\frac{n_h}{2} \ln(2\pi) - \frac{n_h}{2} \ln(\sigma^2) - \frac{1}{2} \ln \left| \mathbf{\Psi}(\mathbf{x}_h, \mathbf{x}_h) \right| - \frac{(\mathbf{y}_h - \beta_0 \mathbf{F})^T \, \mathbf{\Psi}(\mathbf{x}_h, \mathbf{x}_h)^{-1} \, (\mathbf{y}_h - \beta_0 \mathbf{F})}{2\sigma^2} \tag{6.30}$$

By differentiating and equating to zero, the MLE for β_0 and σ^2 can be defined as:

$$\begin{aligned} \hat{\beta}_0 &= \frac{\mathbf{F}^T \mathbf{\Psi}^{-1} \mathbf{y}_h}{\mathbf{F}^T \mathbf{\Psi}^{-1} \mathbf{F}} \\ \hat{\sigma}^2 &= \frac{(\mathbf{y}_h - \beta_0 \mathbf{F})^T \mathbf{\Psi}(\mathbf{x}_h, \mathbf{x}_h)^{-1} (\mathbf{y}_h - \beta_0 \mathbf{F})}{n_h} \end{aligned} \tag{6.31}$$

The assessment of $\hat{\beta}_0$ is very similar to the evaluation of $\hat{\mu}$ in Ordinary Kriging (OK), as shown in Eq. (5.22). However, instead of a constant additive term, β_0 represents a scaling factor applied to $\hat{y}_l$, which then serves as the HK trend.

Substituting Eq. (6.31) in Eq. (6.30), we can derive the Hierarchical Kriging (HK) concentrated ln-likelihood function:

$$\ln L \approx -\frac{n_h}{2}\ln(\hat{\sigma}^2) - \frac{1}{2}\ln|\boldsymbol{\Psi}(\mathbf{x}_h,\mathbf{x}_h)| \tag{6.32}$$

Then, the rest of HK hyper-parameters can be set by maximizing this function. Considering $p_k = 2.0$ for simplification, the MLE for $\boldsymbol{\theta}$ can be found by the following unconstrained optimization problem:

$$\begin{cases} \text{find} & \boldsymbol{\theta} = \{\theta_1, \theta_2, \ldots, \theta_m\} \\ \text{that minimizes} & -\ln L(\boldsymbol{\theta}) \\ \text{with} & \boldsymbol{\theta}_L \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_U \end{cases} \tag{6.33}$$

Similar to Co-Kriging, the building of the Hierarchical Kriging (HK) model also requires two optimization processes: one for the building of a Kriging model based on the LF sample, and other for the building of the HK itself. While the former considers n_l design variables and presents a $n_l \times n_l$ correlation matrix $\boldsymbol{\Psi}_l$, the other considers n_h design variables and presents a $n_h \times n_h$ correlation matrix $\boldsymbol{\Psi}$. Note that the size of these matrices is the same as those from Co-Kriging, which means that there should be little difference in the building cost for these two methods, as their time complexity is very similar: $\mathrm{O}(N_l\, n_l^3\, m + N_h\, n_h^3\, m)$. Here, N_l is the number of times Eq. (6.11) is evaluated, while N_h is the number of times Eq. (6.32) is evaluated.

Once the HK model is built, its prediction can be evaluated by:

$$\hat{y}(\mathbf{x}) = \hat{\beta}_0\, \hat{y}_l(\mathbf{x}) + \boldsymbol{\psi}^T\, \boldsymbol{\Psi}^{-1}\, (\mathbf{y}_h - \hat{\beta}_0\, \mathbf{F}) \tag{6.34}$$

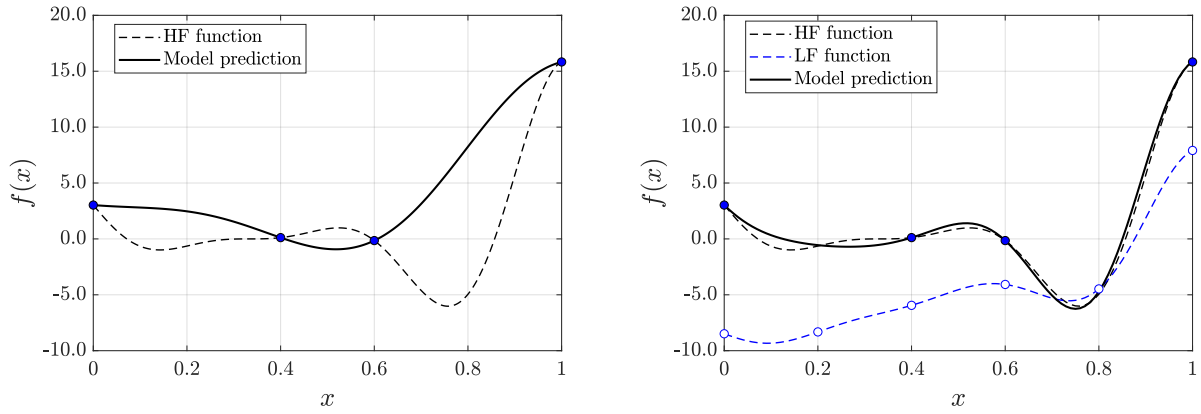
where $\boldsymbol{\psi}$ is given by:

$$\psi_i(\mathbf{x}) = \exp\left(-\sum_{k=1}^{m} \theta_k\, |x_{i,k} - x_k|^{p_k}\right) \tag{6.35}$$

It is worth noting that, while the evaluation of the Co-Kriging model uses a $(n_l + n_h) \times (n_l + n_h)$ covariance matrix $\mathbf{C}$, the evaluation of the HK uses two correlation matrices: $\boldsymbol{\Psi}$ and $\boldsymbol{\Psi}_l$ (for the evaluation of $\hat{y}_l$), which are $n_h \times n_h$ and $n_l \times n_l$, respectively. This means that the evaluation of the HK model is cheaper than the evaluation of the Co-Kriging model.

Figure 52 presents an example of the Hierarchical Kriging (HK) prediction. Although simpler than the Co-Kriging, the HK also seems to attain a good approximation quality, representing the HF source almost perfectly while only performing four HF evaluations.

Figure 52 – Improvement of the model accuracy by the use of the Hierarchical Kriging (HK).



(a) Ordinary Kriging (OK) prediction

(b) Hierarchical Kriging (HK) prediction

Source: the author

6.2.1 Adaptive sampling

For Hierarchical Kriging (HK), the assessment of the posterior variance is given by:

$$\hat{s}^2(\mathbf{x}) = \hat{\sigma}^2 \left[1 - \boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1} \boldsymbol{\psi} + \frac{(\hat{y}_l - \mathbf{F}^T \boldsymbol{\Psi}^{-1} \boldsymbol{\psi})^2}{\mathbf{F}^T \boldsymbol{\Psi}^{-1} \mathbf{F}} \right] \tag{6.36}$$

This time, the third term might be relevant and, thus, should not be neglected. Using the model prediction $\hat{y}(\mathbf{x})$ and the posterior variance $\hat{s}^2(\mathbf{x})$, all infill criteria discussed in Section 5.3 can be employed, such as the EI and the PI. Similar to Co-Kriging, the choice of the infill point's fidelity can be performed arbitrarily, or simply based on the relative expense of each source.

However, to better deal with that matter using HK, Zhang, Han and Zhang[73] proposed the VF-EI. Here, along with the HK posterior variance, the LF model posterior variance is also evaluated by:

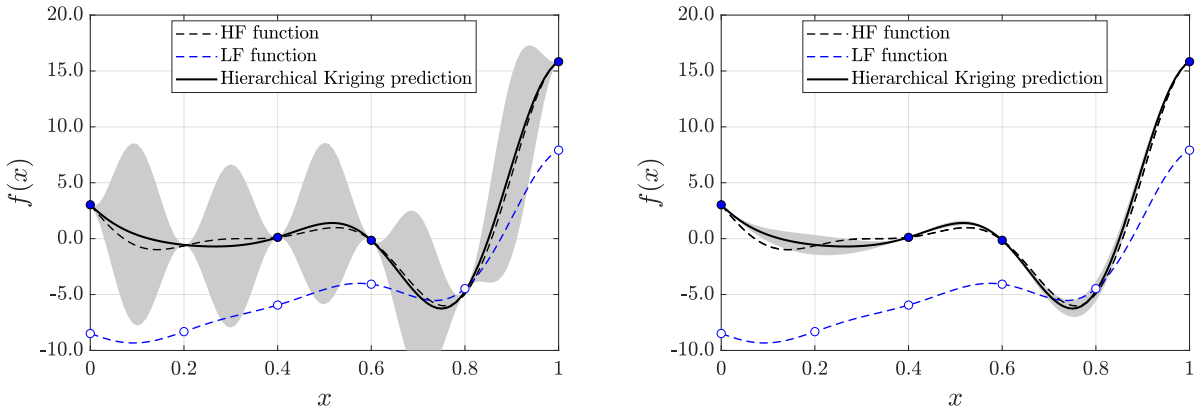
$$\hat{s}_l^2(\mathbf{x}) = \hat{\sigma}_l^2 \left[1 - \boldsymbol{\psi}_l^T \boldsymbol{\Psi}_l^{-1} \boldsymbol{\psi}_l + \frac{(1 - \mathbf{1}^T \boldsymbol{\Psi}_l^{-1} \boldsymbol{\psi}_l)^2}{\mathbf{1}^T \boldsymbol{\Psi}_l^{-1} \mathbf{1}} \right] \tag{6.37}$$

which is the same as the Ordinary Kriging (OK) variance, shown in Eq. (5.28). Note that $\hat{s}_l$, by itself, would guide the infilling based on the improvement of the LF source[73], which is not helpful. However, the HK allows for an easy method to guarantee that both posterior variances assist in finding new infill points related to the improvement over the HF source[73]. Here, a variable-fidelity posterior variance $\hat{s}^2(\mathbf{x}, f)$ can be evaluated by:

$$\hat{s}^2(\mathbf{x}, f) = \begin{cases} \beta_0^2 \, \hat{s}_l^2(\mathbf{x}) & \text{, if } f = 1 \\ \hat{s}^2(\mathbf{x}) & \text{, if } f = 2 \end{cases} \tag{6.38}$$

where $f = 1$ for the LF model and $f = 2$ for the HK model. Figure 53 shows the confidence interval $(\hat{y} \pm \hat{s})$ for the HK model, comparing $\hat{s}$ for $f = 1$ or $f = 2$. The scaled LF posterior variance is much greater than that of the HK one. Indeed, Zhang, Han and Zhang[73] comment that, if the LF model is able to correctly capture the trend of the HF source, the uncertainty of the HK model is usually much lower than the LF model. This way, this approach tends to add more LF data during its iterations, which is a preferred feature for improving computational efficiency.

Figure 53 – Confidence interval of the HK model.



(a) LF model ($f = 1$) (b) HK model ($f = 2$)

Source: the author

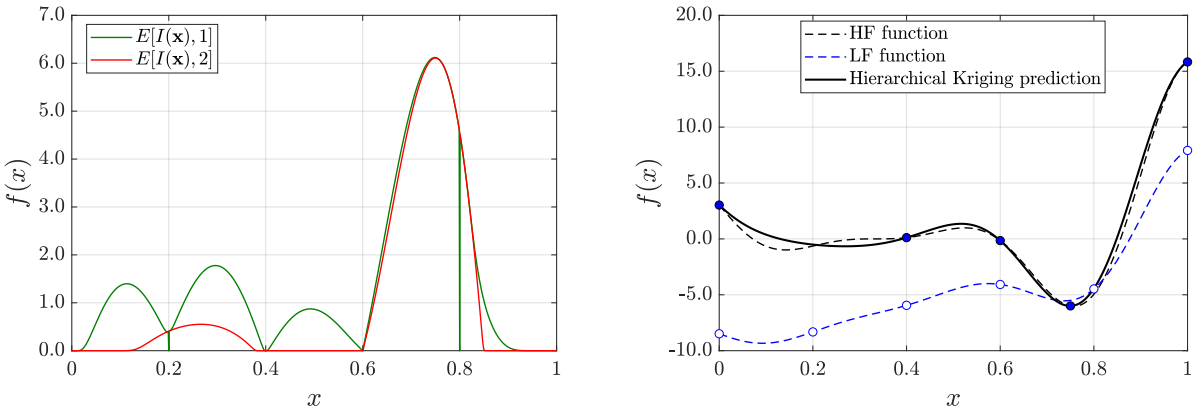
After evaluating the posterior variance for both fidelities, the VF-EI can be computed by:

$$E[I(\mathbf{x}), f] = \begin{cases} (y_{min} - \hat{y}(\mathbf{x}))\,\Phi\left(\dfrac{y_{min} - \hat{y}(\mathbf{x})}{\hat{s}(\mathbf{x}, f)}\right) + \hat{s}(\mathbf{x}, f)\,\phi\left(\dfrac{y_{min} - \hat{y}(\mathbf{x})}{\hat{s}(\mathbf{x}, f)}\right) & , \text{if } \hat{s}(\mathbf{x}, f) > 0 \\ 0 & , \text{if } \hat{s}(\mathbf{x}, f) = 0 \end{cases} \tag{6.39}$$

Then, whichever fidelity attains the highest EI is chosen to evaluate the new infill point. Figure 54 shows how the VF-EI behaves for the same sampling set used in Figure 53. Here, local exploitation is favored, and the location and value of the maximum VF-EI are basically the same for both fidelities ($f = 1, 2$). In these cases, both sources can be evaluated, and the new point is indeed very close to the global optima.

In other cases, when global exploration is important, the difference in fidelities for the evaluation of $\hat{s}^2$ is very important. Figure 55 shows three iterations for the VF-EI using a very poor initial sample. The first point is evaluated only by the LF source. The second one, however, is evaluated by both sources, as it favors exploitation. Indeed, this point is very close to the global optima. For the third point, only the LF source is evaluated, once again.

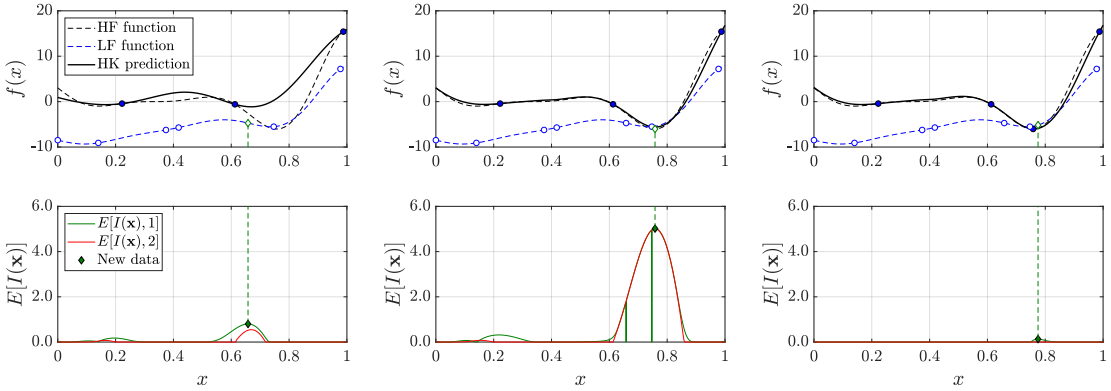
Figure 54 – Addition of a new point by the VF-EI criterion.



(a) LF model ($f = 1$) (b) HK model after the infill

Source: the author

Figure 55 – Variable Fidelity Expected Improvement (VF-EI) criterion for HK.



(a) Iteration 1 (b) Iteration 2 (c) Iteration 3

Source: the author

This approach allows for a much more efficient method, since the expensive HF source does not need to be evaluated in every iteration, but only where it is expected to provide a major improvement to the model. The authors also argue that, in comparison to the AEI (which was discussed shortly for the Co-Kriging method), the VF-EI can be derived analytically and is free of empirical parameters, which means that the method is fully adaptive[73]. Recently, Guo et al.[78] and Guo et al.[79] used the VF-EI in the unconstrained optimization of Variable Stiffness Composites (VSCs), showing very good results.

Ruan et al.[76] proposed the VF-PI, which is very similar to the VF-EI, but the choice of infill points is now performed by the maximization of the $P[I(\mathbf{x},f)]$:

$$P[I(\mathbf{x}),f] = \Phi\left(\frac{y_{min} - \hat{y}(\mathbf{x})}{\hat{s}(\mathbf{x},f)}\right) \tag{6.40}$$

The authors state that the method is able to provide better results when the correlation between

the LF and HF sources is high.

6.2.1.1 *Novel variable fidelity acquisition functions*

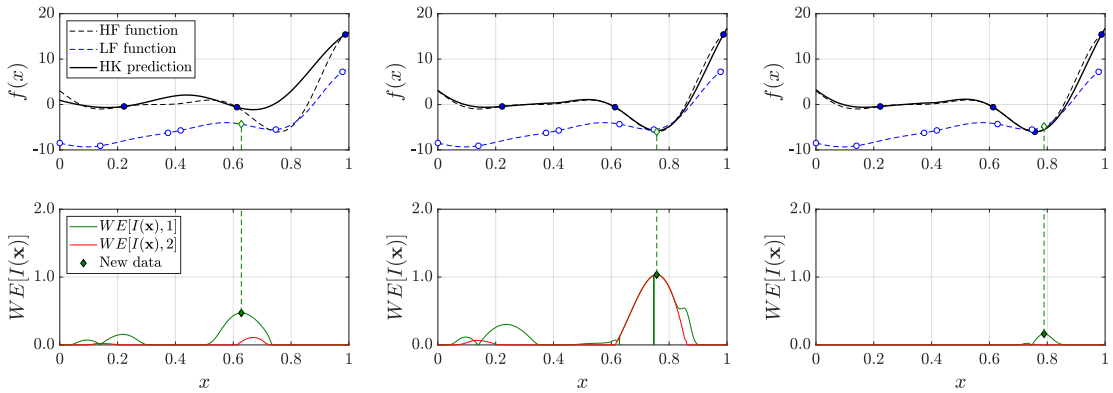
Following the same idea, similar methods can also be derived for other criteria, e.g. the LCB and the WEI. Thus, if $\hat{s}(\mathbf{x}, f) > 0$, we may define a Variable Fidelity Weighted Expected Improvement (VF-WEI) method by:

$$WE[I(\mathbf{x}), f] = w\,(y_{min} - \hat{y}(\mathbf{x}))\,\Phi\left(\frac{y_{min} - \hat{y}(\mathbf{x})}{\hat{s}(\mathbf{x}, f)}\right) + (1-w)\,\hat{s}(\mathbf{x}, f)\,\phi\left(\frac{y_{min} - \hat{y}(\mathbf{x})}{\hat{s}(\mathbf{x}, f)}\right) \tag{6.41}$$

where $w \in [0, 1]$ is an user-defined parameter which allows the user to control the importance of exploration and exploitation aspects. Lower w favor exploration, while higher w favor exploitation. Note that this method is very similar to the VF-EI proposed by Zhang, Han and Zhang[73]. However, the consideration of w makes it so that the user has more control over how the selection of new data points should be performed.

Figure 56 depicts how the VF-WEI behaves for the first three iterations, considering $w = 0.2$. Note that, while the data chosen is very similar to those from the VF-EI method (shown in Figure 55), the acquisition function prioritizes regions further from the current data points. This is especially noticeable in Iterations 1 and 2. This is very important for the optimization of complex multi-modal functions.

Figure 56 – Variable Fidelity Weighted Expected Improvement (VF-WEI) criterion for HK.



(a) Iteration 1 (b) Iteration 2 (c) Iteration 3

Source: the author

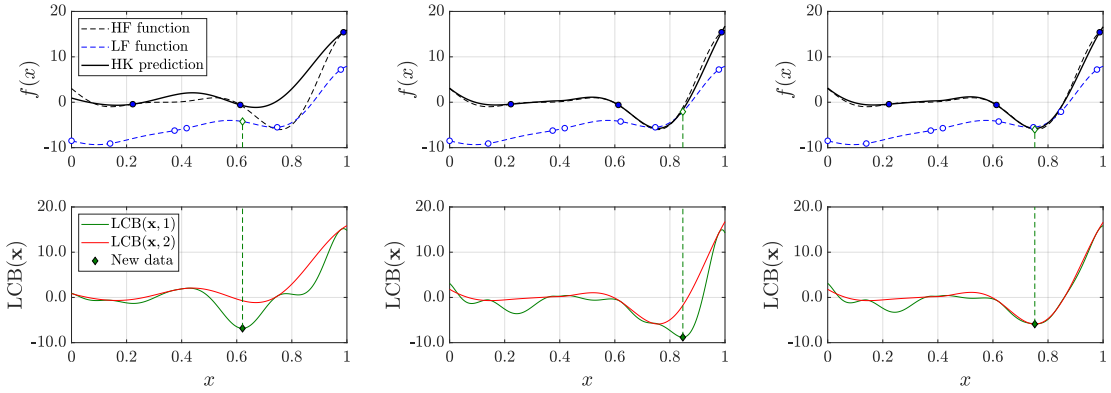
Alternatively, one may also derive a Variable Fidelity Lower Confidence Bound (VF-LCB) approach as in:

$$\text{LCB}(\mathbf{x}, f) = \hat{y}\,(\mathbf{x}) - \beta\,\hat{s}(\mathbf{x}, f) \tag{6.42}$$

where, this time, $\beta \leq 0$ is a user-defined parameter. While lower β favor exploitation, higher β favor exploitation.

Figure 57 shows how the VF-LCB method behaves for $\beta = 3.0$. Note that, in Iterations 1 and 2, the method greatly favors exploration, adding data from the LF source. For Iteration 3, however, exploitation is more important, and both the LF and HF sources are evaluated. Note that, here, the point is very close to the optima.

Figure 57 – Variable Fidelity Lower Confidence Bound (VF-LCB) criterion for HK.



(a) Iteration 1 (b) Iteration 2 (c) Iteration 3

Source: the author

6.3 Computational implementation

In this work, optimization of composite structures will be performed using BIOS, acronym for Biologically Inspired Optimization System. This is an in-house software developed in LMCV to perform structural optimization using meta-heuristics[327].

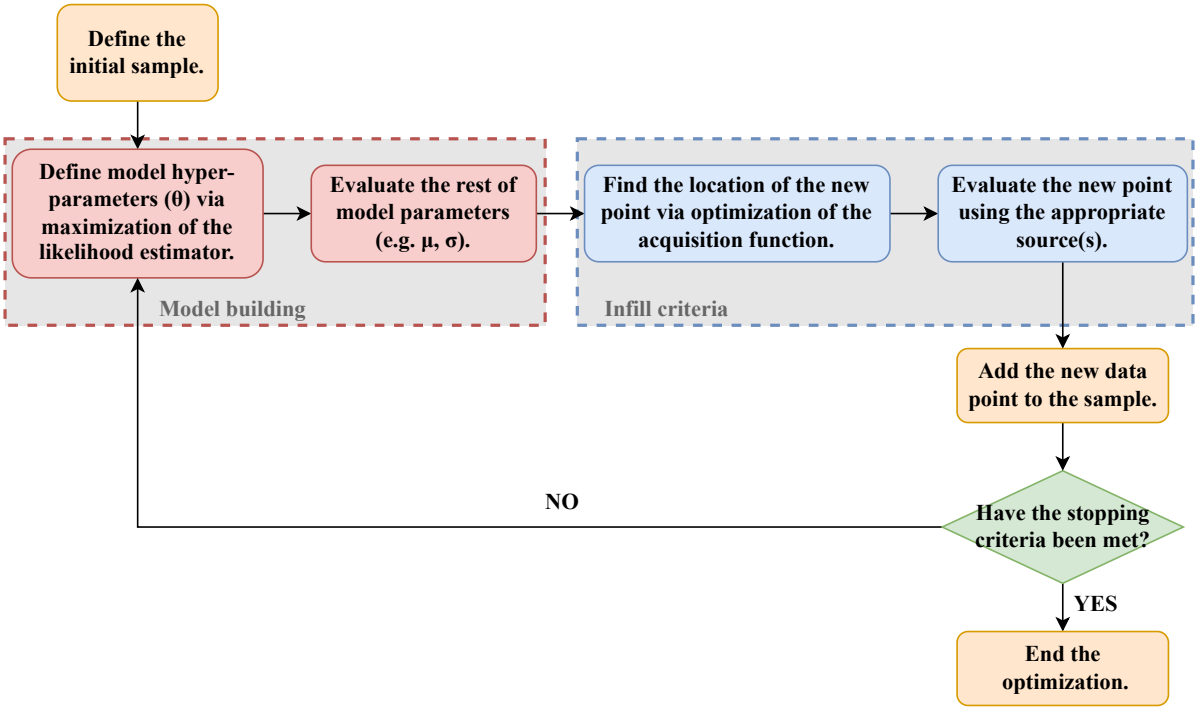
Initially, BIOS was focused on optimization by nature-inspired algorithms[15, 16]. However, recent works have been extending BIOS to SBO, allowing for different single-fidelity models (RBF and Kriging) and acquisition functions (PI, EI, and WEI)[2, 282].

In this work, we focused on the implementation of MFMs, namely the Co-Kriging and the Hierarchical Kriging (HK). Furthermore, we improved the handling of discrete and constrained optimization, adding new methods to BIOS. Also, the LCB acquisition function was added as an alternative to the probabilistic acquisition functions.

7 NUMERICAL EXAMPLES

In this chapter, the results obtained in the optimization of different problems using Sequential Approximate Optimization (SAO) will be presented in three topics. First, the algorithms will be put to test using mathematical and analytical engineering benchmarks. The results found will be compared with those from the literature whenever it is possible. Second, the algorithms will be applied to problems involving FG structures. Finally, these will be applied to laminate problems. Figure 58 shows a general flowchart for the algorithm employed in this work. Several methods will be tested, which will be presented in the following.

Figure 58 – Algorithm employed for SAO in this work.



Source: the author

The initial sample will be defined via the LHS_N method, with $N = 20$. For Multi-Fidelity Models (MFMs), the sample will always be nested, and the HF sample will be drawn from a subset of the LF sample. As for the number of data points used, these will be defined in each problem individually. Regarding the model building phase, a comparative study will be carried out using several modeling criteria. The following models will be tested (along with their respective acronyms):

- Kriging - KRG;
- Cooperative Kriging - COKRG;
- Hierarchical Kriging - HKRG.

Note that, for Multi-Fidelity Models (MFMs), the maximization of the likelihood estimator is performed twice in each iteration, as described in Sections 6.1 and 6.2.

Concerning infill criteria, four main acquisition functions will be used in this work:

- Lower Confidence Bound (LCB);
- Probability of Improvement (PI);
- Expected Improvement (EI);
- Weighted Expected Improvement (WEI).

In these cases, for a more robust process, both fidelities will be evaluated in each iteration. That means that, after using the acquisition function to define the most promising new data point, the new trial design will be evaluated using the HF and LF sources.

For the Hierarchical Kriging, four other acquisition functions will be tried out, which correspond to the variable-fidelity version of each method:

- Variable Fidelity Lower Confidence Bound (VF-LCB);
- Variable Fidelity Probability of Improvement (VF-PI);
- Variable Fidelity Expected Improvement (VF-EI);
- Variable Fidelity Weighted Expected Improvement (VF-WEI).

Since sample is nested, all iterations will add a LF data point. However, the addition of the HF data point will follow what was discussed in Section 6.2.1.

Note that the VF-LCB and the VF-WEI criteria have not yet been proposed in the literature, even though these can be easily derived similar to the VF-EI and the VF-PI, as shown in Section 6.2.1.1. Also, all LCB and WEI based approaches will employ a cyclic approach, where $\beta \in [1.0, 2.0, 3.0]$ for the LCB and $w \in [0.20, 0.35, 0.50]$ for the WEI.

In some problems, it is necessary to use an appropriate constraint-handling technique. For the LCB method, a simple adaptive penalty approach will be employed[186]. For the probabilistic infill criteria (namely the EI, PI, and WEI), different approaches will also be tried out:

- Direct approach (Direct);
- Probability of Feasibility (PF);
- Feasibilibity function proposed by Tutum, Deb and Baran[62] (FFT);
- Feasibilibity function proposed by Bagheri et al.[63] (FFB);
- Feasibilibity function proposed by Sohst, Afonso and Suleman[297] (FFS).

Two main stopping criteria will be considered, one related to the maximum number

of High-Fidelity (HF) points evaluated (N_{max}) and other to the maximum number of iterations without improvement (Gen_{stall}). Unless stated otherwise, this work considers $N_{max} = 100$ and $Gen_{stall} = 10$.

Regarding the building of the models, which involves the maximization of several likelihood functions throughout the process, a PSO algorithm will be employed. For continuous problems, a PSO will also be employed for the optimization of acquisition functions. Table 11 presents the parameters for the PSO employed in each phase. The difference in the algorithm parameters is due to the higher cost of evaluating the likelihood function for model building.

Table 11 – Parameters used for the PSO algorithm employed for model building and infill criteria.

Phase	Model building	Infill criteria
N_{gen}	200	250
N_p	50	100
Gen_{stall}	50	100
w	0.72	0.72
c_1	1.50	1.50
c_2	1.50	1.50
Topology	Ring	Ring

The performance of SAO algorithms will be assessed using appropriate metrics. The accuracy of each approach will be defined by the NRMSE:

$$\text{NRMSE} = \frac{y_{min} - y_{opt}}{y_{opt}} \tag{7.1}$$

which compares the best result found with the true optimum. When necessary, the success rate will also be shown. The efficiency will be defined by the number of HF ($n_{ev,h}$) and LF ($n_{ev,l}$) evaluations performed in the entire process and during the adaptive sampling (which will be shown in parenthesis). The time spent through each approach will also be shown. Since all methods employed in this work are stochastic, $N_r = 10$ runs will be carried out for each problem and the average of these metrics will be shown.

All simulations were carried out on a computer running on an Intel i9-9820 X @ 3.30GHz, 10 cores, and 128 GB RAM. No parallelization procedure was adopted.

7.1 Mathematical and analytical engineering benchmarks

In this section, the results for the mathematical and analytical engineering benchmarks will be shown. Five different problems will be used. The first three are related to

mathematical problems commonly used in the literature for the optimization of multi-fidelity problems. To validate our implementation, our results will be compared to those from different papers using similar methods, namely: KRG-PI, KRG-EI, COKRG-PI, COKRG-EI, HKRG-VF-PI, and HKRG-VF-EI.

The other two problems are related to the optimization of simple beam structures. Here, a variety of methods will be employed, so that these may be compared appropriately. Focus on engineering problems will be further extended in the following sections.

Unless stated otherwise, the initial sample will be defined via the LHS_N method, with $N = 20$, using $n_h = 3m$ and $n_l = 6m$. Since these problems are used for benchmarking purposes, another stopping criterion will be considered: if the algorithm finds the optimum design, it stops. Furthermore, as these have a really fast evaluation, there is no reason to compute their time spent.

7.1.1 Forrester function

The first problem presented in this work is the minimization of the Forrester function:

$$y_h(x) = (6x-2)^2 \sin(12x-4)\,, x \in [0,1] \tag{7.2}$$

This is a well-known problem proposed by Forrester, Sobester and Keane[40] for Surrogate-Based Optimization (SBO), and is commonly used for example purposes in many works related to SAO and MFMs. The optimum of this problem is given by $f(0.757) = -6.021$.

In these works, the Low-Fidelity (LF) function is usually represented by:

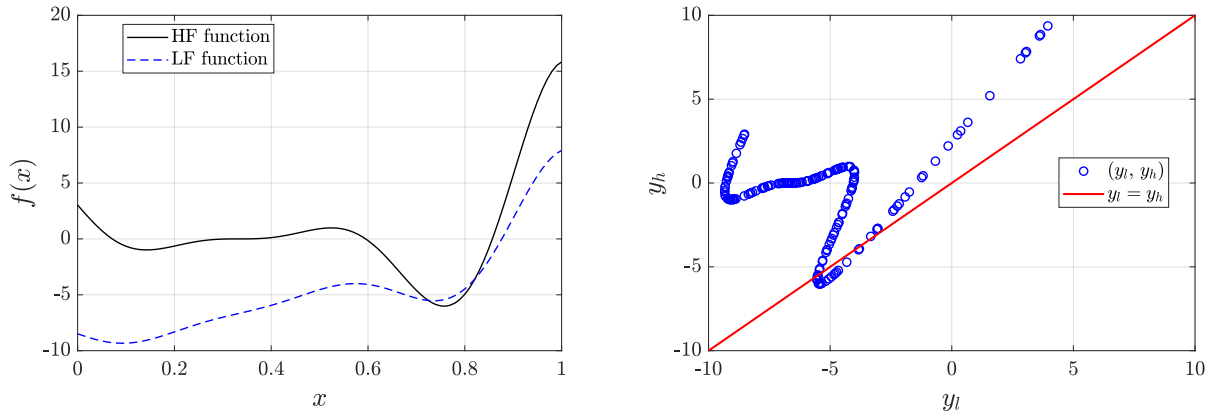
$$y_{l,a}(x) = A\,[6\,(x+D)-2]^2 \sin[12\,(x+D)-4] + B\,(x+D-0.5) + C\,, x \in [0,1] \tag{7.3}$$

with $A = 0.5$, $B = 10.0$, $C = -5.0$, and $D = 0.0$. Figure 59 shows these two functions, along with a representation of the correlation between the responses. Note that these functions show quite a distinct behavior, presenting a $R^2 = 0.55$. This problem will be known in this work as Forrester 1a.

To improve the understanding over the behavior of each method for different kinds of LF sources, other LF functions will also be tried out. For the problem Forrester 1b, the LF function is given by a shift in the y-direction:

$$y_{l,b}(x) = y_h(x) - 5 = (6x-2)^2 \sin(12x-4) - 5, x \in [0,1] \tag{7.4}$$

Figure 59 – Forrester 1a problem.

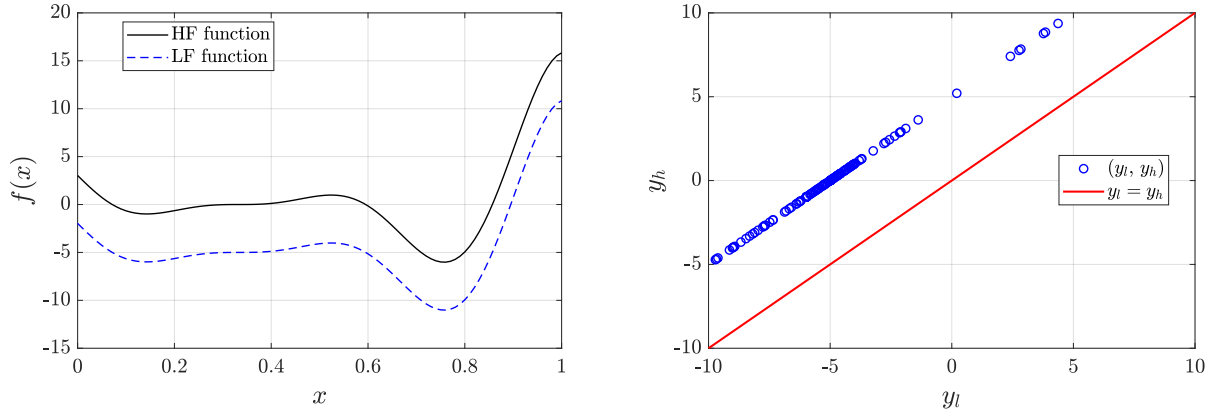


(a) HF and LF function (b) Correlation between fidelities

Source: the author

which means that the absolute difference between the two sources is constant ($y_h - y_l = 5$). This can be seen in Figure 60. Here, these two functions are very well correlated ($R^2 = 1.00$), even though their results are very different.

Figure 60 – Forrester 1b problem.



(a) HF and LF function (b) Correlation between fidelities

Source: the author

For the problem Forrester 1c, the LF function is given by a shift in the x-direction:

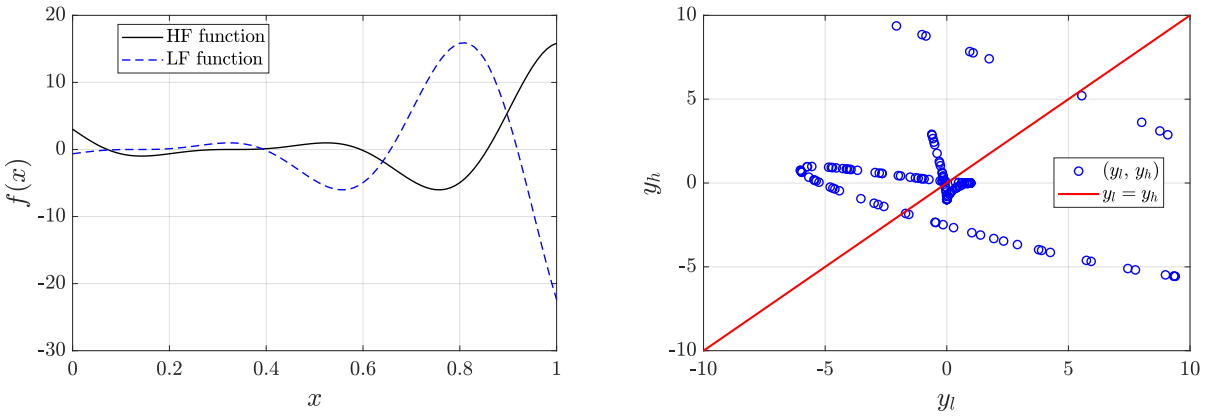
$$y_{l,c}(x) = y_h(x+0.2) = [6\,(x+0.2)-2]^2 \sin[12\,(x+0.2)-4], x \in [0,1] \tag{7.5}$$

which can be seen in Figure 61. This time, the correlation between these functions is poor ($R^2 = 0.57$).

Finally, for the problem Forrester 1d, the LF function is given by the consideration of a constant multiplicative term:

$$y_{l,d}(x) = 3\,y_h(x) = 3\,(6x-2)^2 \sin(12x-4), x \in [0,1] \tag{7.6}$$

Figure 61 – Forrester 1c problem.



(a) HF and LF function (b) Correlation between fidelities

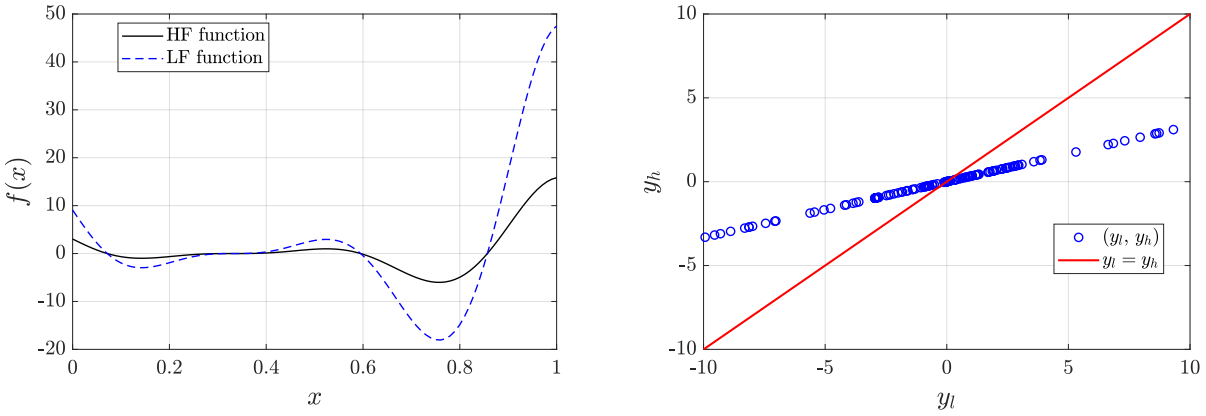
Source: the author

which means that the relative difference between the two fidelities is constant:

$$\frac{y_h - y_l}{y_h} = 2 \tag{7.7}$$

which can be seen in Figure 62. Once again, the correlation between these function is very high ($R^2 = 1.00$).

Figure 62 – Forrester 1d problem.



(a) HF and LF function (b) Correlation between fidelities

Source: the author

Problems Forrester1a, Forrester1b, and Forrester1c have been previously tested by Zhang, Han and Zhang[73], who solved them using three different algorithms: KRG-EI, COKRG-AEI, and HKRG-VF-EI. Forrester 1d, on the other hand, is a novel problem presented in this work. Table 12 shows the results obtained for these problems, using different modeling criteria and acquisition functions, also comparing those with the results found in the literature. In these 4 cases, our results for the KRG model are kept the same, since it is not affected by the different

LF models. It is important to remember that the Augmented Expected Improvement (AEI) is a criterion proposed by Huang et al.[324], where two seemingly arbitrary multiplicative factors are introduced to the EI function, allowing for the definition of which fidelity to add in each iteration. However, the method seems to heavily prioritize the addition of HF sampling points.

Table 12 – Averaged results for the Forrester problem.

Type	Source	SAO algorithm	NRMSE	Success rate	$n_{ev,h}$	$n_{ev,l}$
1a	Present work	KRG-PI	8.36%	90%	13 (10)	-
		KRG-EI	0.00%	100%	11 (8)	-
		COKRG-PI	8.36%	90%	10 (7)	13 (7)
		COKRG-EI	0.00%	100%	8 (5)	11 (5)
		HKRG-VF-PI	0.00%	100%	6 (3)	10 (4)
		HKRG-VF-EI	0.00%	100%	6 (3)	9 (3)
	Zhang, Han and Zhang[73]	KRG-EI	0.00%	-	13 (10)	-
		COKRG-AEI	0.00%	-	11 (8)	9 (3)
		HKRG-VF-EI	0.00%	-	9 (6)	25 (19)
1b	Present work	KRG-PI	8.36%	90%	13 (10)	-
		KRG-EI	0.00%	100%	11 (8)	-
		COKRG-PI	2.10%	30%	20 (17)	23 (17)
		COKRG-EI	1.57%	50%	17 (14)	20 (14)
		HKRG-VF-PI	0.00%	100%	9 (6)	12 (6)
		HKRG-VF-EI	0.00%	100%	6 (3)	10 (4)
	Zhang, Han and Zhang[73]	KRG-EI	0.02%	-	11 (8)	-
		COKRG-AEI	0.00%	-	11 (8)	7 (1)
		HKRG-VF-EI	0.00%	-	8 (5)	28 (22)
1c	Present work	KRG-PI	8.36%	90%	13 (10)	-
		KRG-EI	0.00%	100%	11 (8)	-
		COKRG-PI	8.39%	90%	11 (8)	14 (8)
		COKRG-EI	0.00%	100%	11 (8)	14 (8)
		HKRG-VF-PI	1.31%	80%	12 (9)	17 (11)
		HKRG-VF-EI	0.00%	100%	9 (6)	13 (7)
	Zhang, Han and Zhang[73]	KRG-EI	0.00%	-	11 (8)	-
		COKRG-AEI	0.00%	-	11 (8)	8 (2)
		HKRG-VF-EI	0.00%	-	10 (7)	17 (11)
1d	Present work	KRG-PI	8.36%	90%	13 (10)	-
		KRG-EI	0.00%	100%	11 (8)	-
		COKRG-PI	0.00%	90%	9 (6)	12 (6)
		COKRG-EI	0.00%	100%	7 (4)	10 (4)
		HKRG-VF-PI	0.00%	80%	9 (6)	12 (6)
		HKRG-VF-EI	0.00%	100%	6 (3)	10 (4)

For problem Forrester 1a, all approaches managed to find the optimum design in

all cases, except for the KRG-PI and the COKRG-PI, which presented a success rate of 90%. That is expected: since the problem is multimodal, the Probability of Improvement (PI) criterion should behave worse than the Expected Improvement (EI). However, the HKRG-PI still managed to find the optimum in all cases. Furthermore, the HKRG-based approaches presented the lowest number of evaluations. Our results are very similar to those presented by Zhang, Han and Zhang[73], although ours present a lower number of evaluations. This difference is likely related to the termination criteria employed.

In problem Forrester 1b, HKRG-based methods still managed to find the optimum with ease, even though they required slightly more evaluations (in comparison to Forrester 1a). However, the performance of the COKRG-based methods reduced dramatically, as they behaved worse than the single-fidelity KRG (both in terms of accuracy and number of evaluations). It is worth noting that the COKRG-AEI, used by Zhang, Han and Zhang[73], seemed to not present the same problem. However, we can understand the reasoning for this issue by taking a closer look at the Co-Kriging formulation. In problem Forrester 1b, the LF source is given by the addition of a constant term to the HF function. Considering that $y_h(\mathbf{x}) = y_l(\mathbf{x}) + C$, where C is a constant, the MLE for ρ can be derived from Eq. (6.15) as:

$$\frac{\partial \ln L_d}{\partial \rho} = 0 \quad \Rightarrow \quad \rho = \frac{\mathbf{y}_h - \mathbf{1}\,\mu_d}{\mathbf{y}_h + \mathbf{1}C} \tag{7.8}$$

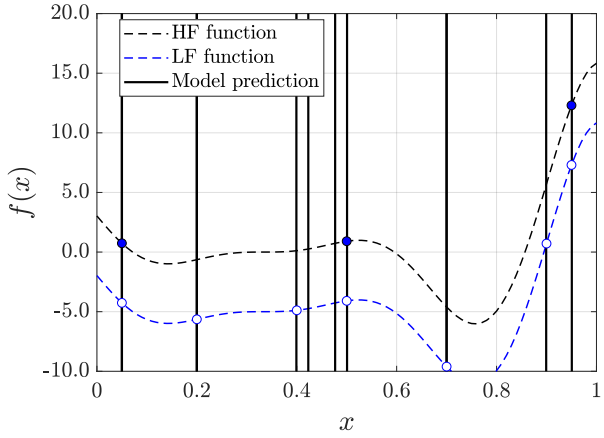
Since the MLE for μ_d is given by Eq. (6.16), to satisfy both equations, $\rho = 1.0$ and $\mu_d = -C$, as in:

$$\rho = \frac{\mathbf{y}_h - \mathbf{1}\,(-C)}{\mathbf{y}_h + \mathbf{1}C} = 1.0 \quad \text{and} \quad \hat{\mu}_d = -\frac{\mathbf{1}^T \boldsymbol{\Psi}_d(\mathbf{x}_h, \mathbf{x}_h)^{-1}\,\mathbf{1}}{\mathbf{1}^T \boldsymbol{\Psi}_d(\mathbf{x}_h, \mathbf{x}_h)^{-1}\mathbf{1}} C = -C \tag{7.9}$$

These also mean that, also by Eq. (6.16), $\sigma_d = 0.0$. Finally, as sample is nested, the covariance matrix, shown in Eq. (6.7), becomes singular, since it presents linearly dependent columns for $\rho = 1.0$ and $\sigma_d = 0.0$. For exemplification purposes, Figure 63 shows an example for the initial prediction of the Co-Kriging model for this problem. The model is not able to represent the true function, even though the correlation between the sources is very high.

The rest of the problems did not present the same issues. The results found for problem Forrester 1c were very similar to the ones found for Forrester 1a, even though, here, the HKRG-VF-PI also presented a slight decrease in accuracy and efficiency. The HKRG-VF-EI was still the most efficient method while maintaining a 100% success rate. This problem clearly shows the superiority of the EI for such multimodal problems.

Figure 63 – Co-Kriging initial prediction for problem Forrester 1b.



Source: the author

Finally, results found for Forrester 1d are very good across all models, even though some of them presented a success rate lower than 100% (namely, the ones which employed the PI-based methods).

Overall, Multi-Fidelity Models (MFMs) were able to reduce the number of HF evaluations required to find the global optimum, with a few occasional exceptions. In particular, the COKRG model suffers the most when the LF source is related to the incorporation of a constant additive term. Moreover, these results show that the correlation factor (R^2) by itself is not sufficient to determine if Bayesian MFMs will improve upon the single-fidelity model response. The absolute and relative difference between fidelities also seems to play a role in the performance of the MFM.

7.1.2 Hartmann3 function

The minimization of the Hartmann3 function is also commonly performed for benchmarking purposes for SBO[53, 73]. This is a 3-dimensional function given by:

$$y_h(\mathbf{x}) = -\sum_{i=1}^{4} c_i \exp\left[-\sum_{j=1}^{3} \alpha_{ij}\left(x_j - p_{ij}\right)^2\right], x_i \in [0,1] \tag{7.10}$$

where

$$\alpha_{ij} = \begin{bmatrix} 3.0 & 10 & 30 \\ 0.1 & 10 & 35 \\ 3.0 & 10 & 30 \\ 0.1 & 10 & 35 \end{bmatrix}, \quad c_i = \begin{bmatrix} 1 \\ 1.2 \\ 3 \\ 3.2 \end{bmatrix}, \quad p_{ij} = \begin{bmatrix} 0.3689 & 0.1170 & 0.2673 \\ 0.4699 & 0.4387 & 0.7470 \\ 0.1091 & 0.8732 & 0.5547 \\ 0.0382 & 0.5743 & 0.8828 \end{bmatrix} \tag{7.11}$$

The optimum of this problem is given by $f(0.114, 0.556, 0.852) = -3.863$.

A multi-fidelity version for the problem was proposed by Huang et al.[324], where the authors use a LF source given by:

$$y_l(\mathbf{x}) = y_h(\mathbf{x}) + 7.6\,\mathrm{MA3}(\mathbf{x}) \tag{7.12}$$

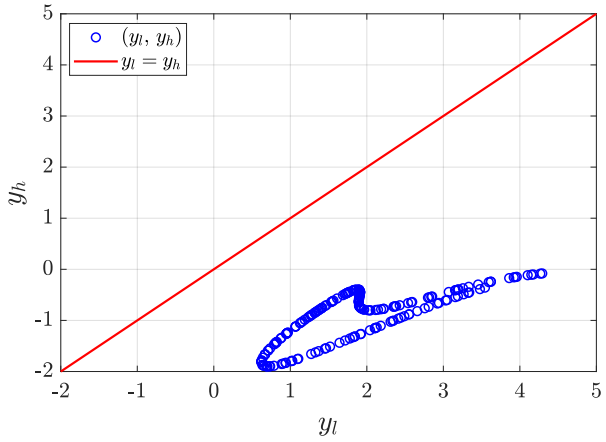
where

$$\mathrm{MA3} = \begin{bmatrix} 1 & \mathbf{x}^T \end{bmatrix} \begin{bmatrix} 0.5850 & -0.1620 & -0.1895 & -0.2155 \\ -0.1620 & 0.2250 & -0.1040 & 0.1630 \\ -0.1895 & -0.1040 & 0.2630 & 0.0965 \\ -0.2155 & 0.1630 & 0.0965 & 0.2740 \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{x} \end{bmatrix} \tag{7.13}$$

where $\mathbf{x}$ is the design variable column vector.

Figure 64 shows the correlation between the responses for these two functions. Again, these two functions present quite a distinct behavior, as the response from the LF source is very different. The correlation factor between these responses is $R^2 = 0.60$.

Figure 64 – Correlation between sources for the Hartmann3 function.



Source: the author

This problem was solved by Zhang, Han and Zhang[73] using the KRG-EI, COKRG-AEI, and HKRG-VF-EI algorithms. Later, Ruan et al.[76] also solved it, this time using the KRG-PI, COKRG-AEI, HKRG-VF-PI, and the HKRG-VF-EI. Finally, Huang et al.[324] also solved it using the COKRG-AEI algorithm, even though the authors considered a different number of initial LF source evaluations ($10m$).

The results found can be seen in Table 13, along with the results obtained by these authors. The best results were obtained by the HKRG-VF-EI, which achieved the lowest NRMSE and the highest success rate, while also performing the lowest amount of HF evaluations. The COKRG-EI also achieved similar NRMSE, even though it needed almost double the amount of

HF evaluations. Here, the advantages of the Variable Fidelity (VF) adaptive sampling are more clear: the HKRG-VF-EI was able to efficiently substitute a number of HF evaluations with LF evaluations, while still finding the optimum design rather quickly. The results found here are also very similar to the ones found by other authors.

Table 13 – Averaged results for the Hartmann3 problem.

Source	SAO algorithm	NRMSE	Success rate	$n_{ev,h}$	$n_{ev,l}$
Present work	KRG-PI	0.05%	50%	42 (33)	-
	KRG-EI	0.05%	60%	35 (26)	-
	COKRG-PI	0.05%	20%	44 (35)	53 (35)
	COKRG-EI	0.01%	60%	34 (25)	43 (25)
	HKRG-VF-PI	0.06%	20%	28 (19)	42 (24)
	HKRG-VF-EI	0.01%	80%	19 (10)	48 (30)
Zhang, Han and Zhang[73]	KRG-EI	0.06%	-	47 (38)	-
	COKRG-AEI	0.05%	-	56 (47)	21 (3)
	HKRG-VF-EI	0.08%	-	21 (12)	32 (14)
Ruan et al.[76]	KRG-PI	0.22%	-	49 (40)	-
	COKRG-AEI	0.10%	-	22 (13)	57 (39)
	HKRG-VF-PI	0.15%	-	27 (18)	43 (25)
	HKRG-VF-EI	0.17%	-	27 (18)	43 (25)
Huang et al.[324]	COKRG-AEI	$< 0.14\%$	-	25 (16)	32 (2)

7.1.3 Ackley5 function

In this section, the minimization of the Ackley function will be performed. Even though this function is not very popular for SBO in general, some authors have optimized it using MFMs[324, 73, 76]. The HF function is given by:

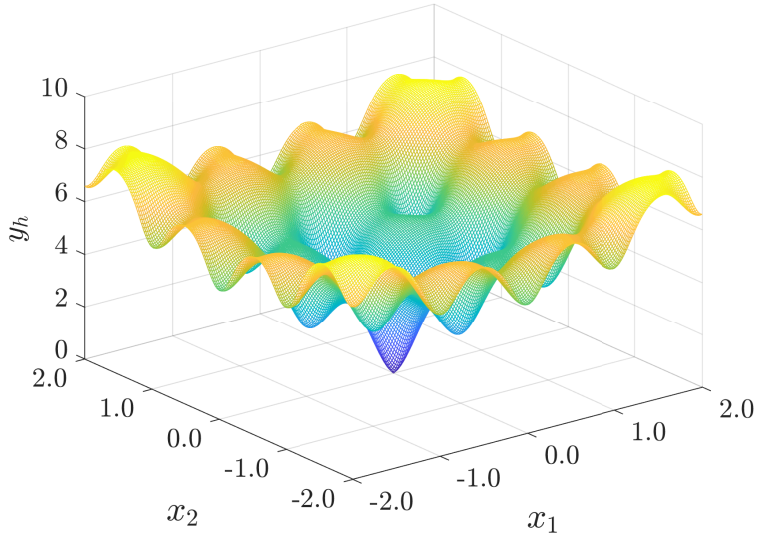
$$y_h(\mathbf{x}) = -a\exp\left[-b\sqrt{\frac{1}{m}\sum_{i=1}^{m}x_i^2}\right] - \exp\left[\frac{1}{m}\sum_{i=1}^{m}\cos(cx_i)\right] + a + \exp(1)\,,\; x_i \in [-2,2] \tag{7.14}$$

where $a = 20$, $b = 0.2$, and $c = 2\pi$. The problem is m-dimensional, and Figure 65 shows its behavior for the two-dimensional case. The problem is highly multimodal, as there are multiple local minima alongside with only one global optima. In this work, the five-dimensional case will be used. The optimum of this problem is given by $f(0,0,0,0,0) = 0.0$.

To derive a LF source, the authors add a quadratic polynomial MA5:

$$y_l(\mathbf{x}) = y_h(\mathbf{x}) + 0.74\,\mathrm{MA5}(\mathbf{x}) \tag{7.15}$$

Figure 65 – Two-dimensional Ackley function.



Source: the author

where

$$\text{MA5} = \begin{bmatrix} 1 & \mathbf{x}^T \end{bmatrix} \begin{bmatrix} 585.0 & -0.635 & -0.565 & -3.315 & -6.450 & -3.055 \\ -0.635 & 0 & 0 & 0 & 2.630 & 5.300 \\ -0.565 & 0 & 0 & 0 & -0.313 & -1.550 \\ -3.315 & 0 & 0 & -0.950 & 0 & 0 \\ -6.450 & 2.630 & -0.313 & 0 & -12.40 & -3.620 \\ -3.055 & 5.300 & -1.550 & 0 & -3.620 & -10.10 \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{x} \end{bmatrix} \times 10^{-3} \tag{7.16}$$

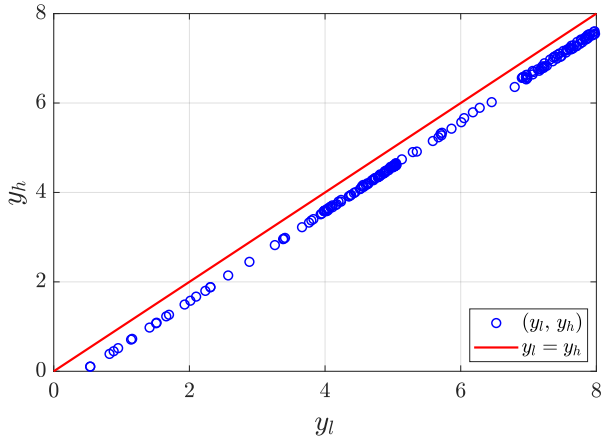
where $\mathbf{x}$ is the design variable column vector.

Figure 66 shows the correlation between responses for these two functions. The behavior is similar to the one seen in Forrester 1b, where the LF source seems to be almost a simple shift in the y-direction. However, the correlation factor is very high ($R^2 = 0.99$).

Similar to the Hartmann3 problem, this problem was solved by Zhang, Han and Zhang[73] (using the KRG-EI, COKRG-AEI, and HKRG-VF-EI), Ruan et al.[76] (using the KRG-PI, COKRG-AEI, HKRG-VF-PI, and HKRG-VF-EI), and Huang et al.[324] (using the COKRG-AEI). Once again, Huang et al.[324] considered a different number of initial LF evaluations ($10m$).

Table 14 shows the results obtained for this problem. Here, since $y_{opt} = 0.0$, the NRMSE is replaced by the absolute difference between the best result found and the optimum. Also, the stopping criterion is replaced by one where the process stops when the absolute

Figure 66 – Correlation between sources for the Ackley5 function.



Source: the author

difference is lower than 0.100. It is interesting to note that, here, the best-performing approach, in terms of absolute difference, was the single-fidelity method, KRG-EI. However, the HKRG-VF-PI and the HKRG-VF-EI managed to find very similar results, and the latter needed a much lower number of HF evaluations. Similar to problem Forrester 1b, the COKRG model behaved very badly for this problem. This model seems to be incapable of providing a good approximation quality when the models are related by an almost constant additive term.

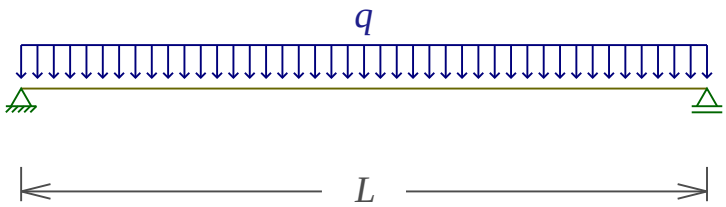
Table 14 – Averaged results for the Ackley5 problem.

Source	SAO algorithm	Absolute difference	$n_{ev,h}$	$n_{ev,l}$
Present work	KRG-PI	0.847	75 (60)	-
	KRG-EI	0.072	55 (40)	-
	COKRG-PI	0.461	66 (51)	81 (51)
	COKRG-EI	0.241	56 (41)	83 (53)
	HKRG-VF-PI	0.090	54 (39)	69 (39)
	HKRG-VF-EI	0.076	35 (20)	73 (43)
Zhang, Han and Zhang[73]	KRG-EI	0.100	86 (71)	-
	COKRG-AEI	0.100	51 (36)	159 (129)
	HKRG-VF-EI	0.100	35 (20)	162 (132)
Ruan et al.[76]	KRG-PI	0.160	88 (73)	-
	COKRG-AEI	0.100	31 (16)	83 (53)
	HKRG-VF-PI	0.152	34 (19)	89 (59)
	HKRG-VF-EI	0.148	36 (21)	95 (65)
Huang et al.[324]	COKRG-AEI	1.776	25 (10)	73 (23)

7.1.4 Optimization of a simply supported beam

This section presents the first engineering benchmark problem, which is related to the optimization of a simply supported beam shown in Figure 67. Here, we consider that the beam length is $L = 5$ m, the distributed load is $q = 10$ kN/m, the Young modulus is $E = 21662$ kPa, and the shear modulus is $G = 8528.35$ kPa.

Figure 67 – Simply supported beam.



Source: the author

In this problem, the aim is to minimize the maximum displacement of the beam, considering the width b and height h of the cross section as design variables. Two constraints will be considered, related to a maximum cross-sectional area $A_{max} = 0.8$ and a restriction on the cross-section aspect-ratio ($h \leq 3b$). Since these constraints are naturally cheap to evaluate, they will be evaluated exactly, as there is no need to build an additional surrogate model to estimate their values. The optimization problem may be formulated as:

$$\begin{cases} \text{find} & \mathbf{x} = \{b, h\} \\ \text{that minimizes} & w(\mathbf{x}) \\ \text{subject to} & g_1(\mathbf{x}) = \dfrac{A}{A_{max}} - 1 \leq 0 \\ & g_2(\mathbf{x}) = \dfrac{h}{3b} - 1 \leq 0 \\ \text{with} & 0.4 \text{ m} \leq b \leq 1.0 \text{ m} \\ & 1.0 \text{ m} \leq h \leq 2.0 \text{ m} \end{cases} \tag{7.17}$$

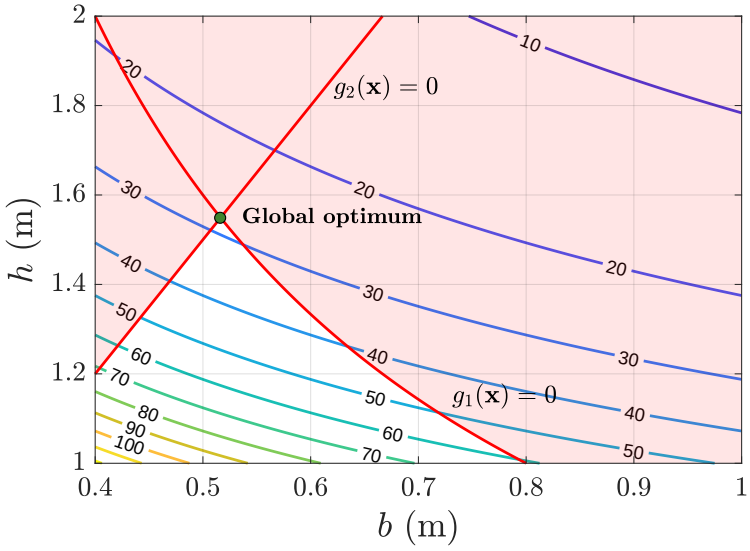
Considering shear deformations, and assuming that the cross-section may not be perpendicular to the neutral axis after deformation, the beam maximum displacement may be obtained using the Timoshenko theory:

$$w_T(\mathbf{x}) = \frac{5qL^4}{384EI} + \frac{qL^2}{8GA} \tag{7.18}$$

Using this function to evaluate the maximum displacement, the optimum is found at $w_T(0.516; 1.549) = 28.1$ mm. Figure 68 shows the constrained space for this problem. Even though this problem is

seemingly simple, the constraints make it so that it is not trivial to find the feasible region with few evaluations.

Figure 68 – Constrained space for the simply supported beam problem.



Source: the author

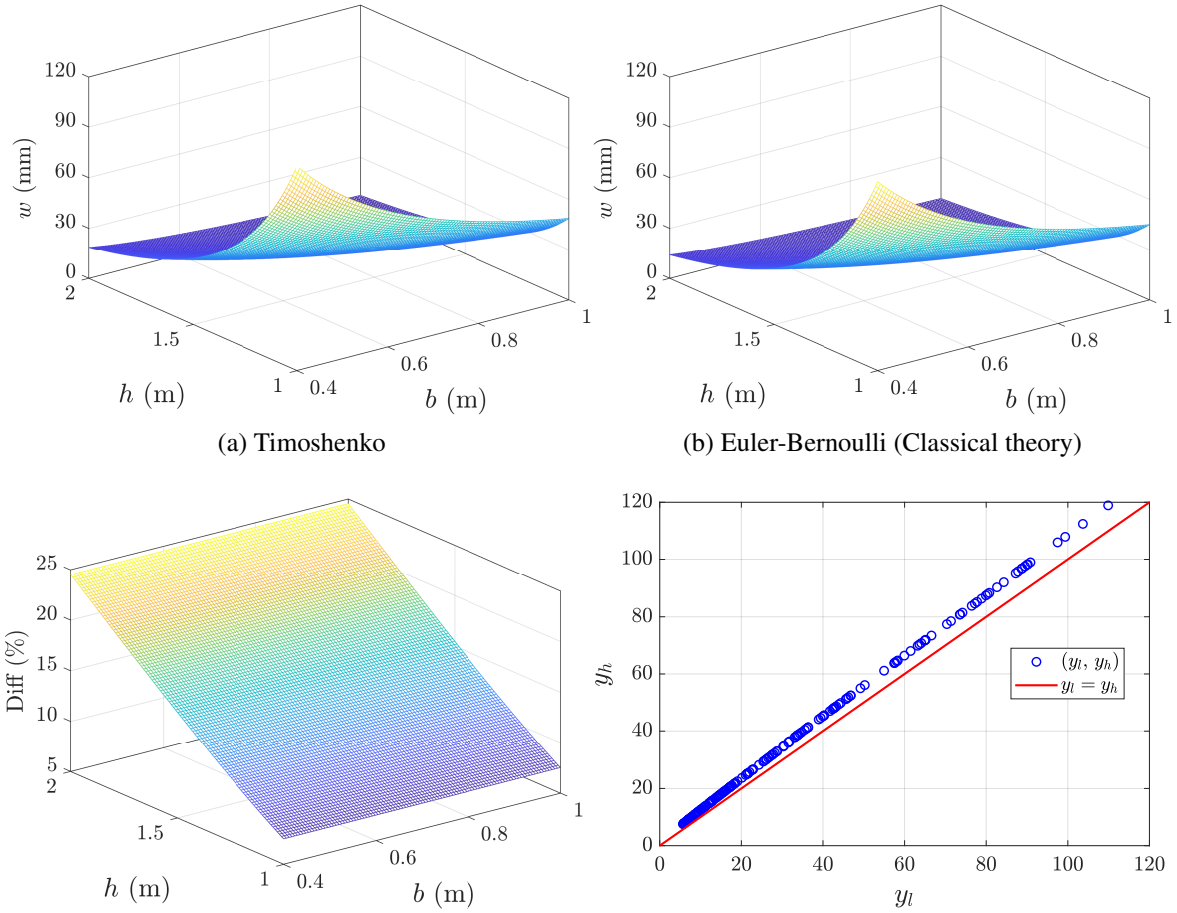
Instead of the Timoshenko solution, we may compute the displacement in a simplified way by the Euler-Bernoulli theory, known as the classical beam theory, which disregards shear deformations. Here, the maximum displacement is given by:

$$w_E(\mathbf{x}) = \frac{5qL^4}{384EI} \tag{7.19}$$

Figure 69 shows how the displacement changes in the design space by each theory. The results obtained by both approaches seem to be very similar. However, when we analyze the relative difference between the two functions, we see that it may be close to 40% for higher h. Indeed, the difference between these two approaches should increase as h/L increases. At the optimum point, the relative difference is 24.23%. Even so, the correlation factor between these two approaches is very high ($R^2 = 0.99$).

In this problem, the initial sample is defined using $n_h = 1.5m = 3$ and $n_l = 3m = 6$. Table 15 shows the results obtained using a variety of approaches. With the exception of the COKRG-WEI, all COKRG-based approaches performed very poorly. Also, no PI-based approach was able to improve upon the single-fidelity version (KRG-PI). Once again, HKRG behaved exceptionally well, especially using VF-LCB and VF-WEI, where, on average, only one extra HF evaluation was required to find the global optima. It is worth remembering that none of these approaches have been proposed in the literature, being variable-fidelity extensions of the LCB and the WEI, respectively.

Figure 69 – Assessment of the displacement of the simply supported beam via different sources.



(a) Timoshenko

(b) Euler-Bernoulli (Classical theory)

(c) Relative difference

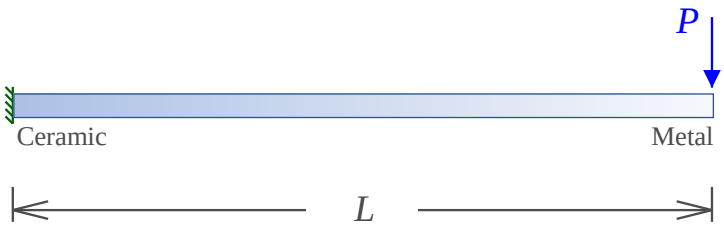
(d) Correlation between sources

Source: the author

7.1.5 Optimization of a Functionally Graded beam

Finally, in this problem, the optimization of a FG clamped beam is performed, which is shown in Figure 70. In this problem, $L = 10$ m and $P = 40$ kN.

Figure 70 – Functionally Graded (FG) clamped beam.



Source: the author

The beam has a circular cross-section of radius r, and the material gradation is given

Table 15 – Averaged results for the simply supported beam problem.

Model	Infill criterion	w_{opt} (mm)	NRMSE	$n_{ev,h}$	$n_{ev,l}$
KRG	LCB	28.10	0.00%	6 (3)	-
	PI	28.31	0.76%	12 (9)	-
	EI	28.10	0.00%	8 (5)	-
	WEI	28.10	0.00%	7 (4)	-
Co-KRG	LCB	31.73	12.93%	7 (4)	10 (4)
	PI	41.42	47.40%	14 (11)	17 (11)
	EI	28.48	1.37%	7 (4)	10 (4)
	WEI	28.10	0.00%	7 (4)	10 (4)
H-KRG	LCB	28.10	0.00%	4 (1)	7 (1)
	PI	28.47	1.33%	18 (15)	21 (15)
	EI	28.10	0.00%	6 (3)	9 (3)
	WEI	28.10	0.00%	6 (3)	9 (3)
	VF-LCB	28.10	0.00%	4 (1)	8 (2)
	VF-PI	28.44	1.19%	16 (13)	20 (14)
	VF-EI	28.10	0.00%	5 (2)	8 (2)
	VF-WEI	28.10	0.00%	4 (1)	7 (1)

by:

$$V_m = \left(\frac{x}{L}\right)^N \text{ and } V_c = 1 - V_m \tag{7.20}$$

where V_m and V_c refer to the metal and ceramic volume fractions, respectively, and N is a parameter which defines the material gradation. Equivalent material properties are evaluated via the rule of mixtures:

$$E(\mathbf{x}) = V_m(\mathbf{x})\,E_m + V_c(\mathbf{x})\,E_c \tag{7.21}$$

where $E_m = 90$ GPa and $E_c = 380$ GPa are the metal and ceramic Young's moduli, respectively.

Once again, the aim is to minimize the maximum displacement of the beam, considering r and N as design variables. Two constraints will be considered, now related to a maximum cost $C_{max} = 50$ M.U. (Monetary Units), and a minimum percentage of ceramic in the beam $\overline{V}_{c,min} = 50\%$. Again, these constraints are cheap to evaluate, and they will be evaluated exactly.

The optimization problem may be formulated as:

$$\begin{cases} \text{find} & \mathbf{x} = \{r, N\} \\ \text{that minimizes} & w(\mathbf{x}) \\ \text{subject to} & g_1(\mathbf{x}) = 1 - \dfrac{\overline{V}_c}{\overline{V}_{c,min}} \leq 0 \\ & g_2(\mathbf{x}) = \dfrac{C}{C_{max}} - 1 \leq 0 \\ \text{with} & 0.20 \text{ m} \leq r \leq 1.00 \text{ m} \\ & 0.00 \leq N \leq 10.0 \end{cases} \tag{7.22}$$

where $\overline{V}_c$ is the percentage of ceramic in the beam, evaluated as:

$$\overline{V}_c = \frac{1}{L}\int_L V_c \, dx = \frac{N}{N+1} \tag{7.23}$$

and C is the total cost of the beam, evaluated as:

$$C = V\left[(1-\overline{V}_c)C_M + \overline{V}_c C_C\right] \tag{7.24}$$

where V is the beam volume, and $C_M = 1$ M.U./m^3 and $C_C = 5$ M.U./m^3 are the cost of the metal and the ceramic.

Using the unit load method and considering the Euler-Bernoulli theory, the tip displacement of the beam can be evaluated as:

$$w(\mathbf{x}) = \int_0^L \frac{M(\mathbf{x})\overline{M}(\mathbf{x})}{E(\mathbf{x})I}\, dx = \int_0^L \frac{P(L-x)^2}{E(\mathbf{x})I}\, dx \tag{7.25}$$

This integral can be solved numerically using a 10-point Gauss quadrature rule. Using this function, the optimum of the problem is found at $w(0.728, 1.0) = 0.1025$ mm. Figure 71 shows the constrained space for this problem.

Alternatively, we may use the Rayleigh-Ritz method to find an approximation for the beam displacement. The weak form of the problem is given by:

$$\Pi = \frac{1}{2}\int_0^L EI\, w_{,xx}\, dx + P\, w(L) \tag{7.26}$$

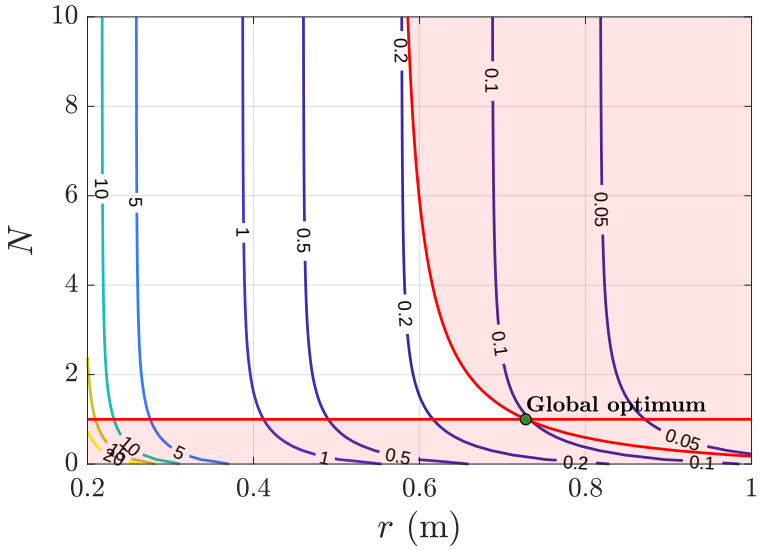
Assuming a displacement described by:

$$w(x) = ax^2 \tag{7.27}$$

we may find a considering the principle of stationary total potential energy:

$$\frac{\partial \Pi}{\partial a} = 0 \quad \Rightarrow \quad a = \frac{PL(N+1)}{4I(E_c N + E_m)} \tag{7.28}$$

Figure 71 – Constrained space for the FG beam problem.



Source: the author

Thus, the displacement may be estimated by:

$$w(x) = \frac{PL(N+1)}{4I(E_c N + E_m)} x^2 \tag{7.29}$$

and the maximum beam displacement (at $x = L$) is given by:

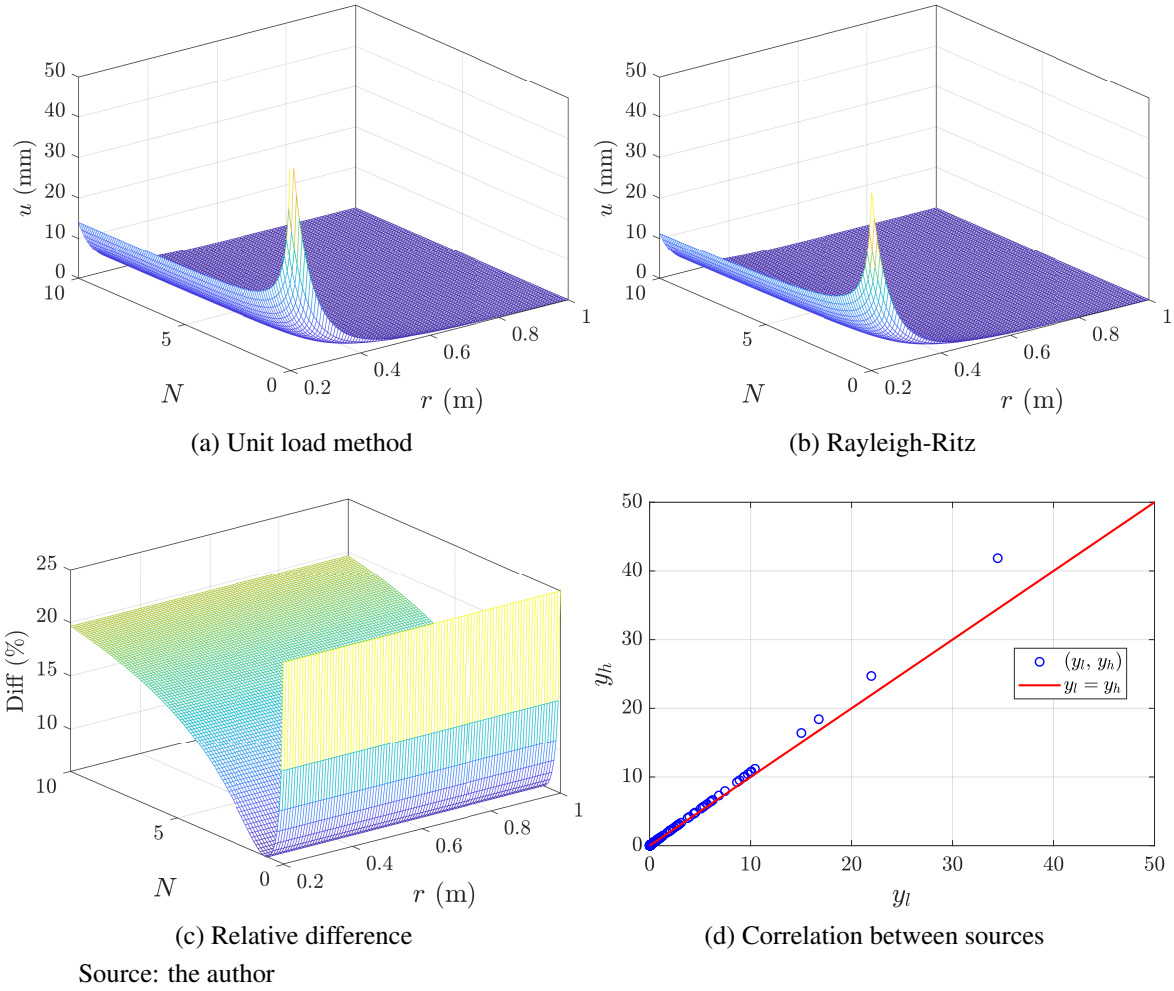
$$w(x) = \frac{PL^3(N+1)}{4I(E_c N + E_m)} \tag{7.30}$$

This function can be used as a LF function for the beam displacement.

Figure 72 shows how the displacement changes in the design space according to each source. The unit load method presents slightly higher displacements, and the relative difference between the two approaches is close to 20% for higher N. In the optimum, the difference is 6.05%. Once again, the correlation factor between these two approaches is very high ($R^2 = 0.99$).

The initial sample for this problem is defined using $n_h = 1.5m = 3$ and $n_l = 3m = 6$. Table 16 shows the results obtained using each SAO approach. This example clearly shows that the use of Multi-Fidelity Models (MFMs) is no panacea for the improvement of the optimization process. It is easy to see that the single-fidelity KRG-based approaches performed better than most MFMs. Even so, when Variable Fidelity (VF) approaches are used, the HKRG model excels at finding the optimum using few HF evaluations. The best approach was the VF-LCB, where, on average, only one extra HF evaluation was necessary to find the optimum. Moreover, no PI-based approach was able to achieve good results for this problem.

Figure 72 – Assessment of the displacement of the FG beam via different sources.



(a) Unit load method

(b) Rayleigh-Ritz

(c) Relative difference

(d) Correlation between sources

Source: the author

Table 16 – Averaged results for the FG beam problem.

Model	Infill criterion	w_{opt} (mm)	NRMSE	$n_{ev,h}$	$n_{ev,l}$
KRG	LCB	0.1025	0.00%	5 (2)	-
	PI	0.1033	0.84%	32 (29)	-
	EI	0.1025	0.00%	6 (3)	-
	WEI	0.1025	0.00%	6 (3)	-
COKRG	LCB	0.1025	0.00%	7 (4)	10 (4)
	PI	0.1030	0.51%	22 (19)	25 (19)
	EI	0.1027	0.21%	10 (7)	13 (7)
	WEI	0.1025	0.00%	6 (3)	9 (3)
HKRG	LCB	0.1025	0.00%	8 (5)	11 (5)
	PI	0.1037	1.26%	28 (25)	31 (25)
	EI	0.1027	0.25%	13 (10)	16 (10)
	WEI	0.1025	0.02%	10 (7)	13 (7)
	VF-LCB	0.1025	0.00%	4 (1)	11 (5)
	VF-PI	0.1034	0.95%	30 (27)	36 (30)
	VF-EI	0.1025	0.00%	5 (2)	11 (5)
	VF-WEI	0.1025	0.00%	6 (3)	14 (8)

7.2 Numerical FGM problems

In this section, problems related to the optimization of FG plates and shells will be solved. These structures often require numerical approaches to evaluate the structural response. In this work, Isogeometric Analysis (IGA) will be employed to evaluate quantities of interest through both solid continuum and shallow shell theories. Unless stated otherwise, the initial sample will be defined via the LHS_N method, with $N = 20$. The number of HF and LF initial sampling points will depend on the cost of each source and on the problem's dimensionality. Since time spent is a major concern for these problems, it will also be used to measure the efficiency of each SAO approach, along with the number of HF and LF evaluations. Material properties found in this Section can be found in Table 1.

7.2.1 Maximization of the buckling load of a unidirectional FG plate

The first example deals with the maximization of the buckling load factor of a simply supported square plate made of Aluminum (Al) as metal and Alumina (Al_2O_3) as ceramic, with $a/h = 10$, subject to a unidirectional uniform compressive load N_x. Equivalent material properties are estimated via the Mori-Tanaka model and the material gradation is defined by 9 control points through the thickness, symmetric with respect to the mid-plane. Due to symmetry, there are only 5 design variables. A constraint will be used to limit the maximum percentage of ceramic material $\overline{V}_{c,max}$. Thus, the optimization problem can be formulated as:

$$\begin{cases} \text{find} & \mathbf{x} = \{x_1, x_2, \ldots, x_5\} \\ \text{that minimizes} & -\lambda_{norm}(\mathbf{x}) \\ \text{subject to} & g_1(\mathbf{x}) = \dfrac{\overline{V}_c(\mathbf{x})}{\overline{V}_{c,max}} - 1 \leq 0 \\ \text{with} & 0 \leq x_i \leq 1 \end{cases} \tag{7.31}$$

where λ_{norm} is evaluated by Eq. (3.54) and:

$$\overline{V}_c = \frac{1}{h} \int_{-h/2}^{h/2} V_c \, dz \tag{7.32}$$

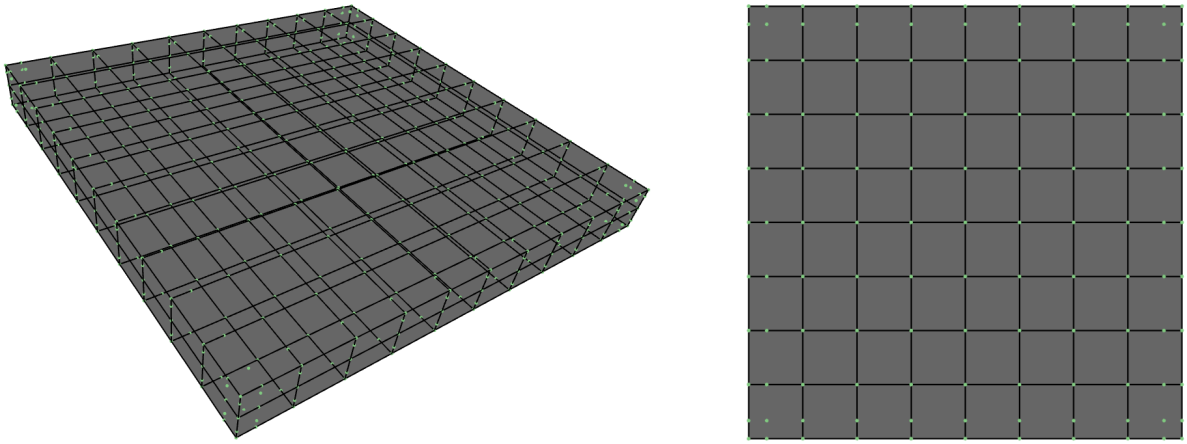
This integral can be evaluated numerically, using a 10-point Gaussian quadrature rule.

This problem was first proposed by Do, Lee and Lee[1], where a DNN was used as a surrogate model to improve the efficiency of the process. The authors used 10,000 sampling points to train and validate the model, and then optimized the problem considering three different $\overline{V}_{c,max}$: 35%, 50%, and 65%. Structural analyses were performed considering the HSDT. Later,

Ribeiro et al.[2] used a single-fidelity RBF-based SAO to optimize the problem, showing more accurate results while only requiring around 50 sampling points. This time, the authors used the FSDT to carry out the structural analyses.

In this work, the HF source will be represented by a 3D model using a $10 \times 10 \times 2$ cubic NURBS mesh, while the LF source will be represented by a FSDT model using a 8×8 cubic NURBS mesh. These two meshes are depicted in Figure 73.

Figure 73 – Mesh used for each source for the maximization of the buckling load of a FG plate.



(a) HF source, $10 \times 10 \times 2$ cubic NURBS mesh
Source: the author

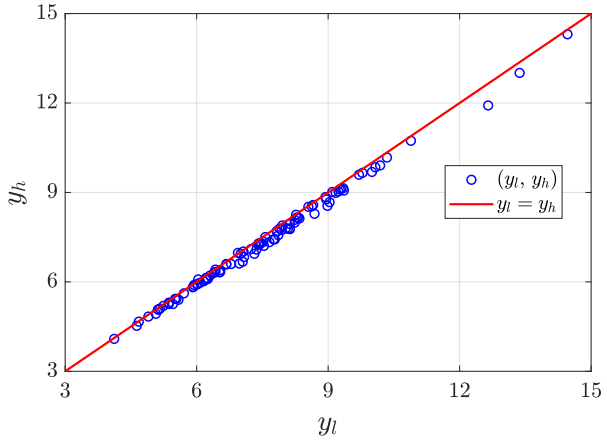
(b) LF source, 8×8 cubic NURBS mesh

Figure 74 shows the correlation between responses for these two cases. Here, the correlation factor is very high ($R^2 = 0.99$), and the average relative difference between sources is close to 2.27%. Each HF evaluation takes, on average, 28.75 s, while each LF evaluation takes only 0.64 s. This means that the relative expense is $C_r = 0.022$. The 3D model offers slightly lower critical buckling load factors than those obtained via the 2D model. Thus, the LF source, although cheaper, may lead to less conservative designs when compared to the HF source.

Finally, this problem will be solved using single and multi-fidelity SAO approaches. Due to the cost of each HF evaluation, for the single-fidelity case (KRG model), only 5 initial sampling points will be used. For the MFMs, some points will be replaced with a number of LF sampling points: the initial sample will have 3 HF and 20 LF data points.

In this work, similar to other researchers, three different versions for the problem will be considered, with three different $\overline{V}_{c,max}$: 35%, 50%, and 65%. Table 17 shows the optimum design for each case. Figure 75 shows how the optimum design ceramic volume fraction changes along the structure thickness for each case. The optimum buckling load factor found by Do, Lee and Lee[1] and by Ribeiro et al.[2] can be found in Table 3. It is worth mentioning that these

Figure 74 – Correlation between sources for the optimization of a unidirectional FG plate.



Source: the author

works used different plate theories to evaluate the buckling load, while the present work uses a 3D model as the HF source.

Table 17 – Optimum design for different $\overline{V}_{c,max}$ [1, 2]

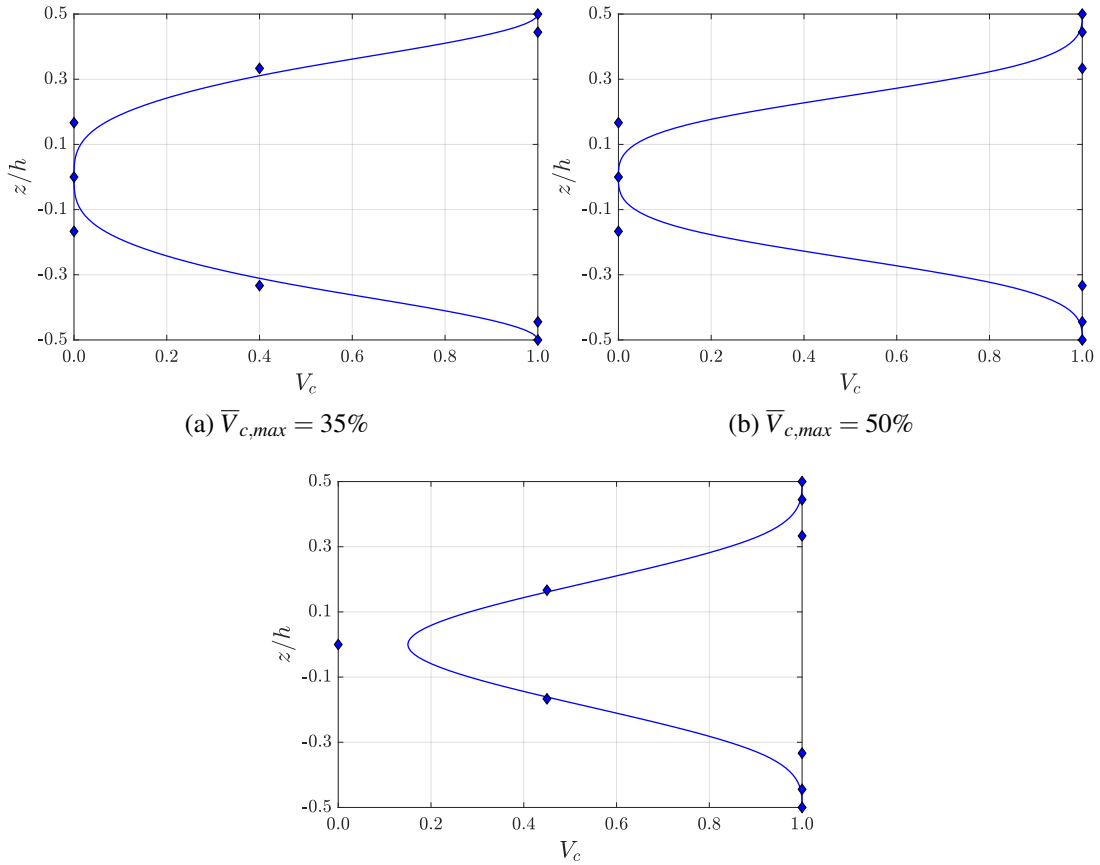
$\overline{V}_{c,max}$	Design variables					λ_{norm}
	x_1	x_2	x_3	x_4	x_5	
35%	1.00	1.00	0.40	0.00	0.00	9.893
50%	1.00	1.00	1.00	0.00	0.00	12.58
65%	1.00	1.00	1.00	0.45	0.00	14.09

Table 18 shows the results found for $\overline{V}_{c,max} = 35\%$. The results show that almost all approaches were able to accurately find the optimum. However, PI-based approaches showed a higher NRMSE in all cases. The KRG model required more HF evaluations to find the optimum, as at least 20 HF evaluations were performed in all cases. For the COKRG and the HKRG, however, most cases managed to achieve a stopping criteria after only 15 HF evaluations. This difference, although seemingly negligible, is able to reduce the time spent in the multi-fidelity approaches in about 25%.

For the VF approaches, it is important to note that the VF-LCB and the VF-PI were not able to efficiently determine the need to perform a HF evaluation, as the number of HF points was very close to the other methods. For the VF-EI and the VF-WEI, however, there was a major decrease in the number of HF evaluations performed. The time spent was up to 65% faster in these cases, in comparison to the results from the KRG model.

It is important to note that, while MFMs require less HF evaluations, they also present a higher complexity, and model building and evaluation are way more expensive. Figure

Figure 75 – Optimum design for each case for the optimization of a unidirectional FG plate.



(a) $\overline{V}_{c,max} = 35\%$ (b) $\overline{V}_{c,max} = 50\%$

(c) $\overline{V}_{c,max} = 65\%$

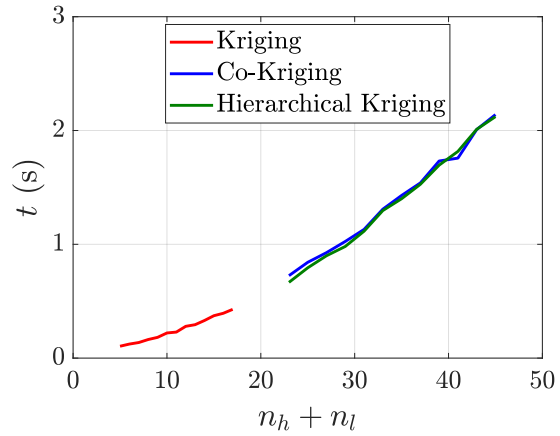
Source: the author

Table 18 – Averaged results for the maximization of the buckling load of a unidirectional FG plate considering $\overline{V}_{c,max} = 35\%$.

Model	Infill criterion	λ_{norm}	NRMSE	$n_{ev,h}$	$n_{ev,l}$	Time spent (s)
KRG	LCB	9.893	0.00%	22 (17)	-	625
	PI	9.823	0.70%	27 (22)	-	772
	EI	9.893	0.00%	21 (16)	-	607
	WEI	9.893	0.00%	20 (15)	-	586
COKRG	LCB	9.893	0.00%	15 (12)	32 (12)	480
	PI	9.814	0.80%	21 (18)	38 (18)	639
	EI	9.893	0.00%	14 (11)	31 (11)	436
	WEI	9.893	0.00%	14 (11)	31 (11)	440
HKRG	LCB	9.893	0.00%	14 (11)	31 (11)	443
	PI	9.782	1.12%	23 (20)	40 (20)	729
	EI	9.893	0.00%	14 (11)	31 (11)	436
	WEI	9.893	0.00%	14 (11)	31 (11)	430
	VF-LCB	9.893	0.00%	14 (11)	32 (12)	436
	VF-PI	9.761	1.34%	29 (26)	46 (26)	927
	VF-EI	9.892	0.01%	7 (4)	31 (11)	225
	VF-WEI	9.893	0.00%	6 (3)	31 (11)	194

76 shows how the cost of the building phase change for different models, depending on the sample size ($n_h + n_l$). The building phase cost seems to be heavily influenced by the sample size, which is in agreement with what is expected. It is worth to remember that time complexity for building these models is proportional to the cube of the number of sampling points.

Figure 76 – Time spent for the building phase for each model for the maximization of the buckling load of a unidirectional FG plate.



Source: the author

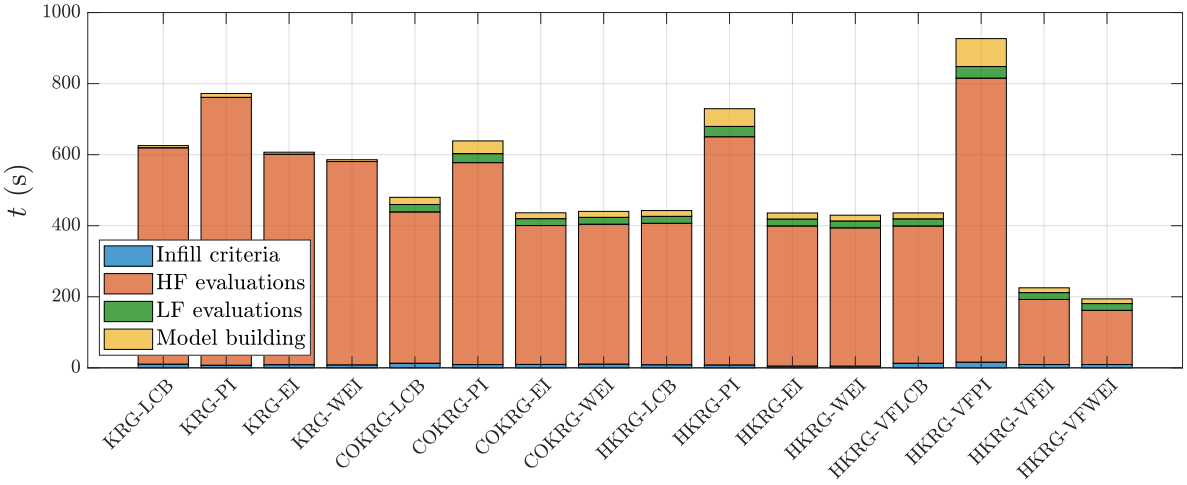
As MFMs present a higher sample size due to the incorporation of LF sampling points, they also present a more expensive model building. This means that, when choosing the number of LF sampling points, regardless of how cheap these evaluations are, the user should not introduce an overwhelmingly large amount of points, as these might make the model too complex.

That being said, the most expensive phase of the process is still the evaluation of data points. Figure 77 shows how the time spent is divided between different phases for SAO: model building, HF and LF data points evaluation, and definition of the new data via the infill criterion. Indeed, the optimization process is usually more efficient when MFMs are employed, especially when the HKRG model is used with the VF-EI or VF-WEI approaches. However, this does not occur when PI-based approaches are employed. For this problem, these approaches are both less accurate and less efficient than the other methods.

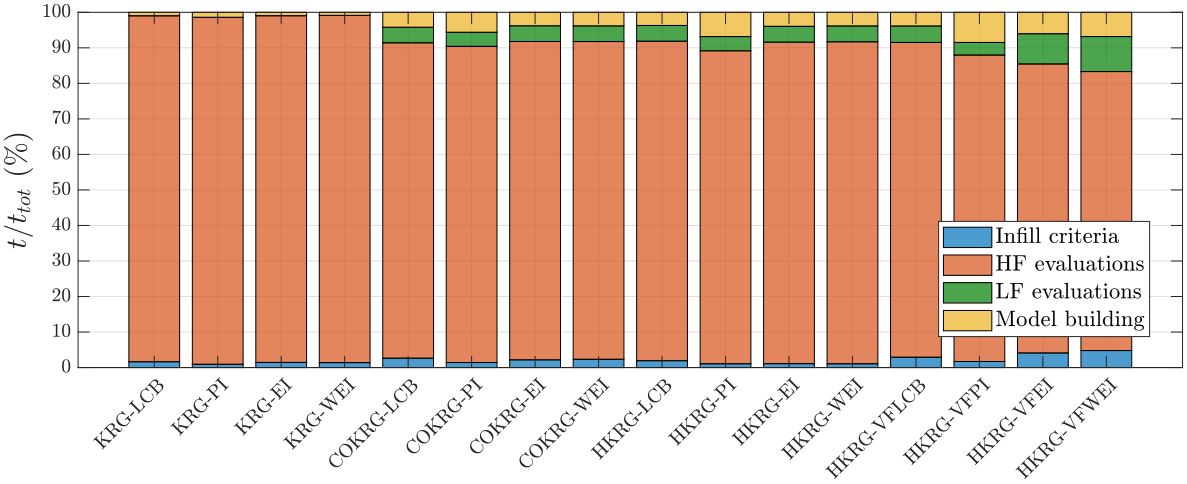
It is also interesting to discuss the relative contribution of each phase. For the Kriging model, the data points evaluation corresponds to roughly 97.5% of the total cost, as all analysis are performed using the High-Fidelity (HF) source. Furthermore, as the total number of sampling points is lower, model building and evaluation are cheaper.

On MFMs, however, model building and, to a lower degree, model evaluation are

Figure 77 – Cost of each phase of the process for the maximization of the buckling load of a unidirectional FG plate.



(a) Total time spent in each phase



(b) Percentage contribution relative to the total cost of each method

Source: the author

much more expensive. Even so, the evaluation of HF data points is still the most expensive part of the process, being responsible to at least 80% of the total cost of the process. This also helps to understand why the VF-EI and VF-WEI approaches are so much cheaper than the others: as the evaluation of data points is, by far, the most expensive phase, the total cost of the process can be greatly decreased by reducing the number of HF evaluations required.

Table 19 shows the results for $\overline{V}_{c,max} = 50\%$. Overall, similar results were found. Once again, MFMs were able to show a gain in efficiency when compared to the KRG model, where VF-EI and VF-WEI showed an even greater reduction in the time spent, as they only required 4 and 3 extra evaluations on average, respectively. Also, PI-based approaches struggled to find the optimum. Very similar conclusions can be made for the results considering $\overline{V}_{c,max} = 65\%$, which can be found in Table 20.

Table 19 – Averaged results for the maximization of the buckling load of a unidirectional FG plate considering $\overline{V}_{c,max} = 50\%$.

Model	Infill criterion	λ_{norm}	NRMSE	$n_{ev,h}$	$n_{ev,l}$	Time spent (s)
KRG	LCB	12.576	0.00%	21 (16)	-	593
	PI	12.443	1.06%	28 (23)	-	819
	EI	12.576	0.00%	20 (15)	-	578
	WEI	12.576	0.00%	22 (17)	-	637
COKRG	LCB	12.576	0.00%	14 (11)	31 (11)	439
	PI	12.491	0.68%	25 (22)	42 (22)	772
	EI	12.576	0.00%	14 (11)	31 (11)	430
	WEI	12.576	0.00%	14 (11)	31 (11)	429
HKRG	LCB	12.576	0.00%	14 (11)	31 (11)	429
	PI	12.433	1.14%	23 (20)	40 (20)	700
	EI	12.576	0.00%	14 (11)	31 (11)	423
	WEI	12.576	0.00%	14 (11)	31 (11)	423
	VF-LCB	12.576	0.00%	14 (11)	32 (12)	439
	VF-PI	12.415	1.28%	26 (23)	43 (23)	798
	VF-EI	12.576	0.00%	7 (4)	31 (11)	236
	VF-WEI	12.576	0.00%	6 (3)	31 (11)	197

Table 20 – Averaged results for the maximization of the buckling load of a unidirectional FG plate considering $\overline{V}_{c,max} = 65\%$.

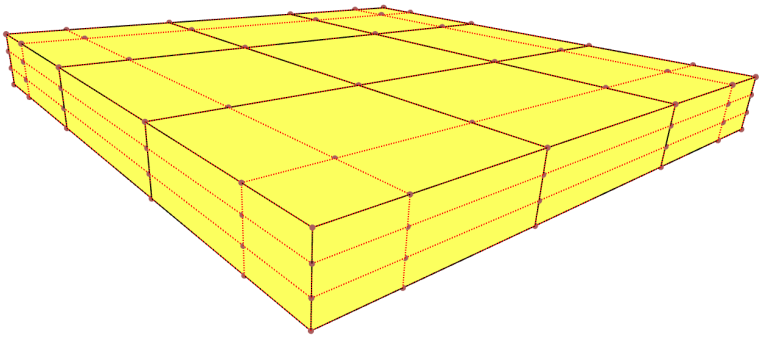
Model	Infill criterion	λ_{norm}	NRMSE	$n_{ev,h}$	$n_{ev,l}$	Time spent (s)
KRG	LCB	14.091	0.00%	22 (17)	-	642
	PI	14.043	0.34%	30 (25)	-	875
	EI	14.091	0.00%	22 (17)	-	626
	WEI	14.091	0.00%	22 (17)	-	625
COKRG	LCB	14.091	0.00%	15 (12)	32 (12)	475
	PI	14.046	0.32%	23 (20)	40 (20)	733
	EI	14.091	0.00%	14 (11)	31 (11)	446
	WEI	14.091	0.00%	14 (11)	31 (11)	430
HKRG	LCB	14.033	0.41%	14 (11)	31 (11)	431
	PI	14.040	0.36%	29 (26)	46 (26)	905
	EI	14.091	0.00%	14 (11)	31 (11)	423
	WEI	14.091	0.00%	14 (11)	31 (11)	423
	VF-LCB	14.091	0.00%	14 (11)	32 (12)	440
	VF-PI	13.979	0.79%	29 (26)	47 (27)	929
	VF-EI	14.091	0.00%	7 (4)	31 (11)	222
	VF-WEI	14.091	0.00%	6 (3)	31 (11)	217

7.2.2 Maximization of the buckling load of a tridirectional FG plate

This example deals with the maximization of the buckling load factor of a simply supported square plate made of stainless steel (SUS304) as metal and silicon nitride (Si_3N_4) as ceramic. Once again, the plate is subject to a unidirectional uniform compressive load N_x and its

equivalent material properties are estimated via the Mori-Tanaka model. This time, however, material gradation is defined by a trivariate B-Spline function with 144 control points. The design mesh is shown in Figure 78.

Figure 78 – Design mesh for the optimization of a tridirecional FG plate.



Source: the author

The design variables are, once again, the ceramic volume fraction in the control points. However, we consider that the volume fraction distribution is symmetric at the center of the plate in all three dimensions. Thus, there are only 18 design variables. Similar to the previous example, a limitation on the maximum percentage of ceramic material will be imposed. This way, the optimization problem can be described as:

$$\begin{cases} \text{find} & \mathbf{x} = \{x_1, x_2, \ldots, x_{18}\} \\ \text{that minimizes} & -\lambda_{norm}(\mathbf{x}) \\ \text{subject to} & g_1(\mathbf{x}) = \dfrac{\overline{V}_c(\mathbf{x})}{\overline{V}_{c,max}} - 1 \leq 0 \\ \text{with} & 0 \leq x_i \leq 1 \end{cases} \tag{7.33}$$

where λ_{norm} is evaluated by Eq. (3.53) and:

$$\overline{V}_c = \frac{1}{V} \int_V V_c \, dv \tag{7.34}$$

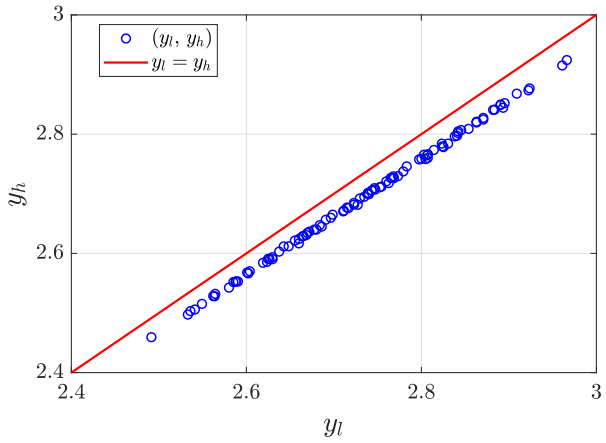
Again, this integral can be evaluated numerically, using an adequate Gaussian quadrature rule.

It is important to note that this problem is much more complex than the previous one, especially due to the higher dimensionality. Do, Nguyen-Xuan and Lee[18] employed a DNN to fit this problem, using 10,000 sampling points to train and validate the model. The authors show that it took 189,983 s to evaluate all sampling points and 656 s to train the model. Then, the authors conducted the optimization process considering the HSDT and using three different $\overline{V}_{c,max}$: 30%, 50%, and 70%. On average, it took 21 s to perform the optimization process on the static single-fidelity model. Thus, for one optimization process, it took, on average, 190,660 s,

or roughly 53 hours, where 99.6% of this time was spent evaluating all data points. Note that a very high number of data points is required in this case because the authors used a single-fidelity model to perform the optimization process. For a reliable optimization process, the authors need to make sure that the model has a sufficiently good global accuracy. That being said, the DNN still struggled to find the global optimum for this problem, as, for $\overline{V}_{c,max} = 50\%$ and $\overline{V}_{c,max} = 70\%$, the optimum found actually violated the ceramic fraction constraint.

In this work, optimization was performed using single and multi-fidelity models, using the sources presented in Figure 73. Thus, for the HF source, a 3D solid continuum mesh was used, while a 2D mesh (considering the FSDT) was used for the LF source. Figure 79 shows the correlation between these two sources. Once again, the correlation is very high ($R^2 = 0.99$), even though, this time, there seems to be an almost constant absolute difference between the response found. This time, each HF evaluation takes around 28.85 s, while the LF source takes only 0.69 s, which means that the relative expense is $C_r = 0.024$.

Figure 79 – Correlation between sources for the optimization of a tridirecional FG plate.



Source: the author

Table 21 shows the optimum designs obtained in this work, and compares them to the ones found by Do, Nguyen-Xuan and Lee[18]. Note that designs obtained by our algorithm are better than the ones of Do, Nguyen-Xuan and Lee[18]. Here, SS1 and SS2 refer to the different boundary conditions for simply supported plates, as shown in Figure 10. The SS2 condition was adopted for the optimization process. For reference, the response found using the SS1 condition is also shown.

Figure 80 shows the optimum design for $\overline{V}_{c,max} = 30\%$. Note that, in the loaded faces, there is more ceramic than the non-loaded ones. Also, there is more ceramic near the center of the plate, where displacements are higher. In the inner side of the plate there is gradually less

Table 21 – Optimum design for different $\overline{V}_{c,max}$.

$\overline{V}_{c,max}$		30%		50%		70%	
Control points		Ref.[18]	This work	Ref.[18]	This work	Ref.[18]	This work
1		1	1	1	1	1	1
2		1	0	1	0	1	1
3		0	0	0.548	1	1	1
4		1	1	1	1	1	1
5		1	0.318	1	1	1	1
6		0	1	1	1	1	1
7		0	1	0.998	1	1	1
8		0.526	0	1	1	1	1
9		1	1	1	1	1	1
10		0	0	1	1	1	1
11		0	0	0	0	1	0
12		0	0	0	0	0	0
13		0	0	0	0	1	1
14		0	0	0	0	0.242	0
15		0	0	0	0	0	1
16		0	0	0	0.615	0	1
17		0	0	0	0	0	0
18		0	0	0	0	1	0.408
λ_{norm}	2D, HSDT, SS1[18]	2.906	-	3.241	-	3.506	-
	2D, FSDT, SS1	2.832	2.869	3.244	3.256	3.462	3.504
	2D, FSDT, SS2	2.610	2.635	2.975	2.984	3.171	3.207
	3D	2.564	2.587	2.911	2.923	3.111	3.147
$\overline{V}_c$		29.86%	30.00%	50.28%	50.00%	71.29%	70.00%

ceramic than in the outer side.

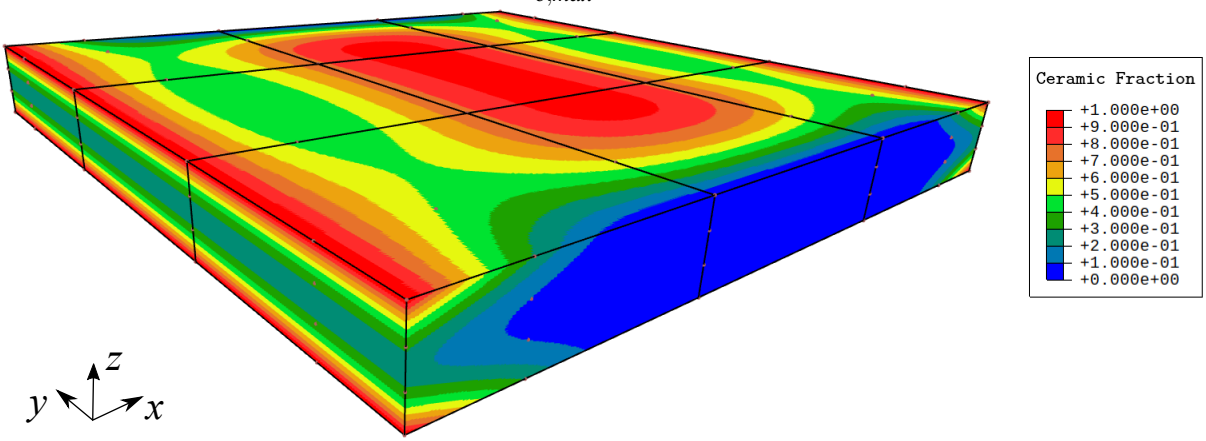
Figure 81 now shows the optimum design for $\overline{V}_{c,max} = 50\%$. Aside from the fact that there is more ceramic in the plate, similar observations can be brought up. Here it is more clear that the outer side of the plate has gradually more ceramic than the inner side.

Finally, Figure 82 shows the optimum design for $\overline{V}_{c,max} = 70\%$. Here, almost the entire outer side of the plate is covered in ceramic. The exception is the non-loaded face. However, there is a gradual reduction as one gets closer to the internal side of the plate.

For single-fidelity approaches (using the Kriging model), 18 initial sampling points were used, while, for MFMs, the initial sample had 10 HF and 24 LF data points. Table 22 shows the results found for $\overline{V}_{c,max} = 30\%$. It is clear that this problem is indeed much harder than the previous example, which had only 5 design variables. After all, surrogate models, and especially GP models, often struggle in approximating high-dimensional spaces. However, with exception to the PI-based methods, all approaches managed to find designs very close to the optima.

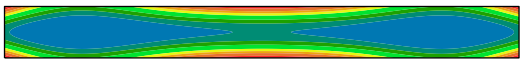
In terms of accuracy, LCB-based approaches were, consistently, the best-performing

Figure 80 – Optimum design for the optimization of a tridirectional FG plate considering $\overline{V}_{c,max} = 30\%$.



(a) Ceramic distribution throughout the plate

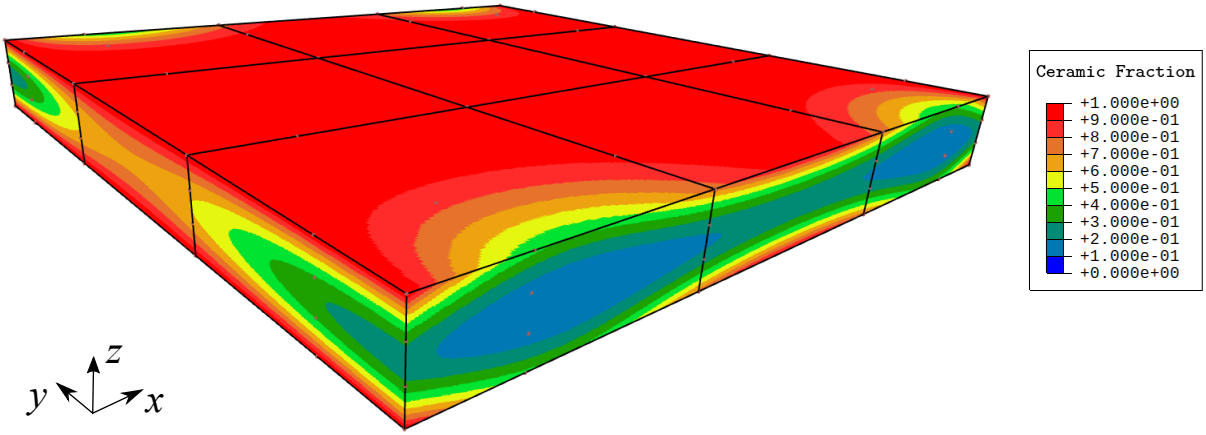
(b) Ceramic distribution at $x = a/3$



(c) Ceramic distribution at $y = a/3$

Source: the author

Figure 81 – Optimum design for the optimization of a tridirectional FG plate considering $\overline{V}_{c,max} = 50\%$.



(a) Ceramic distribution throughout the plate

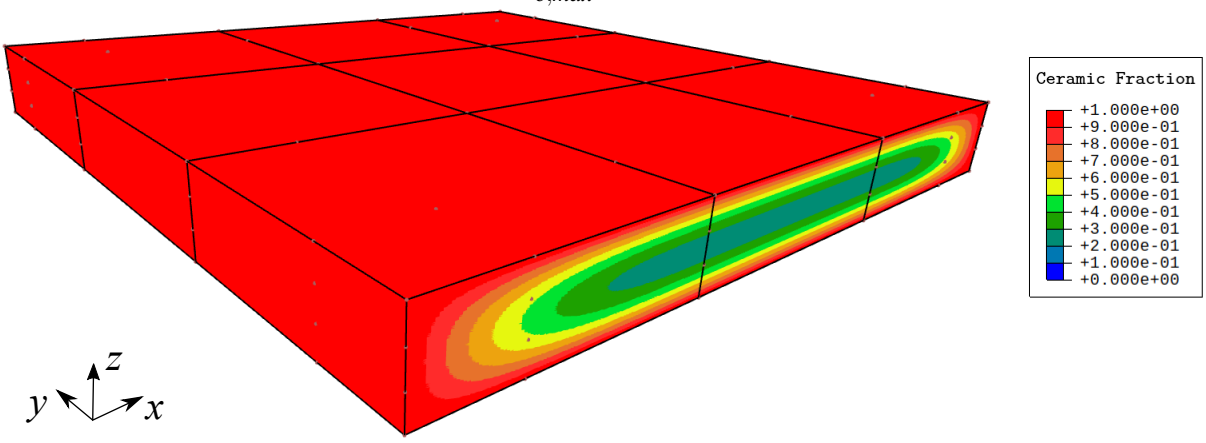
(b) Ceramic distribution at $x = a/3$

(c) Ceramic distribution at $y = a/3$

Source: the author

methods. In that matter, the HKRG-LCB stood out, showing an average NRMSE of only 0.03%. However, LCB-based approaches were the most time-consuming methods in almost all cases. Also, in all cases, MFMs managed to reduce the number of HF evaluations required. However, in this case, the gain in efficiency for these models is very slight or nonexistent in most cases. The exception are the Variable Fidelity (VF) approaches, where the reduction in the number of HF

Figure 82 – Optimum design for the optimization of a tridirectional FG plate considering $\overline{V}_{c,max} = 70\%$.



(a) Ceramic distribution throughout the plate

(b) Ceramic distribution at $x = a/3$

(c) Ceramic distribution at $y = a/3$

Source: the author

Table 22 – Averaged results for the maximization of the buckling load of a tridirectional FG plate considering $\overline{V}_{c,max} = 30\%$.

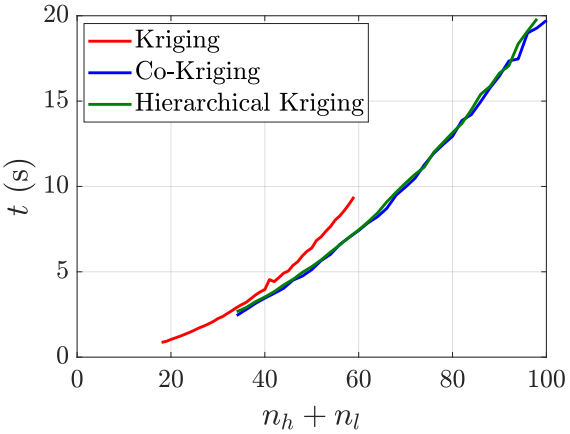
Model	Infill criterion	λ_{norm}	NRMSE	$n_{ev,h}$	$n_{ev,l}$	Time spent (s)
KRG	LCB	2.584	0.12%	46 (28)	-	1473
	PI	2.549	1.48%	32 (14)	-	959
	EI	2.580	0.26%	41 (23)	-	1257
	WEI	2.580	0.28%	40 (22)	-	1226
COKRG	LCB	2.585	0.08%	39 (29)	53 (29)	1596
	PI	2.539	1.85%	26 (16)	40 (16)	866
	EI	2.580	0.26%	32 (22)	46 (22)	1152
	WEI	2.582	0.19%	28 (18)	42 (18)	990
HKRG	LCB	2.586	0.03%	35 (25)	49 (25)	1308
	PI	2.526	2.36%	31 (21)	45 (21)	1088
	EI	2.564	0.88%	22 (12)	36 (12)	737
	WEI	2.561	0.99%	23 (13)	37 (13)	756
	VF-LCB	2.584	0.10%	21 (11)	48 (24)	853
	VF-PI	2.536	1.97%	26 (16)	41 (17)	910
	VF-EI	2.577	0.40%	15 (5)	38 (14)	557
	VF-WEI	2.576	0.43%	17 (7)	39 (15)	641

evaluations was even more noticeable. These approaches managed to perform the optimization process 1.5 to 2.2 times faster when compared to their single-fidelity counterparts. It is worth to notice that, while the HKRG-VF-EI and HKRG-VF-WEI were the most efficient methods, these also showed a high average NRMSE. Overall, the best performing approach was the HKRG-VF-LCB method, which showed a low NRMSE while also reducing the number of HF

evaluations required to find the optimum.

Figure 83 shows how the building cost changes as the sample size ($n_h + n_l$) increases. Note that the cost quickly increases, and, soon, it is almost as expensive as the evaluation cost. This is the reason why, in this problem, there is very little gain in efficiency using most MFM methods. However, if evaluation costs were higher, these methods would certainly be more appealing in terms of efficiency. It is interesting to note that, for the same sample size ($n_h + n_l$), building of the Kriging model is more expensive than the MFMs. This likely occurs because, for the building of Kriging, multiple full size matrices need to be inverted, while building of MFMs require inversion of two separate matrices, one $n_h \times n_h$ and one $n_l \times n_l$.

Figure 83 – Time spent for the building phase for each model for the maximization of the buckling load of a tridirectional FG plate.



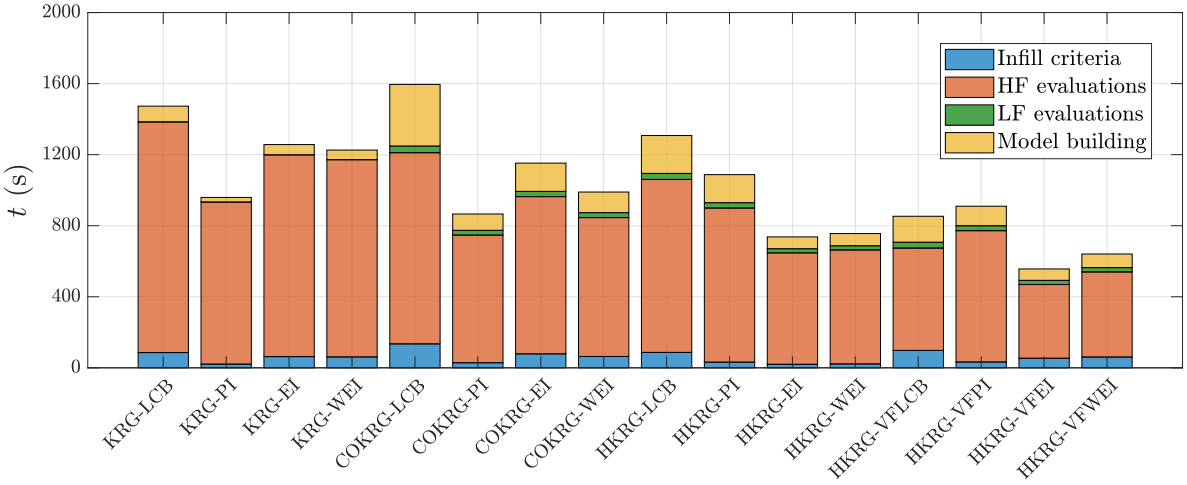
Source: the author

Figure 84 compares the time spent for each approach, showing how the computational cost is divided for each phase of the process. Notice that, while LF evaluations present a very minor overall cost, these also contribute to higher building and infilling costs, as sample size increases. However, we can see more easily that, indeed, MFMs are slightly more efficient overall, especially when we look at Variable Fidelity (VF) approaches.

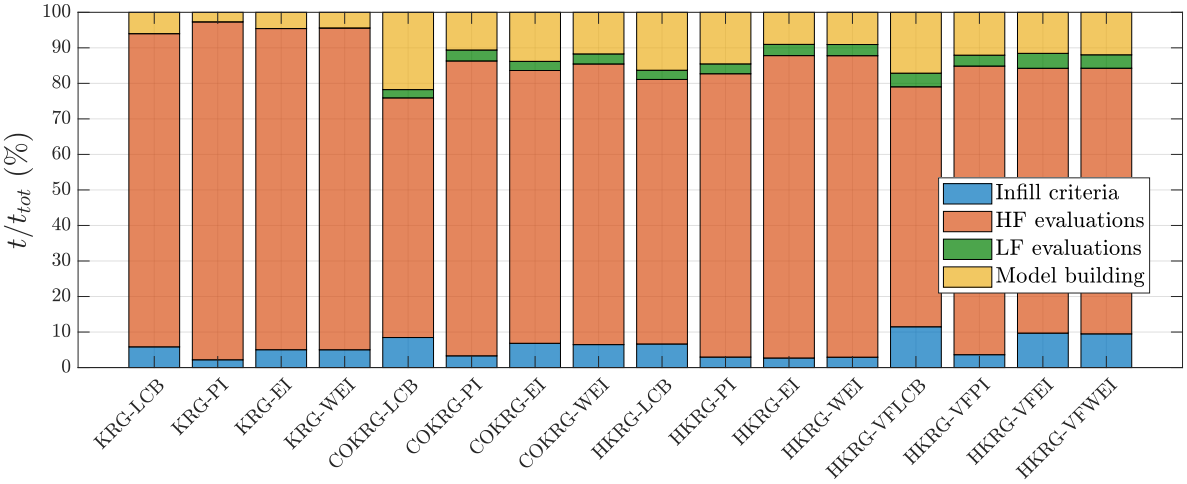
Regarding the relative cost of each phase, on Kriging, HF evaluations correspond to 90% to 95% of the total cost, while, on MFMs, these only correspond to 65% to 85%. On the other hand, while model building corresponds to around 5% of the total cost on Kriging, it may take up to 22% of the total cost on MFMs. This clearly shows that the trade-off between HF and LF data points should be made with caution, and not an overwhelmingly large amount of points should be used in the model, as this would make it too expensive to build and evaluate.

Table 23 shows the results for $\overline{V}_{c,max} = 50\%$. Overall, the results were very similar to

Figure 84 – Cost of each phase of the process for the maximization of the buckling load of a tridirectional FG plate.



(a) Total time spent in each phase



(b) Percentage contribution relative to the total cost of each method

Source: the author

the ones for $\overline{V}_{c,max} = 30\%$. Here, the KRG-LCB showed a higher NRMSE, and was outperformed by KRG-EI and KRG-WEI. On the other hand, for the HKRG, the opposite happened: HKRG-LCB showed a very low error (NRMSE of 0.00%), while HKRG-EI and HKRG-WEI performed poorly. Interestingly enough, all VF-based approaches, with exception to the VF-PI, performed very well in terms of NRMSE. Here, considering both the accuracy and the efficiency of the method, the best performing approaches were indeed the HKRG-VF-EI and the HKRG-VF-WEI, even though the HKRG-VF-LCB did seem to find the optimum in almost all cases, showing an NRMSE of 0.00%.

Finally, Table 24 shows the results for $\overline{V}_{c,max} = 70\%$. Here, once again, the best performing methods, in terms of accuracy, were the LCB-based ones. EI and WEI-based approaches also performed well, even though the HKRG-EI and HKRG-WEI showed poor

Table 23 – Averaged results for the maximization of the buckling load of a tridirectional FG plate considering $\overline{V}_{c,max} = 50\%$.

Model	Infill criterion	λ_{norm}	NRMSE	$n_{ev,h}$	$n_{ev,l}$	Time spent (s)
KRG	LCB	2.895	0.96%	45 (27)	-	1459
	PI	2.870	1.81%	36 (18)	-	1082
	EI	2.922	0.02%	43 (25)	-	1372
	WEI	2.913	0.32%	41 (23)	-	1304
COKRG	LCB	2.923	0.01%	36 (26)	50 (26)	1414
	PI	2.872	1.72%	26 (16)	40 (16)	874
	EI	2.922	0.02%	29 (19)	43 (19)	1060
	WEI	2.922	0.04%	28 (18)	42 (18)	995
HKRG	LCB	2.923	0.00%	36 (26)	50 (26)	1403
	PI	2.860	2.16%	33 (23)	47 (23)	1174
	EI	2.905	0.60%	22 (12)	36 (12)	720
	WEI	2.877	1.55%	22 (12)	36 (12)	721
	VF-LCB	2.923	0.00%	29 (19)	50 (26)	1159
	VF-PI	2.854	2.35%	29 (19)	44 (20)	1014
	VF-EI	2.921	0.04%	14 (4)	40 (16)	548
	VF-WEI	2.921	0.06%	15 (5)	40 (16)	582

results, in terms of accuracy. One more time, the VF-LCB, VF-EI, and VF-WEI showed excellent results, where the VF-LCB presents a lower NRMSE, while VF-EI and VF-WEI present a lower number of HF evaluations and, therefore, time spent.

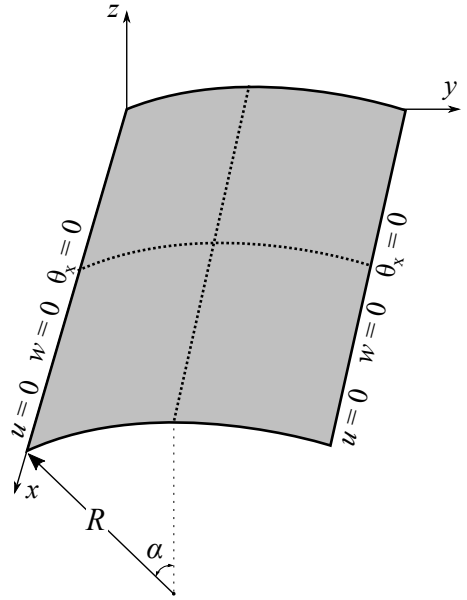
Table 24 – Averaged results for the maximization of the buckling load of a tridirectional FG plate considering $\overline{V}_{c,max} = 70\%$.

Model	Infill criterion	λ_{norm}	NRMSE	$n_{ev,h}$	$n_{ev,l}$	Time spent (s)
KRG	LCB	3.147	0.01%	45 (27)	-	1416
	PI	3.118	0.93%	39 (21)	-	1156
	EI	3.143	0.13%	42 (24)	-	1287
	WEI	3.146	0.05%	43 (25)	-	1304
COKRG	LCB	3.147	0.01%	40 (30)	54 (30)	1587
	PI	3.119	0.90%	31 (21)	45 (21)	1055
	EI	3.145	0.08%	29 (19)	43 (19)	997
	WEI	3.145	0.09%	29 (19)	43 (19)	1039
HKRG	LCB	3.147	0.00%	36 (26)	50 (26)	1347
	PI	3.106	1.32%	33 (23)	47 (23)	1141
	EI	3.121	0.84%	25 (15)	39 (15)	829
	WEI	3.098	1.57%	22 (12)	36 (12)	725
	VF-LCB	3.147	0.01%	25 (15)	42 (18)	927
	VF-PI	3.112	1.13%	31 (21)	47 (23)	1109
	VF-EI	3.144	0.12%	17 (7)	42 (18)	650
	VF-WEI	3.145	0.08%	16 (6)	40 (16)	605

7.2.3 Maximization of the fundamental frequency of a FG shallow shell

In this section, the maximization of the natural frequency of a SUS304/SI_3N_4 FG shallow shell will be performed. The shell is depicted in Figure 85, and it has length $L = 0.508$ m, $R = 2.540$ m, and $\alpha = 0.1$ rad. Shell thickness is considered to be $h = 0.0127$ m. Equivalent material properties are estimated via the Mori-Tanaka model, and the material gradation is given by a univariate B-Spline function with 9 control points in the thickness direction.

Figure 85 – Shallow shell considered for the maximization of the fundamental frequency.



Source: the author

The goal of this problem is to maximize the natural frequency of the shell. No symmetry constraint is imposed and, thus, as the design variables are the B-Spline control points, there are 9 design variables. Here, a limitation on the maximum cost of the shell is imposed, and the optimization problem is described as:

$$\begin{cases} \text{find} & \mathbf{x} = \{x_1, x_2, \ldots, x_9\} \\ \text{that minimizes} & -\omega_{norm}(\mathbf{x}) \\ \text{subject to} & g_1(\mathbf{x}) = \dfrac{C_{tot}(\mathbf{x})}{C_{max}} - 1 \leq 0 \\ \text{with} & 0 \leq x_i \leq 1 \end{cases} \tag{7.35}$$

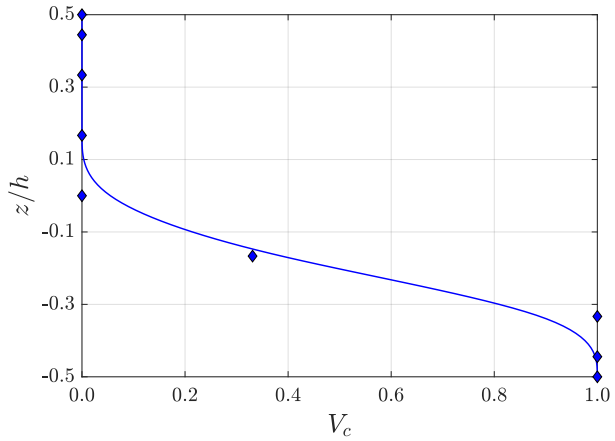
where ω_{norm} is the normalized natural frequency, given by Eq. (3.55), and:

$$C_{tot} = C_c V \frac{1}{h} \int_{-h/2}^{h/2} V_c dz + C_m V \frac{1}{h} \int_{-h/2}^{h/2} V_m dz \tag{7.36}$$

where V is the shell volume, C_{tot} is the total cost, $C_{max} = 0.7$ M.U. is the maximum cost, and $C_c =$ 20 M.U./m^3 and $C_m = 1$ M.U./m^3 are the costs per unit volume for the ceramic and the metal, re-

spectively. The optimum of this problem is found at $\mathbf{x} = \{1.0, 1.0, 1.0, 0.331, 0.0, 0.0, 0.0, 0.0, 0.0\}$, where $\omega_{norm} = 4.143 \times 10^{-3}$. Figure 86 presents the B-Spline function for the optimum design.

Figure 86 – Optimum design for the maximization of the fundamental frequency of a FG shallow shell.



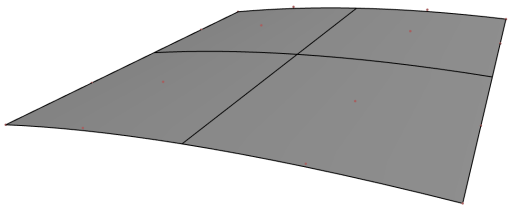
Source: the author

In this problem, only the FSDT will be used. The HF source will be given by a 64×64 cubic NURBS mesh. However, two different LF sources will be tested: a 2×2 quadratic NURBS mesh and a 8×8 quadratic NURBS mesh. Thus, the problem will be optimized two times, each using a different LF source. Figure 87 depicts these three meshes. On average, the analysis time for each mesh is: 0.11 s for the 2×2 mesh, 0.23 s for the 8×8 mesh, and 54.81 s for the 64×64 mesh. In this work, for the single-fidelity models, 9 HF sampling points will be used. For the MFMs, 5 HF and 16 LF sampling points will be used.

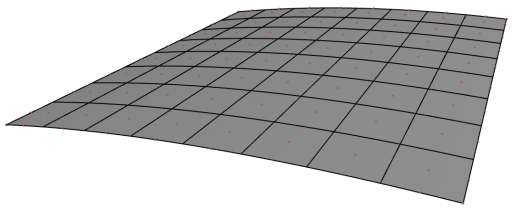
Figure 88 shows the correlation between the HF source and each LF source. In the first case, using the 2×2 mesh as the LF source, the average error is 5.52%. For the 8×8 mesh, the average error is much lower (0.29%). As the LF source becomes more refined, the difference between sources is reduced. That being said, for both cases, the correlation between responses is very high ($R^2 \approx 0.999$).

Table 25 shows the results found when the 2×2 mesh is used as the LF source. Once again, no PI-based approaches were able to find good results. In this problem, since the HF analysis is more expensive, the gain in efficiency for MFMs is more noticeable. EI and WEI-based approaches performed worse using MFMs, especially for the HKRG. These showed higher NRMSE than their single-fidelity counterparts. However, the gain in efficiency of VF approaches was very high. LCB-based approaches showed very high accuracy overall, where both HKRG-LCB and HKRG-VF-LCB showed an NRMSE of 0.00%. In particular, the

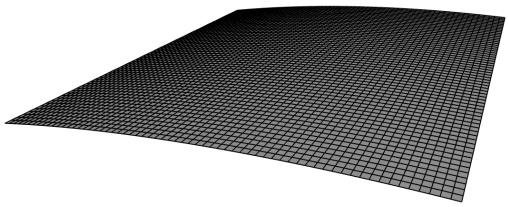
Figure 87 – Meshes used for each source for the maximization of the natural frequency of a FG shallow shell.



(a) 2×2 quadratic NURBS mesh



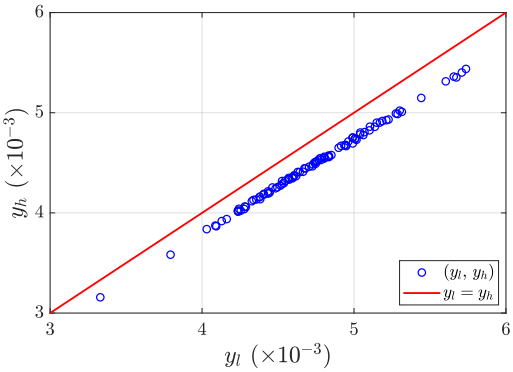
(b) 8×8 quadratic NURBS mesh



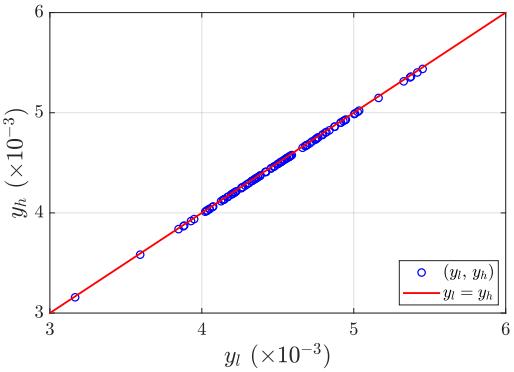
(c) 64×64 cubic NURBS mesh

Source: the author

Figure 88 – Correlation between sources for the maximization of the natural frequency of a FG shallow shell.



(a) LF source: 2×2 mesh



(b) LF source: 8×8 mesh

Source: the author

HKRG-VF-LCB also showed a relevant gain in efficiency, as it performed less HF evaluations due to the VF approach.

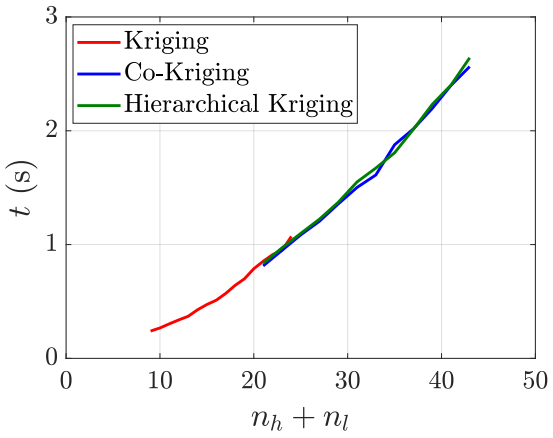
Figure 89 shows the time spent for the building phase for each model for this case. MFMs showed a higher building cost, as it is related to the cube of the sample size. Still, the time spent is not very high, as few sampling points are employed in this problem, due to the low number of design variables.

This aspect is more clear in Figure 90. Here, we note that, for these problems, model building and infill criteria costs are almost negligible when compared to the cost of performing the HF evaluations, even for MFMs. This also occurs due to the higher cost of the linearized

Table 25 – Averaged results for the maximization of the fundamental frequency of a FG shallow shell using a 2×2 mesh as the LF source.

Model	Infill criterion	ω_n $(\times 10^{-3})$	NRMSE	$n_{ev,h}$	$n_{ev,l}$	Time spent (s)
KRG	LCB	4.143	0.01%	39 (30)	-	2201
	PI	4.102	0.99%	29 (20)	-	1600
	EI	4.139	0.10%	28 (19)	-	1514
	WEI	4.141	0.06%	30 (21)	-	1615
COKRG	LCB	4.142	0.02%	25 (20)	36 (20)	1395
	PI	4.062	1.96%	25 (20)	36 (20)	1368
	EI	4.137	0.14%	20 (15)	31 (15)	1105
	WEI	4.141	0.07%	22 (17)	33 (17)	1231
HKRG	LCB	4.143	0.00%	23 (18)	34 (18)	1256
	PI	4.059	2.03%	25 (20)	36 (20)	1396
	EI	4.123	0.49%	19 (14)	30 (14)	1048
	WEI	4.124	0.47%	18 (13)	29 (13)	995
	VF-LCB	4.143	0.00%	16 (11)	37 (21)	939
	VF-PI	4.077	1.60%	21 (16)	33 (17)	1175
	VF-EI	4.108	0.84%	8 (3)	31 (15)	452
	VF-WEI	4.117	0.64%	8 (3)	33 (17)	485

Figure 89 – Time spent for the building phase for each model for the maximization of the critical temperature of a FG shallow shell using a 2×2 mesh as the LF source.

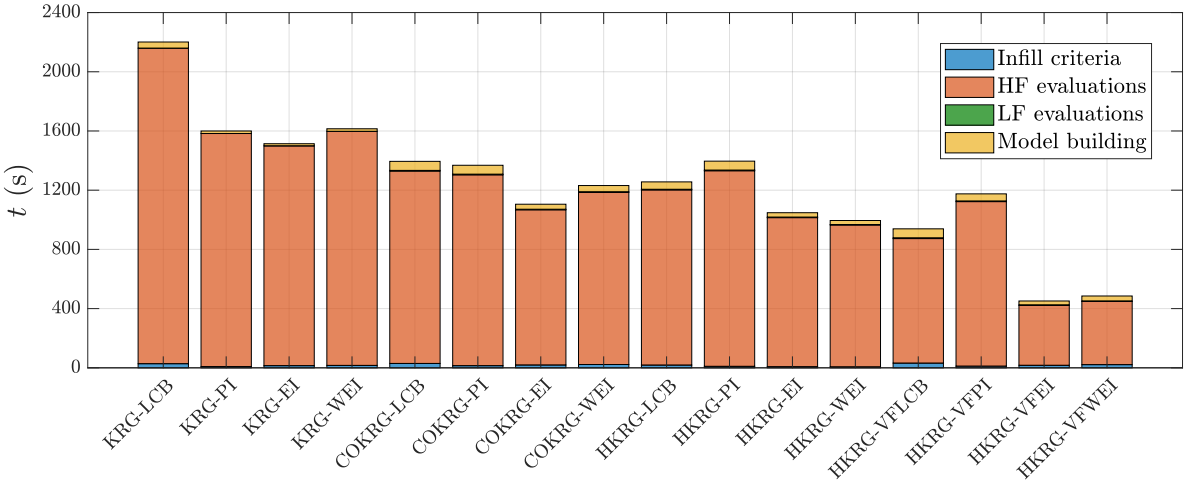


Source: the author

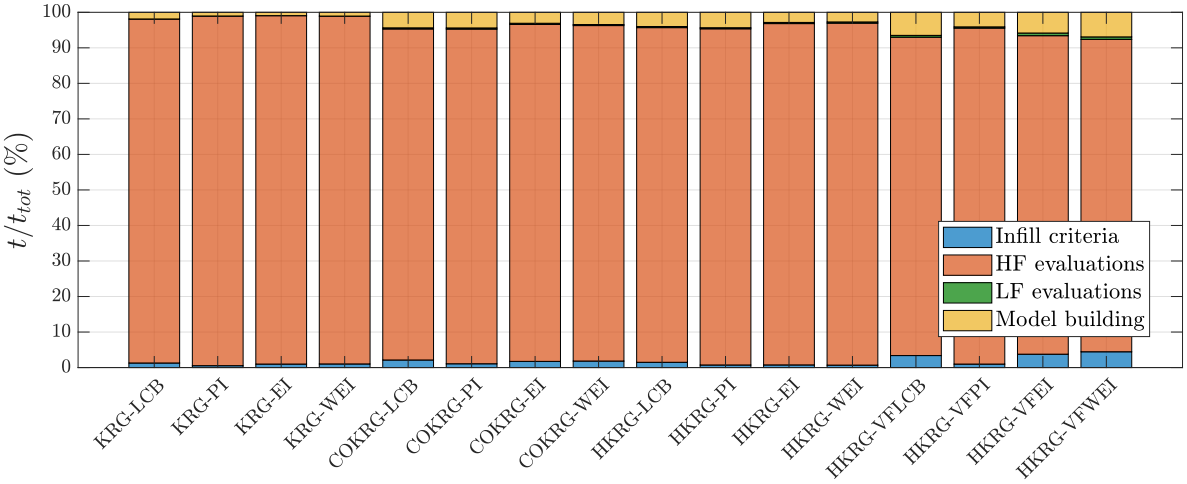
buckling analysis for this problem. On average, for the Kriging model, HF analyses cost is related to 97.8% of the total optimization cost, while, for MFMs, these are related to 94.7%. Again, model building presents a higher contribution to the total cost in MFMs.

Table 26 shows the results found when the 8×8 mesh is used as the LF source. PI-based approaches still behaved poorly, even though we are using a more refined LF source. However, results found using the HKRG-VF-EI and the HKRG-VF-WEI were much better, especially for the former. At the same time, results for the HKRG-VF-LCB were not as accurate. Overall, the best approach was the HKRG-VF-EI, as it was able to achieve a very low NRMSE

Figure 90 – Cost of each phase of the process for the maximization of the fundamental frequency of a FG shallow shell using a 2×2 mesh as the LF source.



(a) Total time spent in each phase



(b) Percentage contribution relative to the total cost of each method

Source: the author

while presenting a major gain in efficiency. Even though other methods presented a slightly better accuracy, they showed a much higher computational cost.

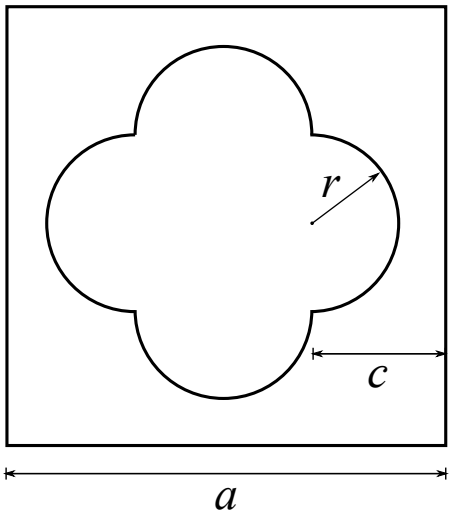
7.2.4 Maximization of the critical buckling temperature of a FG plate with a complicated cutout

This time, the maximization of the critical buckling temperature of a clamped SUS304/Si_3N_4 FG square plate with a complicated cutout will be performed. The plate is depicted in Figure 91. For this problem, $a = 10$ m, $r = 2$ m, $c = 3$ m, and $a/h = 100$. Once again, equivalent material properties will be estimated via the Mori-Tanaka scheme, and material properties will be given by a univariate B-Spline function with 13 control points.

Table 26 – Averaged results for the maximization of the fundamental frequency of a FG shallow shell using a 8×8 mesh as the LF source.

Model	Infill criterion	ω_n $(\times 10^{-3})$	NRMSE	$n_{ev,h}$	$n_{ev,l}$	Time spent (s)
KRG	LCB	4.143	0.00%	40 (31)	-	2178
	PI	4.071	1.74%	27 (18)	-	1457
	EI	4.141	0.05%	28 (19)	-	1525
	WEI	4.141	0.06%	29 (20)	-	1569
COKRG	LCB	4.143	0.00%	23 (18)	34 (18)	1332
	PI	4.080	1.53%	23 (18)	34 (18)	1304
	EI	4.140	0.08%	22 (17)	33 (17)	1205
	WEI	4.141	0.04%	23 (18)	34 (18)	1267
HKRG	LCB	4.141	0.06%	28 (23)	39 (23)	1550
	PI	4.060	2.02%	28 (23)	39 (23)	1587
	EI	4.099	1.07%	19 (14)	30 (14)	1076
	WEI	4.112	0.75%	18 (13)	29 (13)	1034
	VF-LCB	4.105	0.91%	16 (11)	39 (23)	976
	VF-PI	4.039	2.51%	24 (19)	36 (20)	1399
	VF-EI	4.141	0.06%	10 (5)	34 (18)	565
	VF-WEI	4.128	0.38%	10 (5)	34 (18)	570

Figure 91 – Square plate with a complicated cutout.



Source: the author

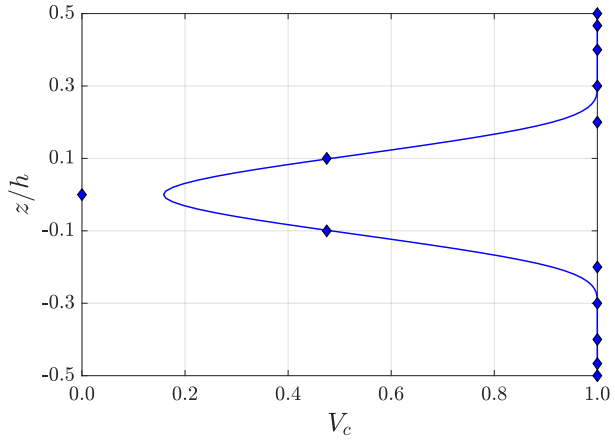
Here, we aim to maximize the critical buckling temperature for the plate. Similar to the first problem, material gradation is symmetric with respect to the mid-plane, and the B-Spline control points are the design variables. This time, there are 7 design variables, and two expensive

constraints will be considered. The optimization problem is given by:

$$\begin{cases} \text{find} & \mathbf{x} = \{x_1, x_2, \ldots, x_7\} \\ \text{that minimizes} & -T_{crit}(\mathbf{x}) \\ \text{subject to} & g_1(\mathbf{x}) = \dfrac{\omega_{norm}(\mathbf{x})}{\omega_{norm,max}} - 1 \leq 0 \\ & g_2(\mathbf{x}) = 1 - \dfrac{\omega_{norm}(\mathbf{x})}{\omega_{norm,min}} \leq 0 \\ \text{with} & 0 \leq x_i \leq 1 \end{cases} \tag{7.37}$$

where T_{crit} is the critical buckling temperature and ω_{norm} is given by Eq. (3.55). For this problem, $\omega_{norm,min} = 0.009$ and $\omega_{norm,max} = 0.015$. The optimum for this problem is found at $\mathbf{x} = \{1.0, 1.0, 1.0, 1.0, 1.0, 0.475, 0.0\}$, where $T_{crit} = 316.285$ K and $\omega_{norm} = 0.015$. Figure 92 shows the B-Spline for the optimum design.

Figure 92 – Optimum design for the maximization of the critical buckling temperature of a FG plate with a complicated cutout.

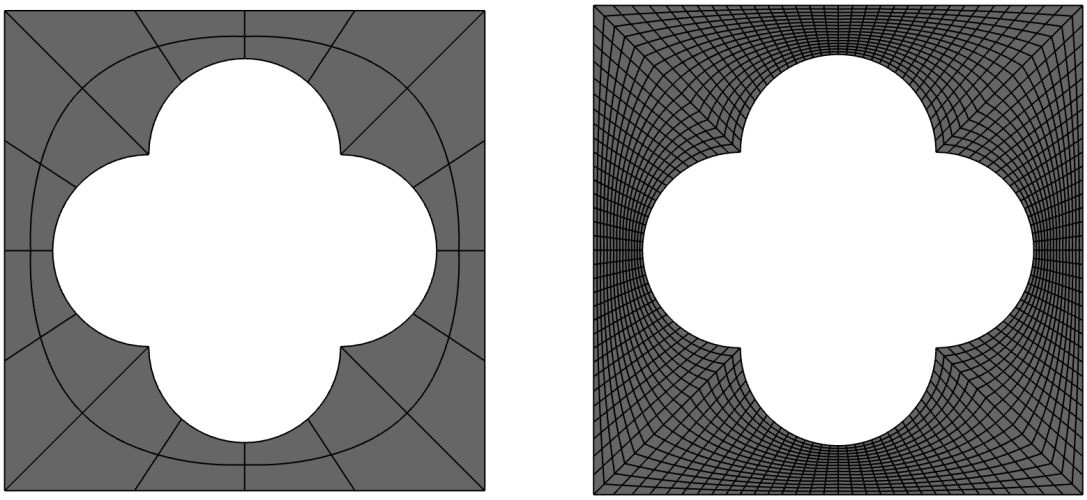


Source: the author

Here, only the FSDT will be employed. HF analyses will be performed using a 2048-element cubic NURBS mesh and, for MFMs, a 32-element cubic NURBS mesh will be considered as the LF source. Figure 93 depicts both of these meshes. Using the HF source, linearized buckling and modal analyses take, on average, 28.83 s and 18.42 s. On the other hand, using the LF source, linearized buckling and modal analyses take 0.42 s and 0.32 s. Thus, it takes 47.25 s to evaluate one design via the HF source, while it takes only 0.75 s to perform one LF evaluation. The relative expense is $C_r = 0.016$.

Figure 94 depicts the correlation between these two sources for both linearized buckling and modal analyses. Correlation between sources is very high for these two analyses ($R^2 \approx 1.0$). However, there is an almost constant relative difference between these. For the linearized buckling analysis, the average error is 2.76%, while it is 1.82% for the modal analysis.

Figure 93 – Meshes used for each source for the maximization of the buckling load of a FG plate with a complicated cutout.

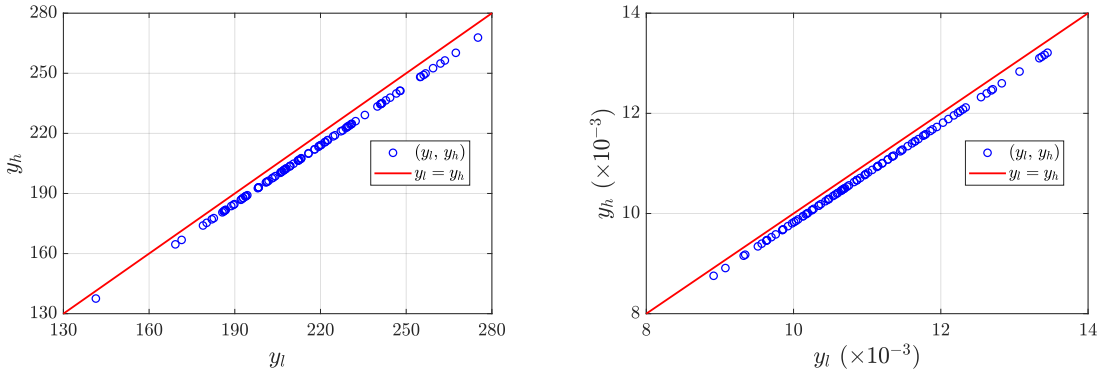


(a) LF source, 32-element cubic NURBS mesh

(b) HF source, 2048-element cubic NURBS mesh

Source: the author

Figure 94 – Correlation between sources for the analysis of a FG plate with a cutout.



(a) Linearized buckling analysis

(b) Modal analysis

Source: the author

To avoid overwhelming the results, PI and WEI-based methods will not be employed in this problem. It is important to note that, since the constraint function is expensive, constraint-handling methods discussed in Section 5.3.5 will be used. Thus, as previously stated, the following approaches will be employed: the direct approach (Direct), the Probability of Feasibility (PF), the feasibility function proposed by Tutum, Deb and Baran[62] (FFT), the feasibility function proposed by Bagheri et al.[63] (FFB), and the feasibility function proposed by Sohst, Afonso and Suleman[297] (FFS). These methods will be used for SAO with EI-based methods. For LCB-based methods, however, a simple adaptive penalty will be used, where unfeasible individuals are penalized. The surrogate model for the constraint function will be used to determine the feasibility of the design.

Note that both expensive constraint functions are related to the same structural response. Thus, it makes no sense to build two models for the two different functions. In this work, for a more efficient process, only one model will be built. This will be made by considering only one constraint function, which is given by:

$$g_p(\mathbf{x}) = \max(g_1, g_2) \leq 0 \tag{7.38}$$

Thus, constraint-handling methods will be applied considering a model for the constraint function $g_p(\mathbf{x})$.

Table 27 shows the results found using the single-fidelity Kriging model. These results already show how the consideration of expensive constraints make the problem much more complex. Optimization using the LCB with an adaptive penalty approach was not able to achieve good results for T_{crit}, where the error found was higher than 10%. While EI-based methods were able to find better results, most constraint-handling methods achieved designs that were still not close to the optima. In terms of accuracy, the best approaches were the PF, proposed by Schonlau, Welch and Jones[60], and the FFS, proposed by Sohst, Afonso and Suleman[297]. Here, only the case where $n = 0.50$ was able to achieve good results.

Table 27 – Averaged results for the maximization of the critical temperature of a FG plate with a complicated cutout using the single-fidelity Kriging model.

Infill criterion	Constraint-handling	T_{crit}	NRMSE	$n_{ev,h}$	Time spent (s)
LCB	Adaptive penalty	281.41	11.03%	27 (13)	1279
EI	Direct	305.20	3.50%	35 (21)	1723
	PF	316.13	0.05%	39 (25)	1905
	FFT	307.69	2.72%	37 (23)	1768
	FFB	308.37	2.50%	37 (23)	1813
	FFS ($n = 0.15$)	304.26	3.80%	34 (20)	1636
	FFS ($n = 0.50$)	316.06	0.07%	37 (23)	1812

In terms of the number of evaluations performed, these methods required close to 40 HF evaluations, which means that, in addition to the initial sample, more than 20 HF evaluations were performed during the adaptive sampling process.

Table 28 now shows the results found using MFMs. Here, the use of LCB-based methods with an adaptive penalty approach was still not able to achieve good results. Even though the LCB showed good results for the optimization using cheap constraint functions, appropriate methods should be derived and tested for optimization using expensive constraints. For the EI-based methods, results were slightly better in terms of accuracy. Using the COKRG,

the FFS with $n = 0.15$ also showed good results. With the HKRG, when VF methods are not used, most approaches found designs close to the optima, even though the NRMSE for the PF and the FFS with $n = 0.50$ were higher than those from the single-fidelity KRG.

Table 28 – Averaged results for the maximization of the critical temperature of a FG plate with a complicated cutout using the Multi-Fidelity Models (MFMs).

Model	Infill	Constr. hand.	T_{crit}	NRMSE	$n_{ev,h}$	$n_{ev,l}$	Time spent (s)
COKRG	LCB	Adapt. pen.	249.34	21.17%	18 (11)	32 (11)	945
	EI	Direct	307.15	2.89%	29 (22)	43 (22)	1634
		PF	316.19	0.03%	32 (25)	46 (25)	1868
		FFT	305.89	3.29%	29 (22)	43 (22)	1618
		FFB	309.63	2.13%	32 (25)	46 (25)	1865
		FFS ($n = 0.15$)	315.52	0.25%	30 (23)	44 (23)	1691
		FFS ($n = 0.50$)	316.09	0.09%	34 (27)	48 (27)	1929
HKRG	LCB	Adapt. pen.	238.03	24.74%	18 (11)	32 (11)	933
	EI	Direct	307.09	2.91%	25 (18)	39 (18)	1313
		PF	315.26	0.34%	28 (21)	42 (21)	1494
		FFT	304.10	3.85%	25 (18)	39 (18)	1344
		FFB	315.39	0.30%	28 (21)	42 (21)	1496
		FFS ($n = 0.15$)	314.38	0.61%	26 (19)	40 (19)	1354
		FFS ($n = 0.50$)	315.00	0.42%	26 (19)	40 (19)	1384
	VF-LCB	Adapt. pen.	245.88	22.26%	16 (9)	32 (11)	826
	VF-EI	Direct	315.47	0.27%	21 (14)	44 (23)	1222
		PF	316.28	0.02%	20 (13)	45 (24)	1198
		FFT	315.52	0.25%	21 (14)	43 (22)	1214
		FFB	316.23	0.04%	20 (13)	45 (24)	1180
		FFS ($n = 0.15$)	307.39	2.82%	25 (18)	44 (23)	1458
		FFS ($n = 0.50$)	316.22	0.04%	21 (14)	45 (24)	1270

That being said, the HKRG-VF-EI showed very good results for most methods. Using the PF, the FFB, and the FFS with $n = 0.50$, the average NRMSE found was very close to 0.00%. Even the FFT and the Direct approach were able to achieve results close to the optima. Only the FFS with $n = 0.15$ showed a bad performance in this case. Also, VF approaches required a lower number of HF evaluations to find the optimum. These methods were able to find very accurate results only using close to 20 HF evaluations.

7.3 Laminate problems

In this section, problems related to the optimization of laminate structures will be described and solved using the proposed algorithms. Note that these require adequate methods due to the discrete nature of these problems. Thus, for the definition of the initial sample, a mapping approach will be adopted, as discussed in Section 5.1.3. Again, unless stated otherwise, the initial sample will be defined via the LHS_N method, with $N = 20$.

Regarding the use of surrogate models, not much will be changed. Since GP models are built based on a measure of a similarity between data points, one can easily train a model simply by ignoring the discrete space. Also, no modification will be made to the correlation functions. Even though the problem is discrete by nature, the Maximum Likelihood Estimator (MLE) for GPs is still a continuous optimization problem, and, thus, a conventional PSO will be used to maximize the MLE. For model building, parameters are the same as those presented in Table 11. However, for the infill process, the discrete nature of data should be considered, so that only feasible points are added. This way, the modified laminate GA presented in Section 4.2.1 will be employed[15]. The parameters considered for the laminate GA can be found in Table 29.

Table 29 – Parameters used for the laminate GA algorithm employed for infill criteria on discrete problems.

N_{gen}	250
N_p	100
Gen_{stall}	100
C_r	0.80
p_{mut}	10%
p_{del}	0%
p_{add}	0%
p_{swap}	5%

7.3.1 Maximization of the buckling load of a laminate considering different boundary conditions

The first laminate problem deals with the maximization of the buckling load of a 20-ply square plate. Here, $a/h = 50$ and, thus, the thickness of each ply is $h_i = h/20 = a/1000$. In this problem, the optimal stacking sequence should be found, and ply orientations are considered as design variables. Only one material is used, which is Material I from Table 7, and the laminate is balanced and symmetric. Thus, there are five design variables. Also, a constraint limits the

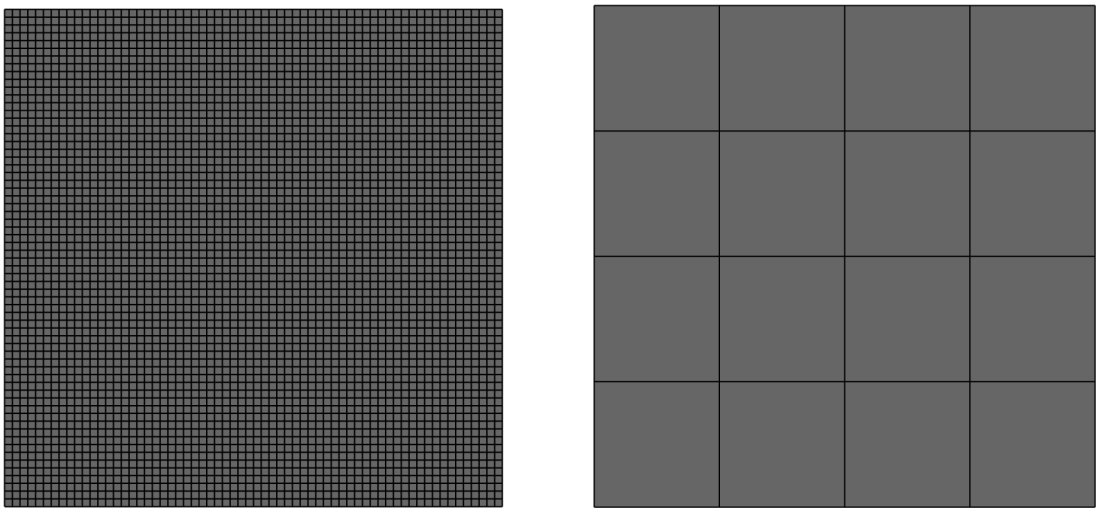
number of maximum contiguous plies ($N_{cont,max}$). The optimization problem is given by:

$$\begin{cases} \text{find} & \mathbf{x} = \{\theta_1, \theta_2, \ldots, \theta_5\} \\ \text{that minimizes} & -\lambda_{norm}(\mathbf{x}) \\ \text{subject to} & g_1(\mathbf{x}) = \dfrac{N_{cont}}{N_{cont,max}} - 1 \leq 0 \\ \text{with} & \theta_i \in \{0°, 15°, 30°, 45°, 60°, 75°, 90°\} \end{cases} \tag{7.39}$$

where λ_{norm} is given by Eq. (3.56) and, for this problem, $N_{cont,max} = 4$.

This problem will be solved with the assistance of single and multi-fidelity models. Meshes used for each source are depicted in Figure 95. The HF source will be given by the analysis of a 64×64 cubic NURBS mesh, while the LF source will be evaluated using a 4×4 quadratic NURBS mesh. Both cases consider the FSDT. Due to the multimodal nature of laminate problems, slightly more data points will be used. Thus, in this problem, $n = 2m = 10$ initial HF sampling points will be used to fit a Kriging model. For MFMs, the initial sample will have $n_h = 5$ and $n_l = 15$.

Figure 95 – Mesh used for each source for the maximization of the buckling load of a laminate square plate.



(a) HF source, 64×64 cubic NURBS mesh (b) LF source, 4×4 quadratic NURBS mesh

Source: the author

This problem will be solved considering two cases. First, a simply supported squared plate will be optimized, considering the SS1 boundary condition. Then, the same problem will be solved considering a clamped (CC) square plate. Table 30 shows the optimum design found for these two cases.

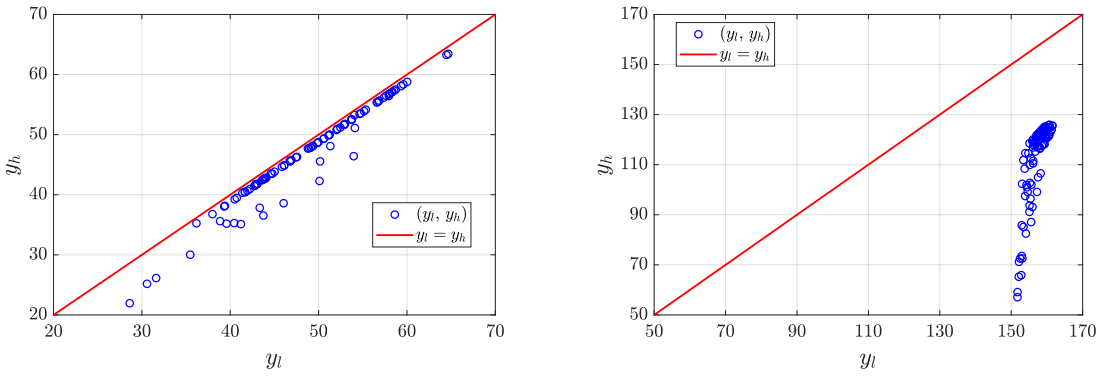
Figure 96 shows the correlation between HF and LF response for these two cases.

Table 30 – Optimum design for the maximization of the buckling load of a laminate square plate considering different boundary conditions[3]

Boundary condition	Layup	λ_{norm}
Simply supported (SS1)	$[\pm 45^\circ_5]_s$	64.987
Clamped (CC)	$[0^\circ_4/90^\circ_2/0^\circ_2/\pm 15^\circ]_s$	126.39

For the simply supported case, the correlation factor is very high ($R^2 = 0.96$), similar to previous examples. The average error is 4.78% and, in the optimum, the error is 1.90%. However, depending on the layup, errors can be up to 30.41%. For the clamped boundary condition, however, sources show a much higher difference in their responses. Correlation factor is only $R^2 = 0.6643$. The error in the optimum design is 27.42%, the average error is 45.47%, and, depending on the layup, the relative difference can be up to 165.94%. This is due to the slightly more complex displacement patterns seem in the buckling of a clamped plate, which are not easily reproduced using the LF mesh. That being said, while the HF source takes 38.43 s to be evaluated, the LF source takes only 0.26 s. The relative expense is $C_r = 0.007$.

Figure 96 – Correlation between sources for the maximization of the buckling load of a laminate square plate considering different boundary conditions.



(a) Simply supported (b) Clamped

Source: the author

Table 31 shows the results for the simply supported case. Just by looking at the NRMSE, we can see that laminate problems are more complex than FG problems. This becomes more clear when we compare the results to those from the 5-variable FG problem presented in Section 7.2.1. It is interesting to note that, here, the PI-based approaches behaved better than the others, with exception to the VF-PI.

In terms of accuracy, the best approach was the HKRG-PI, which showed an NRMSE of 0.00%, while also showing a rather good efficiency. The HKRG-VF-EI outperformed other approaches in terms of efficiency and, unlike HKRG-VF-WEI, it did not present a major loss in

Table 31 – Averaged results for the maximization of the buckling load of a simply supported laminate square plate.

Model	Infill criterion	λ_{norm}	NRMSE	$n_{ev,h}$	$n_{ev,l}$	Time spent (s)
KRG	LCB	64.192	1.23%	40 (30)	-	1538
	PI	64.975	0.03%	33 (23)	-	1283
	EI	64.803	0.29%	37 (27)	-	1420
	WEI	64.148	1.30%	39 (29)	-	1486
COKRG	LCB	64.558	0.67%	30 (25)	40 (25)	1239
	PI	64.989	0.01%	27 (22)	37 (22)	1074
	EI	64.774	0.34%	24 (19)	34 (19)	947
	WEI	64.665	0.51%	29 (24)	39 (24)	1195
HKRG	LCB	63.856	1.75%	23 (18)	33 (18)	910
	PI	64.994	0.00%	23 (18)	33 (18)	919
	EI	64.847	0.23%	25 (20)	35 (20)	1007
	WEI	64.049	1.45%	25 (20)	35 (20)	990
	VF-LCB	63.779	1.87%	17 (12)	41 (26)	712
	VF-PI	64.443	0.85%	20 (15)	32 (17)	790
	VF-EI	64.905	0.14%	12 (7)	38 (23)	518
	VF-WEI	63.248	2.69%	10 (5)	36 (21)	419

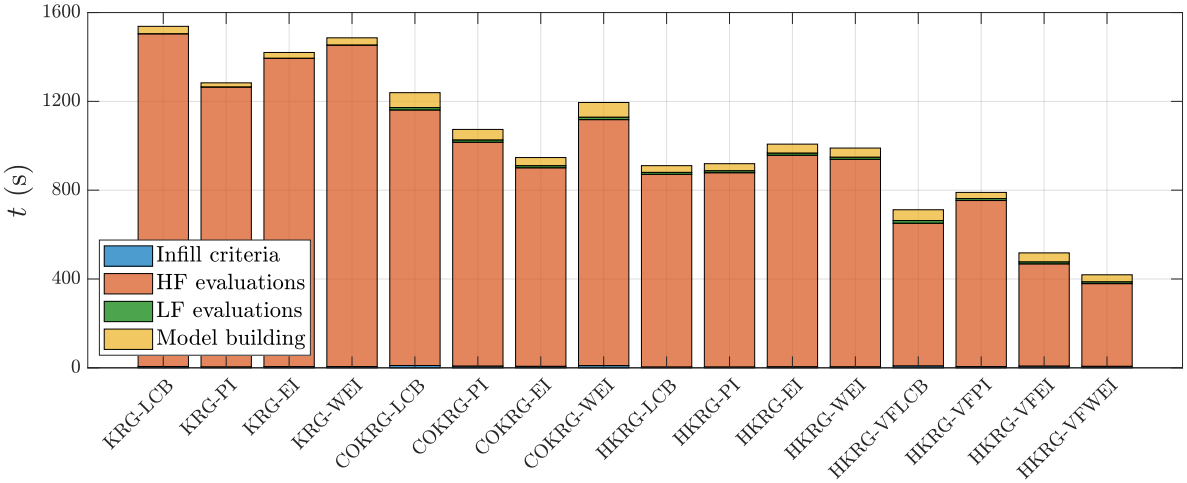
accuracy. This approach required only 12 HF evaluations, while its single-fidelity counterpart required 37 HF evaluations.

Figure 97 shows the cost of each phase of the process. Once again, most of the time is spent evaluating the HF source. Here, time spent in the infill criteria phase and in the evaluation of the LF source is almost negligible. However, model building phase takes almost 10% of the total cost for some multi-fidelity approaches.

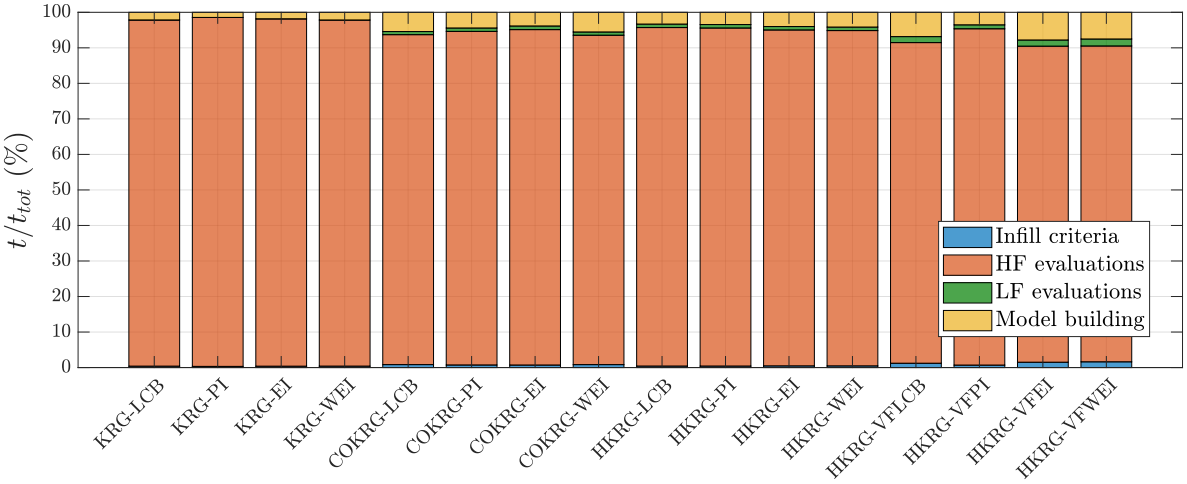
Table 32 shows the results for the clamped case. It is worth to remember that, here, correlation between sources is very low, as shown in Figure 96. That being said, most approaches still managed to find good designs, achieving a very low NRMSE. However, one may note that, here, Variable Fidelity (VF) approaches were not able to present any gain in efficiency over the usual approaches. The exception was the VF-PI, which, in turn, was not able to find designs close to the optimum. The best approach in terms of accuracy was the COKRG-LCB. Despite the poor correlation, all COKRG and non-VF HKRG approaches managed to find good designs while also presenting a gain in efficiency when compared to KRG.

Figure 98 shows the cost of each phase of the process, now for the clamped case. Here, it is even more clear that the VF approaches were not able to present any gain in efficiency over the others. This is most likely due to the very poor correlation between sources for this case. Besides, once again, most of the optimization cost is related to the HF evaluations, while the

Figure 97 – Cost of each phase of the process for the maximization of the buckling load of a simply supported laminate plate.



(a) Total time spent in each phase



(b) Percentage contribution relative to the total cost of each method

Source: the author

model building phase also represents a sizable contribution to the total cost.

7.3.2 Maximization of the buckling load of a laminate considering different number of plies

In this section, the maximization of the buckling load of a simply supported laminate square plate will be performed once again. This time, the problem will be solved considering two different cases: one, for a 2-ply general laminate, and other for a 10-ply symmetric laminate. All layers present the same material properties: $E_1 = 1$ GPa, $E_2 = E_3 = E_1/40$, $G_{12} = G_{13} = 0.6E_2$, $G_{23} = 0.5E_2$, and $\nu_{12} = \nu_{13} = \nu_{23} = 0.25$. Again, the optimal stacking sequence should be found, but both the orientation and the thickness of each ply are considered as design variables. Thus, for the 2-ply case, there are 4 design variables, while, for the 10-ply symmetric case, there

Table 32 – Averaged results for the maximization of the buckling load of a clamped laminate square plate.

Model	Infill criterion	λ_{norm}	NRMSE	$n_{ev,h}$	$n_{ev,l}$	Time spent (s)
KRG	LCB	126.285	0.08%	29 (19)	-	1114
	PI	126.069	0.25%	33 (23)	-	1267
	EI	126.340	0.04%	31 (21)	-	1185
	WEI	126.245	0.11%	30 (20)	-	1151
COKRG	LCB	126.371	0.01%	25 (20)	35 (20)	985
	PI	126.274	0.09%	28 (23)	38 (23)	1137
	EI	126.314	0.06%	22 (17)	32 (17)	870
	WEI	126.277	0.09%	24 (19)	34 (19)	926
HKRG	LCB	126.195	0.15%	20 (15)	30 (15)	765
	PI	126.209	0.14%	27 (22)	37 (22)	1053
	EI	126.340	0.04%	25 (20)	35 (20)	971
	WEI	126.360	0.02%	24 (19)	34 (19)	909
	VF-LCB	126.329	0.05%	20 (15)	31 (16)	790
	VF-PI	124.690	1.34%	7 (2)	29 (14)	303
	VF-EI	126.329	0.05%	24 (19)	36 (21)	947
	VF-WEI	126.359	0.02%	22 (17)	36 (21)	871

are 10 design variables. The optimization problem may be given by:

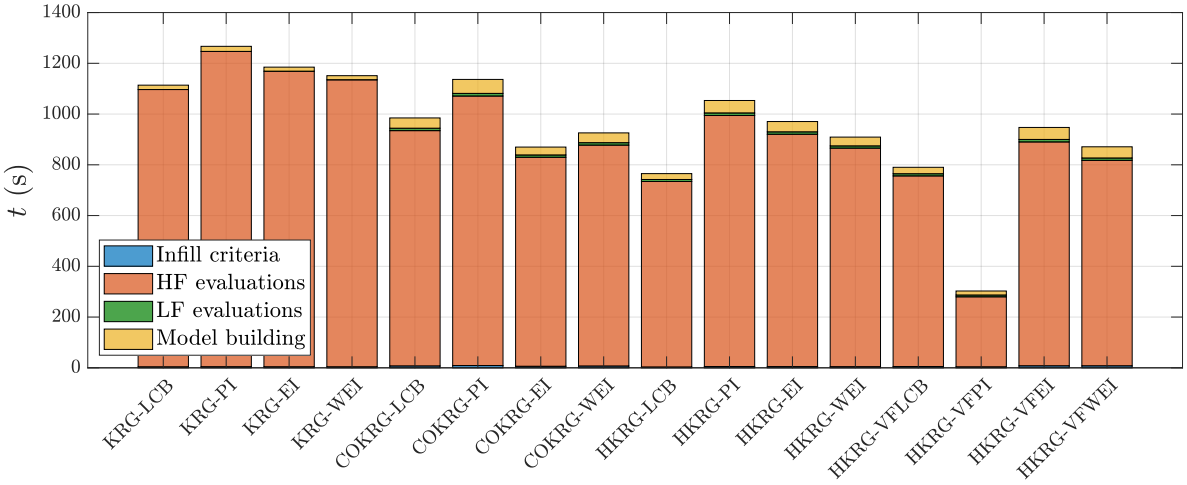
$$\begin{cases} \text{find} & \mathbf{x} = \{h_1, \ldots, h_{m/2}, \theta_1, \ldots, \theta_{m/2}\} \\ \text{that minimizes} & -\lambda_{norm}(\mathbf{x}) \\ \text{subject to} & g_1(\mathbf{x}) = \dfrac{h}{h_{max}} - 1 = 0 \\ \text{with} & \theta_i \in \{-90°, -89°, -88°, \ldots, 88°, 89°, 90°\} \\ & h_i/a \in \{0.005, 0.006, 0.007, \ldots, h_{i,max}/a\} \end{cases} \tag{7.40}$$

where h is the total plate thickness, $h_{i,max}$ is the maximum ply thickness, and λ_{norm} is given by Eq. (3.56). For the 2-ply case, the maximum ply thickness $h_{i,max}/a = 0.095$, while, for the 10-ply case, $h_{i,max}/a = 0.025$. Note that, here, an equality constraint should be satisfied. To guarantee that all trial designs are feasible, the repair algorithm discussed in Section 4.1 will be used.

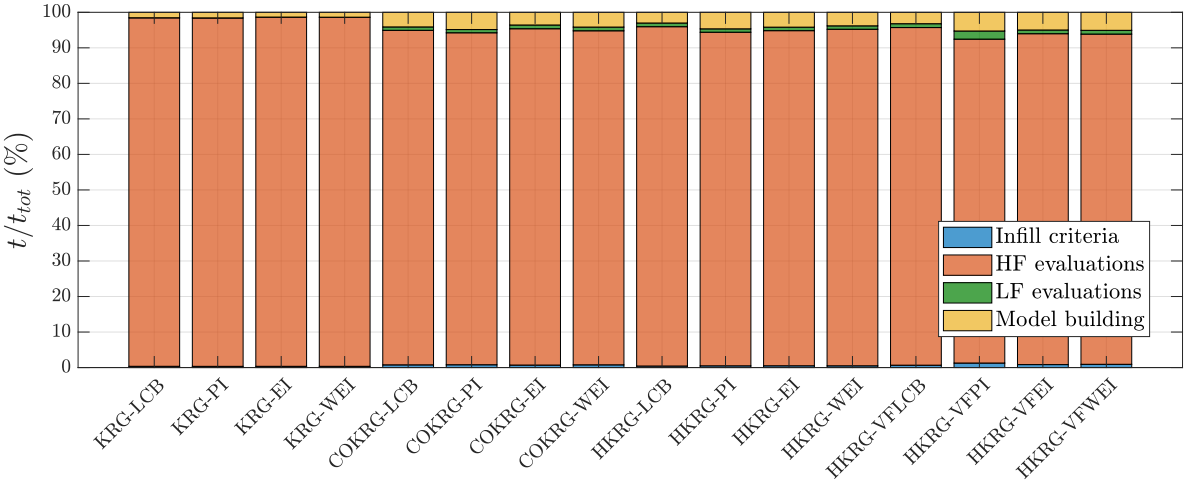
In this problem, we will, once again, consider the fidelity sources depicted in Figure 95. Thus, once again, evaluation of the HF source takes 38.43 s, while evaluation of the LF source takes 0.26 s. Figure 99 shows how these two sources correlate with each other for both cases. For the 2-ply case, the sources are well-correlated ($R^2 = 0.96$) and present an average error of 6.35%. For the 10-ply case, correlation is slightly higher ($R^2 = 0.99$) and the average error is 4.15%. In both of these, the highest error is close to 20%.

A very similar problem was solved by Ho-Huu et al.[328] and Keshtegar et al.[51].

Figure 98 – Cost of each phase of the process for the maximization of the buckling load of a clamped laminate plate.



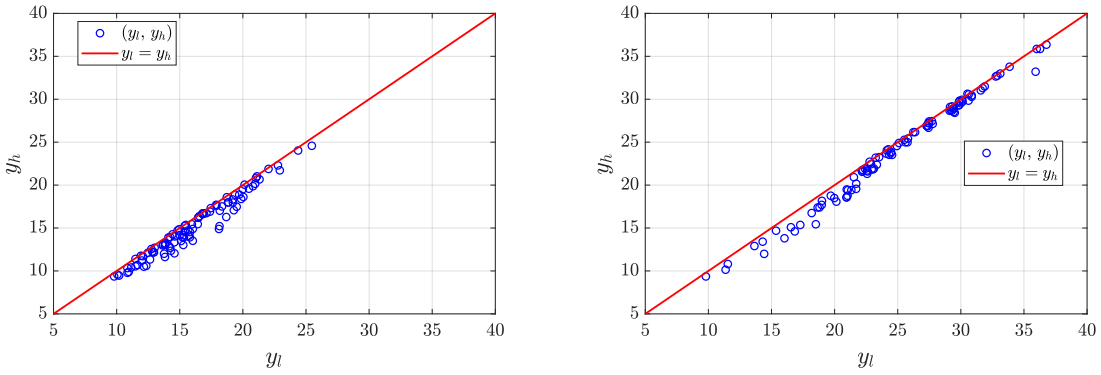
(a) Total time spent in each phase



(b) Percentage contribution relative to the total cost of each method

Source: the author

Figure 99 – Correlation between sources for the maximization of the buckling load of a laminate considering different number of plies.



(a) 2-ply (b) 10-ply

Source: the author

Here, the authors use a simple penalty approach to handle the equality constraint. Also, they consider that $h_{i,max}/a = 0.095$ for all cases. Finally, the authors consider ply thickness as a continuous design variable and, thus, this is a mixed optimization problem. These authors evaluate the buckling load of laminates using a recently proposed smoothed Mindlin plate element, the CS-DSG3.

Ho-Huu et al.[328] uses a modified DE, which uses an adaptive mutation mechanism and is able to handle mixed variable optimization. For the 2-ply and the 10-ply case, the modified DE algorithm performs 1,600 and 7,860 evaluations, respectively. Keshtegar et al.[51], on the other hand, uses a Kriging-improved PSO, where the model is used to assist PSO in defining trial designs. For the 2-ply case, the authors managed to find the optimum with 470 function evaluations. It is very important to note that, by considering thickness as a design variable, the problem becomes much more complex. Kalita, Haldar and Chakraborty[155] comment that thickness is a "hard-to-change" variable, as changing its values may required an extensive design change in the structure and its components.

In this work, Surrogate-Based Optimization (SBO) will be applied using single and multi-fidelity models. For single-fidelity models, the initial sample will have 12 HF points for the 2-ply case and 20 HF points for the 10-ply case. For multi-fidelity models, initial sample will be made of 6 HF and 24 LF points for the 2-ply case, and 10 HF and 40 LF points for the 10-ply case.

Table 33 shows the optimum design obtained in this work for these two cases, where N_{ply} is the number of plies. Here, the optimum designs found by Ho-Huu et al.[328] and Keshtegar et al.[51] are also shown, along with their buckling load factor using the FSDT. Note that, in both cases, the optimum design found in this work was much better than those found on the literature, especially for the 10-ply case. That being said, not all algorithms managed to find the optimum design. For the 2-ply case, the best response was found using the COKRG-LCB. For the 10-ply case, the best response was found using the COKRG-PI. Again, this is a very complex problem, with a very large design space.

Table 34 shows the results for the 2-ply case for SBO using single and multi-fidelity models. Note that, indeed, this problem is much harder than the previous ones for Surrogate-Based Optimization (SBO). The average NRMSE for all cases was much higher. That being said, in almost all cases, the average λ_{norm} was still higher than the one found by the literature. Here, the best approach in terms of accuracy was the HKRG-LCB, which showed an NRMSE of

Table 33 – Optimum design for the maximization of the buckling load of a laminate square plate considering different number of plies.

N_{ply}	Source	θ_i (°) h_i/a ($\times 10^{-3}$)	λ_{norm} CS-DSG3	 FSDT
2	Ho-Huu et al.[328]	[-34/38] [90/10]	26.324	25.331
	Keshtegar et al.[51]	[39/-34] [8.92/91.08]	26.341	25.576
	This work	[-41/40] [8 92]		26.865
10	Ho-Huu et al.[328]	$[37/-37/-36/31/-1]_s$ $[9.47/13.04/11.22/6.77/9.50]_s$	37.884	37.348
	This work	$[33/-36/-35/-33/-32]_s$ $[10/6/13/13/8]_s$		47.495

2.19%, and the best approach in terms of efficiency was the VF-WEI, where the optimization process took only 377 s, requiring only 11 HF function evaluations, on average. Note that, while the VF-LCB showed an even lower time, the method was not able to define adequate designs, as the average NRMSE was much higher than the other methods. With exception to the Variable Fidelity (VF) approaches, the gain in efficiency by the use of MFMs is very marginal, or even non-existent in some cases. Still, almost all multi-fidelity approaches presented higher accuracy than their single-fidelity counterparts.

These averaged results were taken from $N_r = 10$ runs. Table 35 shows the highest buckling load found by each approach. All approaches managed to find a design better than the one from the literature ($\lambda_{opt} = 25.576$ using the FSDT) on at least one run. Also, note that, in all cases, the best design found by multi-fidelity approaches were better than the ones from their single-fidelity counterparts. The optimum design shown in Table 33 was found using the COKRG-LCB approach, but the HKRG-LCB, HKRG-EI, HKRG-VF-LCB, HKRG-VF-EI, and the HKRG-VF-WEI also managed to find designs very close to the optimum.

Figure 100 shows the boxplots for the NRMSE for this case. Here, the blue $\times$ denotes the average NRMSE. In general, the more consistent methods were the WEI-based approaches, especially on MFMs. However, other methods stood out, namely KRG-EI, COKRG-LCB, COKRG-EI, HKRG-LCB, HKRG-WEI, HKRG-VF-EI, and HKRG-WEI. In particular, the HKRG-LCB managed to find designs very close to the optimum and presented a very low variability. Furthermore, the HKRG-VF-EI and HKRG-VF-WEI also presented good results,

Table 34 – Averaged results for the maximization of the buckling load of a 2-ply general laminate.

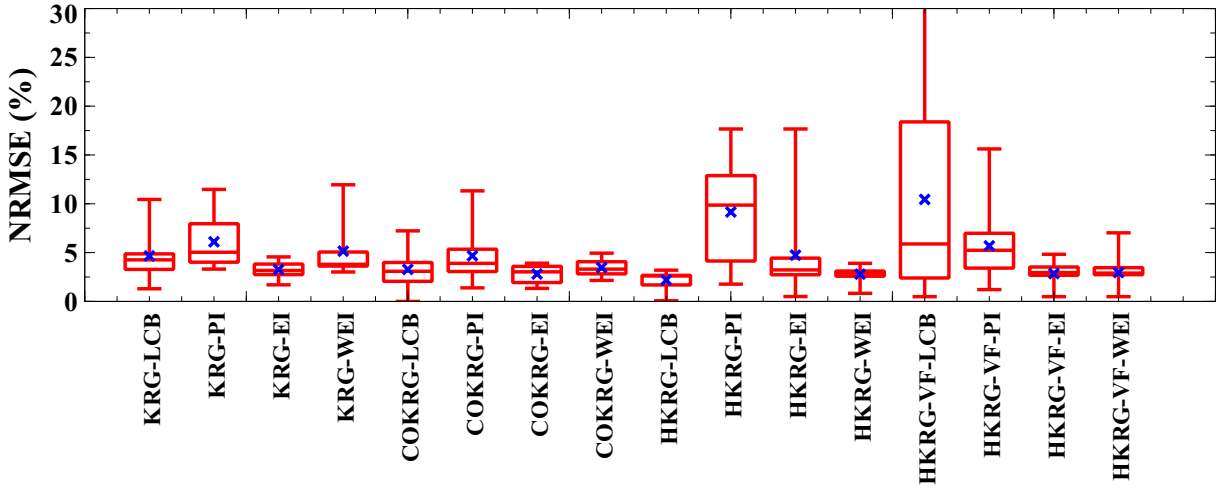
Model	Infill criterion	λ_{norm}	NRMSE	$n_{ev,h}$	$n_{ev,l}$	Time spent (s)
KRG	LCB	25.613	4.66%	37 (25)	-	1061
	PI	25.225	6.11%	37 (25)	-	1038
	EI	25.989	3.26%	32 (20)	-	922
	WEI	25.482	5.15%	37 (25)	-	1063
COKRG	LCB	25.985	3.28%	32 (26)	50 (26)	1029
	PI	25.610	4.67%	29 (23)	47 (23)	887
	EI	26.112	2.80%	29 (23)	47 (23)	917
	WEI	25.934	3.47%	33 (27)	51 (27)	1043
HKRG	LCB	26.277	2.19%	38 (32)	56 (32)	1222
	PI	24.405	9.16%	31 (25)	49 (25)	934
	EI	25.593	4.74%	31 (25)	49 (25)	960
	WEI	26.115	2.79%	38 (32)	56 (32)	1166
	VF-LCB	24.059	10.4%	9 (3)	44 (20)	321
	VF-PI	25.338	5.69%	26 (20)	49 (25)	778
	VF-EI	26.099	2.85%	12 (6)	47 (23)	400
	VF-WEI	26.070	2.96%	11 (5)	49 (25)	377

Table 35 – Highest buckling load ($\lambda_{norm,best}$) found by each approach for a 2-ply general laminate.

Model	LCB	PI	EI	WEI	VF-LCB	VF-PI	VF-EI	VF-WEI
KRG	26.517	25.976	26.409	26.055	-	-	-	-
COKRG	26.865	26.493	26.507	26.287	-	-	-	-
HKRG	26.848	26.390	26.732	26.643	26.735	26.540	26.735	26.735

while also being much more efficient than the others.

Figure 100 – Boxplots for the NRMSE for the maximization of the buckling load of a 2-ply general laminate.



Source: the author

Table 36 shows the results for the 10-ply case. This time, the problem is even harder due to the higher number of variables. Note that, even though the average NRMSE is relatively high for all approaches, the average λ_{norm} of almost all approaches were still higher than the one from the literature, shown in Table 33, where $\lambda_{opt} = 37.348$ when evaluated via the FSDT. The exception were the PI-based approaches and the HKRG-VF-LCB. Here, MFMs were not more efficient than the single-fidelity Kriging model. This is because the sample size for this problem is higher, which means that model building and evaluation are more expensive. Still, MFMs required less HF evaluations, on average, to find the optimum. Since the cost for each evaluation was not very expensive, it did not make up for the higher model complexity. However, in terms of accuracy, MFMs performed exceptionally well when compared to the KRG model. In particular, the HKRG-LCB and the HKRG-WEI showed a very low average NRMSE when compared to the other approaches. In terms of efficiency, the HKRG-VF-WEI outperformed the other approaches. However, it lost in terms of accuracy when compared to the COKRG-EI, the HKRG-LCB, the HKRG-EI, and the HKRG-WEI.

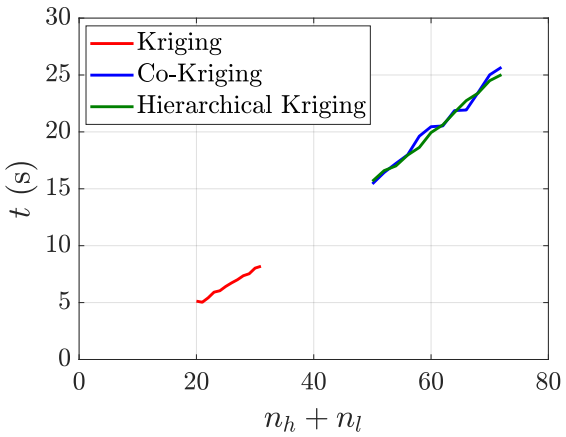
Table 36 – Averaged results for the maximization of the buckling load of a 10-ply symmetric laminate.

Model	Infill criterion	λ_{norm}	NRMSE	$n_{ev,h}$	$n_{ev,l}$	Time spent (s)
KRG	LCB	37.561	20.9%	36 (16)	-	1415
	PI	36.784	22.6%	43 (23)	-	1720
	EI	37.617	20.8%	42 (22)	-	1739
	WEI	37.768	20.5%	36 (16)	-	1394
COKRG	LCB	39.678	16.5%	31 (21)	61 (21)	1760
	PI	37.199	21.7%	31 (21)	61 (21)	1738
	EI	41.084	13.5%	31 (21)	61 (21)	1833
	WEI	40.081	15.6%	30 (20)	60 (20)	1702
HKRG	LCB	44.390	6.5%	32 (22)	62 (22)	1807
	PI	36.657	22.8%	28 (18)	58 (18)	1442
	EI	41.747	12.1%	35 (25)	65 (25)	2037
	WEI	44.559	6.2%	33 (23)	63 (23)	1919
	VF-LCB	32.594	31.4%	10 (0)	52 (12)	672
	VF-PI	37.092	21.9%	26 (16)	57 (17)	1363
	VF-EI	39.364	17.1%	17 (7)	61 (21)	1206
	VF-WEI	40.540	14.6%	14 (4)	55 (15)	862

Figure 101 shows the model building cost for different models. Again, model building cost is much higher on MFMs due to the higher sample size. However, this also contributes to their major gain in accuracy. Therefore, there is a clear trade-off here. While the process for the Kriging model starts with only 20 HF data points, MFMs have 10 HF and 40 LF

points, which means that $n_h + n_l = 50$. Again, model complexity is entirely dependent on the sample size.

Figure 101 – Time spent for the building phase for each model for the maximization of the buckling load of a 10-ply symmetric laminate.

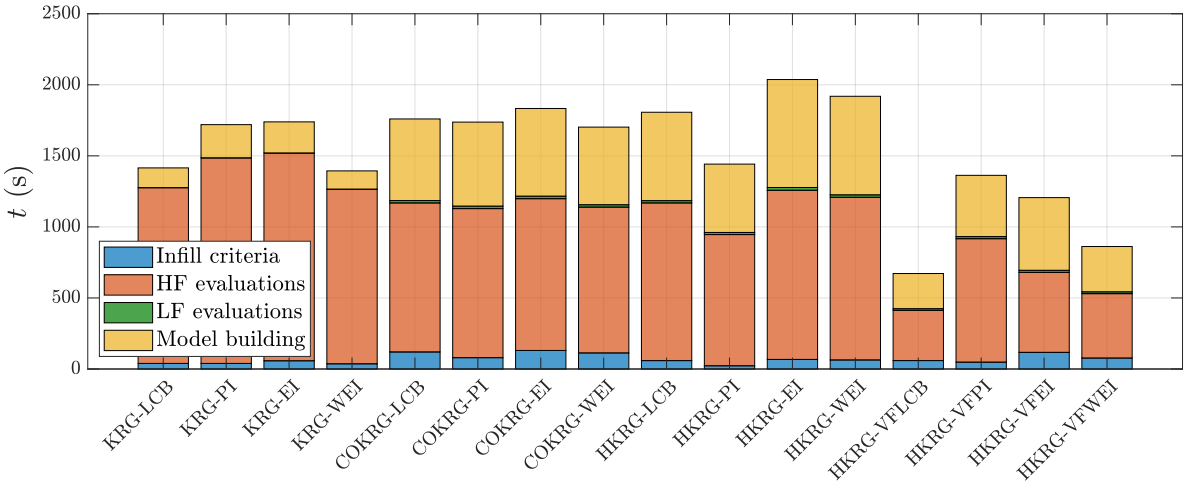


Source: the author

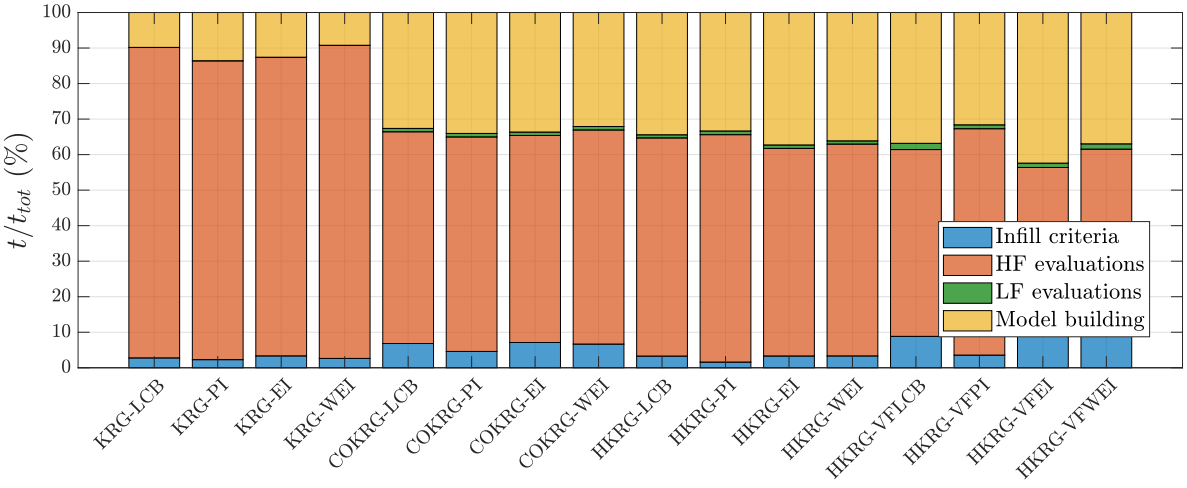
This becomes even more clear in Figure 102. Again, we can not see any gain in efficiency on MFMs, but their gain in accuracy is very noticeable. Still, the VF methods, presented a lower accuracy in this problem, once again. That being said, the VF-WEI presented a major reduction in computational cost, while also presenting a gain in accuracy when compared to some non-VF methods. For this problem, on most MFMs, model building represented almost half of the total process cost. However, one should note that these models still showed a reduction on the number of HF evaluations performed. If HF analysis cost was higher, building cost could be less important for the total optimization cost. Thus, MFMs would be able to show a gain in efficiency even for this problem.

Table 37 shows the best response found by each method for the $N_r = 10$ runs performed. Note that the best design shown in Table 33 was found by the COKRG-PI method. However, a variety of approaches managed to find very close to the optimum designs. The exceptions were the HKRG-PI and the HKRG-VF-LCB. In particular, the latter struggled at defining when a HF data points should be added. This can be seem on Table 36, where, on average, no HF data points was added during the entire optimization process. Finally, Figure 103 shows the boxplots for the NRMSE for all approaches. Here, we can see that the HKRG-LCB and the HKRG-WEI were, indeed, the more robust approaches, showing lower errors in most cases.

Figure 102 – Cost of each phase of the process for the maximization of the buckling load of a 10-ply symmetric laminate.



(a) Total time spent in each phase



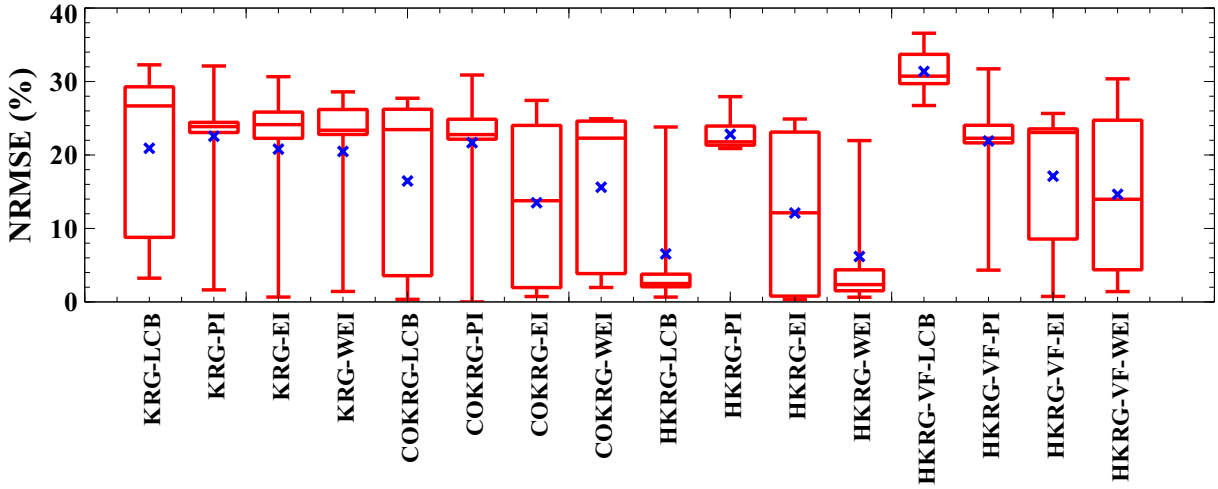
(b) Percentage contribution relative to the total cost of each method

Source: the author

Table 37 – Highest buckling load ($\lambda_{norm,best}$) found by each approach for a 10-ply symmetric laminate.

Model	LCB	PI	EI	WEI	VF-LCB	VF-PI	VF-EI	VF-WEI
KRG	45.959	46.716	47.180	46.818	-	-	-	-
COKRG	47.328	47.495	47.147	46.561	-	-	-	-
HKRG	47.181	37.579	47.324	47.192	34.799	45.440	47.140	46.827

Figure 103 – Boxplots for the NRMSE for the maximization of the buckling load of a 10-ply symmetric laminate.



Source: the author

8 CONCLUSION

This work investigated the efficient optimization of composite structures using Multi-Fidelity Models (MFMs). Two different models, the Co-Kriging and the Hierarchical Kriging, were compared, along with the single-fidelity Kriging model for comparison purposes. Four different acquisition functions were compared: the Lower Confidence Bound (LCB), the Probability of Improvement (PI), the Expected Improvement (EI), and the Weighted Expected Improvement (WEI). For the Hierarchical Kriging, Variable Fidelity (VF) versions of these functions were also tried out, where an effective method was applied for the selection of the fidelity of the new point. When necessary, different constraint-handling methods were also compared, namely the Direct approach, the Probability of Feasibility (PF), and the feasibility functions proposed by Tutum, Deb and Baran[62], Bagheri et al.[63], and Sohst, Afonso and Suleman[297]. These feasibility functions are denominated here as FFT, FFB, and FFS, respectively.

To allow for an effective study over the approaches employed in this work, methods for optimization using MFMs were implemented in BIOS, an academic software developed in C++ on LMCV. Implementation included the creation of classes related to MFMs, namely Co-Kriging and Hierarchical Kriging. Different acquisition functions and constraint-handling methods not yet present on BIOS were added.

Important extensions were also added to the analysis software FAST, also developed in C++ by LMCV collaborators. These allowed for the adequate analysis of FG and laminate structures. Structural responses were assessed using the Isogeometric Analysis (IGA), where cubic NURBS meshes were used. Both 3D solid and 2D theories were employed. For the latter, a shallow shell element based on the Reissner-Mindlin plate theory and Marguerre strains was employed[82].

In this work, applications ranged from simple mathematical benchmarks, to engineering analytical optimization problems, to numerical FG and laminate problems. Here, special focus was given to eigenvalue problems, but the examples managed to study a wide range of relevant aspects for the efficient optimization using MFMs. The effect of dimensionality, correlation between sources, and average error between sources was evaluated. Also, optimization using expensive constraints and discrete designs was performed with great success.

Co-Kriging behaved well, showing better accuracy than the Kriging model. However, we showed that, when the LF source is related to the HF source by the addition of a constant factor, the model is not able to provide accurate approximations. Also, there is no promising

method to add points in an efficient manner. Hierarchical Kriging showed a high average NRMSE for some problems when using non-VF EI or WEI. However, the VF-EI and the VF-WEI showed very good results for most problems, as well as the VF-LCB approach. These also often showed a major reduction in computational time, since the number of HF evaluations required was reduced even further. These methods, however, did not seem to provide any improvement over non-VF methods when sources present a lower correlation.

Regarding adaptive sampling criteria, the LCB method presented a very good accuracy, showing the lowest average NRMSE for most problems. However, it was often the most time consuming method, since it often required more HF evaluations. The EI and the WEI also showed good results. The PI, however, rarely managed to achieve low average NRMSE values, and it was unable to find the optimum for the majority of the analytical and numerical problems.

Regarding optimization using expensive constraints, the LCB method was not able to find the optimum. Here, we considered the adaptive penalty approach, using the model of the constraint function to assess the feasibility of the design. However, the method was not able to show consistent results. Appropriate methods should be derived and tested in future works. For instance, one may try to adapt probabilistic infill criteria to the LCB[297]. On the other hand, excellent results were found using the EI, especially when considering the Probability of Feasibility (PF) or the feasibility functions proposed by Sohst, Afonso and Suleman[297] (with $n = 0.50$).

Optimization considering discrete designs also showed encouraging results using MFMs. Due to the inherent complexity of these problems, it is easier to see the gain in accuracy when comparing MFMs to the single-fidelity Kriging model. Found results were also much better than those from the literature, and the methods seem to be indeed very promising for the optimization of laminate composites. Performing adaptive sampling via the EI showed the most consistent results, but the LCB, the WEI, and even the PI showed good results in many cases. However, further adaptations could still be made to the methodology to better deal with discrete spaces.

In general, Multi-Fidelity Models (MFMs) required a lower number of HF evaluations to find results closer to the optimum. Good results were also found in higher dimensional problems, in optimization considering expensive constraints, and in discrete optimization. On the other hand, depending on the cost of the HF analysis, one may not see a gain in efficiency. This occurs because, since MFMs require more sampling points, model complexity is much

higher. Thus, the cost for evaluating and, especially, building the model is much higher.

The following suggestions can be further explored in future works:

a) Gradually change correlation and/or average relative difference between sources to better understand their importance for the optimization process;
b) Use different ratios of HF and LF sampling points for the same problem;
c) Further test and study how different constraint-handling methods behave for SBO using MFMs. In particular, the feasibility function proposed by Sohst, Afonso and Suleman[297] seems very promising, but further testing with different n is still required;
d) Implement and test different approaches for performing optimization with expensive constraints using the LCB;
e) Implement and test more appropriate methods for dealing with discrete individuals using GP models;
f) Optimize laminate problems considering the material and the thickness as design variables;
g) Implement and test different approaches to assist in optimizing higher dimensional problems, such as Proper Orthogonal Decomposition (POD);
h) Use MFMs for the optimization of Variable Stiffness Composites (VSCs);
i) Use appropriate methods to perform multi-objective optimization using MFMs.

BIBLIOGRAPHY

1 DO, D.; LEE, D.; LEE, J. Material optimization of functionally graded plates using deep neural network and modified symbiotic organisms search for eigenvalue problems. *Composites Part B: Engineering*, v. 159, p. 300–326, 2019.

2 RIBEIRO, L. G. et al. Surrogate based optimization of functionally graded plates using radial basis functions. *Composite Structures*, v. 252, 2020.

3 MAIA, M. A. *Sequential Approximate Optimization of Composite Structures*. Dissertação (Mestrado) — Universidade Federal do Ceará, 2020.

4 ILCEWICZ, L.; HOFFMAN, D.; FAWCETT, A. Composite applications in commercial airframe structures. *Comprehensive Composite Materials*, p. 87–119, 2000.

5 NIKBAKT, S.; KAMARIAN, S.; SHAKERI, M. A review on optimization of composite structures Part I: Laminated Composites. *Composite Structures*, v. 195, p. 158–185, 2018. Available from Internet: https://doi.org/10.1016/j.compstruct.2018.03.063.

6 CAMPBELL, F. C. *Structural composite materials*. [S.l.]: ASM International, 2010.

7 REDDY, J. N.; REDDY, J. N. *Mechanics of laminated composite plates and shells: theory and analysis*. [S.l.]: CRC Press, 2004.

8 SHEN, H. S. *Functionally graded materials: Nonlinear analysis of plates and shells*. [S.l.]: CRC Press, 2009.

9 SHARMA, K.; KUMAR, D. Elastoplastic stability and failure analysis of FGM plate with temperature dependent material properties under thermomechanical loading. *Latin American Journal of Solids and Structures*, v. 14, n. 7, p. 1361–1386, 2017. ISSN 16797825.

10 JHA, D.; KANT, T.; SINGH, R. A critical review of recent research on functionally graded plates. *Composite Structures*, v. 96, p. 833–849, 2013.

11 KOIZUMI, M. FGM activities in japan. *Composites Part B: Engineering*, v. 28, n. 1-2, p. 1–4, 1997.

12 UDUPA, G.; RAO, S. S.; GANGADHARAN, K. Functionally graded composite materials: An overview. *Procedia Materials Science*, v. 5, p. 1291–1299, 2014.

13 NAEBE, M.; SHIRVANIMOGHADDAM, K. Functionally graded materials: A review of fabrication and properties. *Applied Materials Today*, Elsevier Ltd, v. 5, p. 223–245, 2016. ISSN 23529407. Available from Internet: http://dx.doi.org/10.1016/j.apmt.2016.10.001.

14 NIKBAKHT, S.; KAMARIAN, S.; SHAKERI, M. A review on optimization of composite structures part ii: Functionally Graded Materials. *Composite Structures*, v. 214, p. 83–102, 2019.

15 ROCHA, I. B.; PARENTE, E.; MELO, A. M. C. A hybrid shared/distributed memory parallel genetic algorithm for optimization of laminate composites. *Composite Structures*, v. 107, n. 1, p. 288–297, 2014.

16 BARROSO, E. S.; PARENTE, E.; MELO, A. M. C. A hybrid PSO-GA algorithm for optimization of laminated composites. *Structural and Multidisciplinary Optimization*, v. 55, n. 6, p. 2111–2130, 2017.

17 WANG, C. et al. Adaptive chaotic particle swarm algorithm for isogeometric multi-objective size optimization of FG plates. *Structural and Multidisciplinary Optimization*, v. 60, n. 2, p. 757–778, 2019.

18 DO, D. T.; NGUYEN-XUAN, H.; LEE, J. Material optimization of tri-directional functionally graded plates by using deep neural network and isogeometric multimesh design approach. *Applied Mathematical Modelling*, Elsevier Inc., v. 87, n. 107, p. 501–533, 2020. ISSN 0307904X. Available from Internet: https://doi.org/10.1016/j.apm.2020.06.002.

19 KOGISO, N. et al. Genetic algorithms with local improvement for composite laminate design. *Structural Optimization*, v. 7, n. 4, p. 207–218, 1994.

20 JONES, R. M. *Mechanics of composite materials*. [S.l.]: McGraw-Hill, 1999.

21 LI, X.-F.; PENG, X.-L. A pressurized functionally graded hollow cylinder with arbitrarily varying material properties. *Journal of Elasticity*, v. 96, n. 1, p. 81–95, 2009.

22 BOUAZZA, M. et al. Thermoelastic stability analysis of functionally graded plates: An analytical approach. *Computational Materials Science*, v. 49, n. 4, p. 865–870, 2010. ISSN 0927-0256. Available from Internet: https://www.sciencedirect.com/science/article/pii/S092702561000385X.

23 SOFIYEV, A. Review of research on the vibration and buckling of the FGM conical shells. *Composite Structures*, v. 211, p. 301–317, 2019.

24 GU, L. A Comparison of Polynomial Based Regression Models in Vehicle Safety Analysis. In: . [s.n.], 2001. (International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Volume 2A: 27th Design Automation Conference), p. 509–514. Available from Internet: https://doi.org/10.1115/DETC2001/DAC-21063.

25 HAFTKA, R. T.; GüRDAL, Z. Elements of structural optimization. *Solid Mechanics And Its Applications*, 1992.

26 KOU, X.; PARKS, G.; TAN, S. Optimal design of functionally graded materials using a procedural model and particle swarm optimization. *Computer-Aided Design*, v. 44, n. 4, p. 300–310, 2012.

27 LAGAROS, N. D.; PAPADRAKAKIS, M.; KOKOSSALAKIS, G. Structural optimization using evolutionary algorithms. *Computers and Structures*, v. 80, n. 7-8, p. 571–589, 2002.

28 NABIAN, M.; AHMADIAN, M. T. Multi-objective optimization of functionally graded hollow cylinders. *Volume 8: Mechanics of Solids, Structures and Fluids; Vibration, Acoustics and Wave Propagation*, 2011.

29 TAATI, E.; SINA, N. Multi-objective optimization of functionally graded materials, thickness and aspect ratio in micro-beams embedded in an elastic medium. *Structural and Multidisciplinary Optimization*, v. 58, n. 1, p. 265–285, 2018.

30 LOJA, M. A. R. On the use of particle swarm optimization to maximize bending stiffness of functionally graded structures. *Journal of Symbolic Computation*, v. 61-62, p. 12–30, 2014.

31 KIM, N. I. et al. Nurbs-based optimization of natural frequencies for bidirectional functionally graded beams. *Archives of Mechanics*, v. 70, n. 4, p. 337–364, 2018.

32 TRUONG, T. T.; NGUYEN-THOI, T.; LEE, J. Isogeometric size optimization of bi-directional functionally graded beams under static loads. *Composite Structures*, v. 227, p. 111259, 2019.

33 TRUONG, T. T.; LEE, S.; LEE, J. An artificial neural network-differential evolution approach for optimization of bidirectional functionally graded beams. *Composite Structures*, v. 233, 2020.

34 RIBEIRO, L. G. et al. Surrogate based optimization of functionally graded plates using PSO and DE. *XLI Ibero-Latin American Congress on Computational Methods in Engineering*, 2020.

35 REGIS, R. G.; SHOEMAKER, C. A. Constrained global optimization of expensive black box functions using radial basis functions. *Journal of Global Optimization*, v. 31, n. 1, p. 153–171, 2005.

36 STEPONAVIčė, I. et al. On sampling methods for costly multi-objective black-box optimization. *Advances in Stochastic and Deterministic Global Optimization Springer Optimization and Its Applications*, p. 273–296, 2016.

37 ZHAN, D.; XING, H. Expected improvement for expensive optimization: a review. *Journal of Global Optimization*, Springer US, 2020. Available from Internet: https://doi.org/10.1007/s10898-020-00923-x.

38 VALADÃO, M. A.; BATISTA, L. S. A comparative study on surrogate models for SAEAs. *Optimization Letters*, Springer Berlin Heidelberg, v. 14, n. 8, p. 2595–2614, 2020. ISSN 18624480. Available from Internet: https://doi.org/10.1007/s11590-020-01575-2.

39 VENKATARAMAN, S.; HAFTKA, R. T. Structural optimization complexity: What has Moore's law done for us? *Structural and Multidisciplinary Optimization*, v. 28, n. 6, p. 375–387, 2004. ISSN 1615147X.

40 FORRESTER, A. I. J.; SOBESTER, A.; KEANE, A. J. *Engineering design via surrogate modelling: a practical guide*. [S.l.]: Wiley, 2008.

41 BISHOP, C. M. *Pattern Recognition and Machine Learning*. [S.l.]: Springer New York, 2016.

42 TRUONG, T. T.; LEE, J.; NGUYEN-THOI, T. Multi-objective optimization of multi-directional functionally graded beams using an effective deep feedforward neural network-SMPSO algorithm. *Structural and Multidisciplinary Optimization*, Springer, 2021.

43 ZHOU, Q. et al. An adaptive global variable fidelity metamodeling strategy using a support vector regression based scaling function. *Simulation Modelling Practice and Theory*, Elsevier B.V., v. 59, p. 18–35, 2015. ISSN 1569190X. Available from Internet: http://dx.doi.org/10.1016/j.simpat.2015.08.002.

44 XIANG, H. et al. An adaptive surrogate model based on support vector regression and its application to the optimization of railway wind barriers. *Structural and Multidisciplinary Optimization*, v. 55, n. 2, p. 701—-713, 2016.

45 SHI, M. et al. A multi-fidelity surrogate model based on support vector regression. *Structural and Multidisciplinary Optimization*, Springer, v. 61, n. 6, p. 2363–2375, 2020. ISSN 16151488.

46 MONGILLO, M. Choosing basis functions and shape parameters for radial basis function methods. *SIAM Undergraduate Research Online*, v. 4, p. 190–209, 2011.

47 KITAYAMA, S.; YAMAZAKI, K. Simple estimate of the width in gaussian kernel with adaptive scaling technique. *Applied Soft Computing*, v. 11, n. 8, p. 4726–4737, 2011.

48 FOUAIDI, M.; JAMAL, M.; BELOUAGGADIA, N. Nonlinear bending analysis of functionally graded porous beams using the multiquadric radial basis functions and a Taylor series-based continuation procedure. *Composite Structures*, Elsevier Ltd, v. 252, 2020. ISSN 02638223. Available from Internet: https://doi.org/10.1016/j.compstruct.2020.112593.

49 PALAR, P. S.; SHIMOYAMA, K. On efficient global optimization via universal kriging surrogate models. *Structural and Multidisciplinary Optimization*, v. 57, n. 6, p. 2377–2397, 2017.

50 CHENG, Y. C.; JIANG, C. P.; LIN, D. H. Finite element based optimization design for a one-piece zirconia ceramic dental implant under dynamic loading and fatigue life validation. *Structural and Multidisciplinary Optimization*, v. 59, n. 3, p. 835–849, 2018.

51 KESHTEGAR, B. et al. Optimization of buckling load for laminated composite plates using adaptive Kriging-improved PSO: A novel hybrid intelligent method. *Defence Technology*, Elsevier Ltd, v. 17, 2020.

52 SONG, X. et al. A radial basis function-based multi-fidelity surrogate model: exploring correlation between high-fidelity and low-fidelity models. *Structural and Multidisciplinary Optimization*, 2019.

53 JONES, D. R.; SCHONLAU, M.; WELCH, W. J. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, p. 455–492, 1998.

54 GUTMANN, H. M. A radial basis function method for global optimization. *Journal of Global Optimization*, v. 19, n. 3, p. 201–227, 2001. ISSN 09255001.

55 SOBESTER, A.; LEARY, S. J.; KEANE, A. J. On the design of optimization strategies based on global response surface approximation models. *Journal of Global Optimization*, p. 31–59, 2005.

56 KITAYAMA, S.; ARAKAWA, M.; YAMAZAKI, K. Sequential approximate optimization using radial basis function network for engineering optimization. *Optimization and Engineering*, v. 12, n. 4, p. 535–557, 2010.

57 MAIA, M. A. et al. Sequential approximate optimization using kriging and radial basis functions. *XL Ibero-Latin American Congress on Computational Methods in Engineering*, 2019. Submitted for publication.

58 STORK, J. et al. *Open Issues in Surrogate-Assisted Optimization*. Springer International Publishing, 2020. 225–244 p. ISBN 9783030187644. Available from Internet: http://dx.doi.org/10.1007/978-3-030-18764-4_10.

59 BARTZ-BEIELSTEIN, T.; ZAEFFERER, M. Model-based methods for continuous and discrete global optimization. *Applied Soft Computing*, Elsevier B.V., v. 55, p. 154–167, 2017. ISSN 15684946. Available from Internet: http://dx.doi.org/10.1016/j.asoc.2017.01.039.

60 SCHONLAU, M.; WELCH, W. J.; JONES, D. R. Global versus local search in constrained optimization of computer models. *New developments and applications in experimental design*, Institute of Mathematical Statistics, Hayward, CA, Volume 34, p. 11–25, 1998. Available from Internet: https://doi.org/10.1214/lnms/1215456182.

61 SASENA, M. J.; PAPALAMBROS, P.; GOOVAERTS, P. Exploration of metamodeling sampling criteria for constrained global optimization. *Engineering Optimization*, Taylor and Francis, v. 34, n. 3, p. 263–278, 2002. Available from Internet: https://doi.org/10.1080/03052150211751.

62 TUTUM, C. C.; DEB, K.; BARAN, I. Constrained efficient global optimization for pultrusion process. *Materials and Manufacturing Processes*, v. 30, n. 4, p. 538–551, 2014.

63 BAGHERI, S. et al. Constraint handling in efficient global optimization. *Proceedings of the Genetic and Evolutionary Computation Conference*, v. 17, p. 673–680, 2017.

64 DíAZ-MANRíQUEZ, A.; TOSCANO-PULIDO, G.; GóMEZ-FLORES, W. On the selection of surrogate models in evolutionary optimization algorithms. In: *2011 IEEE Congress of Evolutionary Computation (CEC)*. [S.l.: s.n.], 2011. p. 2155–2162.

65 NIK, M. A. et al. A comparative study of metamodeling methods for the design optimization of variable stiffness composites. *Composite Structures*, v. 107, p. 494 – 501, 2014. ISSN 0263-8223. Available from Internet: http://www.sciencedirect.com/science/article/pii/S0263822313004200.

66 GAN, N.; GU, J. Hybrid meta-model-based design space exploration method for expensive problems. *Structural and Multidisciplinary Optimization*, v. 59, n. 3, p. 907–917, 2018.

67 FORRESTER, A. I.; SóBESTER, A.; KEANE, A. J. Multi-fidelity optimization via surrogate modelling. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, v. 463, n. 2088, p. 3251–3269, 2007.

68 PARUSSINI, L. et al. Multi-fidelity Gaussian process regression for prediction of random fields. *Journal of Computational Physics*, Elsevier Inc., v. 336, p. 36–50, 2017. ISSN 10902716. Available from Internet: http://dx.doi.org/10.1016/j.jcp.2017.01.047.

69 YONG, H. K. et al. Multi-fidelity Kriging-assisted structural optimization of whole engine models employing medial meshes. Structural and Multidisciplinary Optimization, p. 1209–1226, 2019.

70 LI, X. et al. A cooperative radial basis function method for variable-fidelity surrogate modeling. *Structural and Multidisciplinary Optimization*, v. 56, n. 5, p. 1077–1092, 2017.

71 KENNEDY, M. C.; O'HAGAN, A. Predicting the output from a complex computer code when fast approximations are available. *Biometrika*, [Oxford University Press, Biometrika Trust], v. 87, n. 1, p. 1–13, 2000. ISSN 00063444. Available from Internet: http://www.jstor.org/stable/2673557.

72 HAN, Z.-H.; GöRTZ, S. Hierarchical kriging model for variable-fidelity surrogate modeling. *AIAA Journal*, v. 50, n. 9, p. 1885–1896, 2012. Available from Internet: https://doi.org/10.2514/1.J051354.

73 ZHANG, Y.; HAN, Z. H.; ZHANG, K. S. Variable-fidelity expected improvement method for efficient global optimization of expensive functions. *Structural and Multidisciplinary Optimization*, Springer, v. 58, n. 4, p. 1431–1451, 2018. ISSN 16151488.

74 RUAN, X. et al. An improved Co-Kriging multi-fidelity surrogate modeling method for non-nested sampling data. *International Journal of Mechanical Engineering and Robotics Research*, v. 8, n. 4, p. 559–564, 2019. ISSN 22780149.

75 GHOSH, S. et al. A strategy for adaptive sampling of multi-fidelity Gaussian process to reduce predictive uncertainty. *arXiv*, p. 1–14, 2019. ISSN 23318422.

76 RUAN, X. et al. Variable-fidelity probability of improvement method for efficient global optimization of expensive black-box problems. *Structural and Multidisciplinary Optimization*, Springer, v. 62, n. 6, p. 3021–3052, 2020. ISSN 16151488.

77 KONTOGIANNIS, S. G.; SAVILL, M. A. A generalized methodology for multidisciplinary design optimization using surrogate modelling and multifidelity analysis. *Optimization and Engineering*, Springer US, v. 21, 2020. Available from Internet: https://doi.org/10.1007/s11081-020-09504-z.

78 GUO, Q. et al. Design optimization of variable stiffness composites by using multi-fidelity surrogate models. *Structural and Multidisciplinary Optimization*, v. 63, 2021.

79 GUO, Q. et al. Buckling optimization of variable stiffness composite cylinders by using multi-fidelity surrogate models. *Thin-Walled Structures*, v. 156, p. 107014, 2020. ISSN 0263-8231. Available from Internet: https://www.sciencedirect.com/science/article/pii/S0263823120308892.

80 GüRDAL RAPHAEL T. HAFTKA, P. H. Z. *Design and optimization of laminated composite materials*. [S.l.]: John Wiley and Sons, 1999.

81 MALLICK, P. K. *Fiber-reinforced composites: materials, manufacturing, and design*. [S.l.]: CRC Press, 2008.

82 PRACIANO, J. S. C. et al. An isogeometric formulation for stability analysis of laminated plates and shallow shells. *Thin-Walled Structures*, v. 143, p. 106224, 2019.

83 HINTON, M. J.; KADDOUR, A. S.; SODEN, P. D. *Failure criteria in fibre reinforced polymer composites: the World-Wide Failure Exercise*. [S.l.]: Elsevier, 2004.

84 DANIEL, I. M.; ISHAI, O. *Engineering mechanics of composite materials*. [S.l.]: Oxford University Press, 2007.

85 CORREIA, V. M. F. et al. Multiobjective optimization of functionally graded material plates with thermo-mechanical loading. *Composite Structures*, v. 207, p. 845–857, 2019.

86 GAUTAM, M.; CHATURVEDI, M. Optimization of FGM composition for better environment material. *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, v. 1017, p. 012024, jan 2021. Available from Internet: https://doi.org/10.1088/1757-899x/1017/1/012024.

87 KUMAR HS, N.; KATTIMANI, S.; NGUYEN-THOI, T. Influence of porosity distribution on nonlinear free vibration and transient responses of porous functionally graded skew plates. *Defence Technology*, China Ordnance Society, 2021. ISSN 22149147.

88 EL-HADAD, S. et al. Fabrication of al-al3ti/ti3al functionally graded materials under a centrifugal force. *Materials*, v. 3, n. 9, p. 4639–4656, 2010.

89 ROQUE, C.; MARTINS, P. Differential evolution for optimization of functionally graded beams. *Composite Structures*, v. 133, p. 1191–1197, 2015. ISSN 0263-8223. Available from Internet: https://www.sciencedirect.com/science/article/pii/S0263822315007187.

90 TSIATAS, G. C.; CHARALAMPAKIS, A. E. Optimizing the natural frequencies of axially functionally graded beams and arches. *Composite Structures*, v. 160, p. 256–266, 2017. ISSN 0263-8223. Available from Internet: https://www.sciencedirect.com/science/article/pii/S0263822316312727.

91 ASHJARI, M.; KHOSHRAVAN, M. Mass optimization of functionally graded plate for mechanical loading in the presence of deflection and stress constraints. *Composite Structures*, v. 110, p. 118–132, 2014.

92 PIEGL, L. A.; TILLER, W. *The NURBS book*. [S.l.]: Springer, 1997. 646 p. ISBN 3540615458.

93 TOULOUKIAN, Y. S.; GERRITSEN, J. K.; SHAFER, W. H. *Thermophysical Properties Research Literature Retrieval Guide*. [S.l.]: Springer, 1974.

94 AKBARZADEH, A. H.; ABEDINI, A.; CHEN, Z. T. Effect of micromechanical models on structural responses of functionally graded plates. *Composite Structures*, Elsevier Ltd, v. 119, p. 598–609, 2015. ISSN 02638223. Available from Internet: http://dx.doi.org/10.1016/j.compstruct.2014.09.031.

95 MEDEIROS, M. S.; PARENTE, E.; MELO, A. M. C. D. Influence of the micromechanics models and volume fraction distribution on the overall behavior of sic/al functionally graded pressurized cylinders. *Latin American Journal of Solids and Structures*, v. 16, n. 4, 2019.

96 CORREIA, V. F. et al. Optimization of metal–ceramic functionally graded plates using the simulated annealing algorithm. *Applied Sciences (Switzerland)*, v. 11, n. 2, p. 1–13, 2021. ISSN 20763417.

97 TAMURA, I. Strength and ductility of fe-ni-c alloys composed of austenite and martensite with various strength. *Proceedings of the third international conference on strength of metals and alloys, 1973*, Cambridge, Institute of Metals, v. 1, p. 611–615, 1973. Available from Internet: https://ci.nii.ac.jp/naid/80013184461/en/.

98 ARSLAN, K. et al. Evaluation of geometrically nonlinear and elastoplastic behavior of functionally graded plates under mechanical loading–unloading. *Mechanics of Advanced Materials and Structures*, Taylor and Francis, p. 1–14, 2020. Available from Internet: https://doi.org/10.1080/15376494.2020.1829760.

99 GIANNAKOPOULOS, A. et al. Elastoplastic analysis of thermal cycling: layered materials with compositional gradients. *Acta Metallurgica et Materialia*, v. 43, n. 4, p. 1335–1354, 1995. ISSN 0956-7151. Available from Internet: https://www.sciencedirect.com/science/article/pii/095671519400360T.

100 JIN, Z.-H.; PAULINO, G. H.; DODDS, R. H. Cohesive fracture modeling of elastic–plastic crack growth in functionally graded materials. *Engineering Fracture Mechanics*, v. 70,

n. 14, p. 1885–1912, 2003. ISSN 0013-7944. Cohesive Models. Available from Internet: https://www.sciencedirect.com/science/article/pii/S0013794403001309.

101 GUNES, R. et al. The elasto-plastic impact analysis of functionally graded circular plates under low-velocities. *Composite Structures*, v. 93, n. 2, p. 860–869, 2011. ISSN 0263-8223. Available from Internet: https://www.sciencedirect.com/science/article/pii/S0263822310002588.

102 HUGHES, T.; COTTRELL, J.; BAZILEVS, Y. Isogeometric analysis: Cad, finite elements, nurbs, exact geometry and mesh refinement. *Computer Methods in Applied Mechanics and Engineering*, v. 194, n. 39, p. 4135–4195, 2005. ISSN 0045-7825. Available from Internet: https://www.sciencedirect.com/science/article/pii/S0045782504005171.

103 DANTAS JR., E. M. *Análise não linear de compósitos laminados utilizando o método dos elementos finitos*. Dissertação (Mestrado) — Universidade Federal do Ceará, 2014.

104 BARROSO, E. S. *Análise e otimização de estruturas laminadas utilizando a formulação isogeométrica*. Dissertação (Mestrado) — Universidade Federal do Ceará, 2015.

105 AUAD, S. P. et al. Isogeometric analysis of fgm plates. *Materials Today: Proceedings*, v. 8, p. 738–746, 2019.

106 LIEU, Q. X. et al. Shape and size optimization of functionally graded sandwich plates using isogeometric analysis and adaptive hybrid evolutionary firefly algorithm. *Thin-Walled Structures*, v. 124, p. 588–604, 2018.

107 COOK, R. D. *Concepts and applications of finite element analysis*. [S.l.]: John Wiley and Sons, 2002.

108 PRACIANO, J. S. C. *Análise da estabilidade de estruturas laminadas e de materiais com gradação funcional utilizando uma formulação isogeométrica*. Dissertação (Mestrado) — Universidade Federal do Ceará, 2018.

109 REISSNER, E. On bending of elastic plates. *Quarterly of Applied Mathematics*, Brown University, v. 5, n. 1, p. 55–68, 1947. ISSN 0033569X, 15524485. Available from Internet: http://www.jstor.org/stable/43633584.

110 BATENI, M.; KIANI, Y.; ESLAMI, M. A comprehensive study on stability of fgm plates. *International Journal of Mechanical Sciences*, v. 75, p. 134–144, 2013.

111 ZHAO, X.; LEE, Y.; LIEW, K. Mechanical and thermal buckling analysis of functionally graded plates. *Composite Structures*, v. 90, n. 2, p. 161–171, 2009. ISSN 0263-8223. Available from Internet: https://www.sciencedirect.com/science/article/pii/S0263822309000725.

112 KIANI, Y.; BAGHERIZADEH, E.; ESLAMI, M. Thermal buckling of clamped thin rectangular fgm plates resting on pasternak elastic foundation (three approximate analytical solutions). *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik*, v. 91, n. 7, p. 581–593, 2011.

113 NGUYEN-THANH, N. et al. Isogeometric analysis using polynomial splines over hierarchical t-meshes for two-dimensional elastic solids. *Computer Methods in Applied Mechanics and Engineering*, v. 200, n. 21, p. 1892–1908, 2011. ISSN 0045-7825. Available from Internet: https://www.sciencedirect.com/science/article/pii/S0045782511000338.

114 SILVA, F. D. P. *Análise Isogeométrica de Placas e Cascas de Material com Gradação Funcional utilizando Elementos de Bézier*. Dissertação (Mestrado) — Universidade Federal do Ceará, 2021.

115 NGUYEN, N.-T. et al. An efficient computational approach for size-dependent analysis of functionally graded nanoplates. *Computer Methods in Applied Mechanics and Engineering*, v. 297, p. 191–218, 2015.

116 NGUYEN, H. M. et al. An alternative approach to avoid overfitting for surrogate models. In: *Proceedings of the 2011 Winter Simulation Conference (WSC)*. [S.l.: s.n.], 2011. p. 2760–2771.

117 SIMULIA. *ABAQUS/Standard user's manual*. Providence, RI, USA: [s.n.], 2012.

118 MATHERN, A. et al. Multi-objective constrained Bayesian optimization for structural design. *Structural and Multidisciplinary Optimization*, 2020. ISSN 16151488.

119 ARORA, J. S. *Introduction to optimum design*. 3. ed. [S.l.]: Academic Press/Elsevier, 2017.

120 SURESH, S.; SUJIT, P.; RAO, A. Particle swarm optimization approach for multi-objective composite box-beam design. *Composite Structures*, v. 81, n. 4, p. 598 – 605, 2007. ISSN 0263-8223. Available from Internet: http://www.sciencedirect.com/science/article/pii/S0263822306003965.

121 PRICE, K. V.; STORN, R. M.; LAMPINEN, J. A. *Differential Evolution: A Pratical Approach to Global Optimization*. [S.l.]: Springer, 2005.

122 LARSON, J.; MENICKELLY, M.; WILD, S. M. Derivative-free optimization methods. *Optimization and Control*, 2019.

123 WANG, D. et al. Structural design employing a sequential approximation optimization approach. *Computers and Structures*, Elsevier Ltd, v. 134, p. 75–87, 2014. Available from Internet: http://dx.doi.org/10.1016/j.compstruc.2013.12.004.

124 OPARA, K. R.; ARABAS, J. Differential evolution: A survey of theoretical analyses. *Swarm and Evolutionary Computation*, v. 44, p. 546–558, 2019.

125 GEN, M.; CHENG, R. *Genetic algorithms and engineering optimization*. [S.l.]: John Wiley and Sons, 2000.

126 KRAMER, O. *Genetic Algorithm Essentials*. [S.l.]: Springer International Publishing, 2017.

127 KENNEDY, J.; EBERHART, R. Particle swarm optimization. In: *Proceedings of ICNN'95 - International Conference on Neural Networks*. [S.l.: s.n.], 1995. v. 4, p. 1942–1948 vol.4.

128 STORN, R.; PRICE, K. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, v. 11, n. 4, p. 341–359, 1997.

129 VO-DUY, T. et al. Multi-objective optimization of laminated composite beam structures using nsga-ii algorithm. *Composite Structures*, v. 168, p. 498–509, 2017.

130 CORREIA, V. M. F. et al. Multiobjective optimization of ceramic-metal functionally graded plates using a higher order model. *Composite Structures*, v. 183, p. 146–160, 2018.

131 ABO-BAKR, H. M. et al. Multi-objective shape optimization for axially functionally graded microbeams. *Composite Structures*, Elsevier Ltd, v. 258, n. August 2020, p. 113370, 2021. ISSN 02638223. Available from Internet: https://doi.org/10.1016/j.compstruct.2020.113370.

132 ABO-BAKR, R. M. et al. Optimal weight for buckling of FG beam under variable axial load using Pareto optimality. *Composite Structures*, Elsevier Ltd, v. 258, n. October 2020, 2021. ISSN 02638223. Available from Internet: https://doi.org/10.1016/j.compstruct.2020.113193.

133 KREMPSER, E.; BARBOSA, H. J. C.; LEMONGE, A. C. C. An Adaptive Constraint Handling Technique for Differential Evolution in Engineering Optimization. 2008.

134 KITAYAMA, S.; ARAKAWA, M.; YAMAZAKI, K. Differential evolution as the global optimization technique and its application to structural optimization. *Applied Soft Computing*, v. 11, n. 4, p. 3792–3803, 2011.

135 LUIS, N. F. et al. Active vibration attenuation in viscoelastic laminated composite panels using multiobjective optimization. *Composites Part B: Engineering*, v. 128, p. 53 – 66, 2017. ISSN 1359-8368. Available from Internet: http://www.sciencedirect.com/science/article/pii/S1359836817303955.

136 CORREIA, V. M. F. et al. Multiobjective design optimization of laminated composite plates with piezoelectric layers. *Composite Structures*, v. 169, p. 10 – 20, 2017. ISSN 0263-8223. Available from Internet: http://www.sciencedirect.com/science/article/pii/S026382231631306X.

137 MOLEIRO, F. et al. Design optimization of functionally graded plates under thermo-mechanical loadings to minimize stress, deformation and mass. *Composite Structures*, Elsevier, v. 245, n. April, p. 112360, 2020. ISSN 02638223. Available from Internet: https://doi.org/10.1016/j.compstruct.2020.112360.

138 KARAKAYA, S.; SOYKASAP, O. Buckling optimization of laminated composite plates using genetic algorithm and generalized pattern search algorithm. *Structural and Multidisciplinary Optimization*, v. 39, n. 5, p. 477–486, 2009.

139 IRISARRI, F.-X. et al. Multiobjective stacking sequence optimization for laminated composite structures. *Composites Science and Technology*, v. 69, n. 7-8, p. 983–990, 2009.

140 LE-MANH, T.; LEE, J. Stacking sequence optimization for maximum strengths of laminated composite plates using genetic algorithm and isogeometric analysis. *Composite Structures*, v. 116, p. 357–363, 2014.

141 SCIUVA, M. D.; GHERLONE, M.; LOMARIO, D. Multiconstrained optimization of laminated and sandwich plates using evolutionary algorithms and higher-order plate theories. *Composite Structures*, v. 59, n. 1, p. 149–154, 2003.

142 CHANG, N. et al. Ply stacking sequence optimization of composite laminate by permutation discrete particle swarm optimization. *Structural and Multidisciplinary Optimization*, v. 41, n. 2, p. 179–187, 2009.

143 ROQUE, C.; MARTINS, P. Maximization of fundamental frequency of layered composites using differential evolution optimization. *Composite Structures*, v. 183, p. 77–83, 2018.

144 VO-DUY, T. et al. Frequency optimization of laminated functionally graded carbon nanotube reinforced composite quadrilateral plates using smoothed FEM and evolution algorithm. *Journal of Composite Materials*, v. 52, n. 14, p. 1971–1986, 2018. ISSN 1530793X.

145 BARGH, H. G.; SADR, M. H. Stacking sequence optimization of composite plates for maximum fundamental frequency using particle swarm optimization algorithm. *Meccanica*, v. 47, n. 3, p. 719–730, 2011.

146 JING, Z. et al. Stacking sequence optimization of composite cylindrical panels by sequential permutation search and Rayleigh-Ritz method. *European Journal of Mechanics, A/Solids*, v. 88, 2021. ISSN 09977538.

147 FAN, H.-T.; WANG, H.; CHEN, X.-H. An optimization method for composite structures with ply-drops. *Composite Structures*, v. 136, p. 650–661, 2016.

148 DEKA, D. J. et al. Multiobjective optimization of laminated composites using finite element method and genetic algorithm. *Journal of Reinforced Plastics and Composites*, v. 24, n. 3, p. 273–285, 2005.

149 LOPEZ, R.; LUERSEN, M.; CURSI, E. Optimization of laminated composites considering different failure criteria. *Composites Part B: Engineering*, v. 40, n. 8, p. 731–740, 2009. ISSN 1359-8368. Available from Internet: https://www.sciencedirect.com/science/article/pii/S1359836809001164.

150 SATHEESH, R.; NAIK, G. N.; GANGULI, R. Conservative design optimization of laminated composite structures using genetic algorithms and multiple failure criteria. *Journal of Composite Materials*, v. 44, n. 3, p. 369–387, 2009.

151 KAMARIAN, S.; SHAKERI, M.; YAS, M. H. Thermal buckling optimisation of composite plates using firefly algorithm. *Journal of Experimental and Theoretical Artificial Intelligence*, v. 29, n. 4, p. 787–794, 2016.

152 SINGHA, M. K.; RAMACHANDRA, L. S.; BANDYOPADHYAY, J. N. Optimum design of laminated composite plates for maximum thermal buckling loads. *Journal of Composite Materials*, v. 34, n. 23, p. 1982–1997, 2000.

153 SPALLINO, R.; THIERAUF, G. Thermal buckling optimization of composite laminates by evolution strategies. *Computers and Structures*, v. 78, n. 5, p. 691–697, 2000.

154 VIJAYACHANDRAN, A. A. et al. Fiber path optimization of a symmetric laminate with a cutout for thermal buckling, using a novel finite element algorithm. *57th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, 2016.

155 KALITA, K.; HALDAR, S.; CHAKRABORTY, S. *A Comprehensive Review on High - Fidelity and Metamodel - Based Optimization of Composite Laminates Box Behnken design*. Springer Netherlands, 2021. ISSN 1886-1784. ISBN 0123456789. Available from Internet: https://doi.org/10.1007/s11831-021-09699-z.

156 DING, S.; WU, C.-P. Optimization of material composition to minimize the thermal stresses induced in fgm plates with temperature-dependent material properties. *International Journal of Mechanics and Materials in Design*, v. 14, n. 4, p. 527–549, 2017.

157 CHIBA, R.; SUGANO, Y. Optimisation of material composition of functionally graded materials based on multiscale thermoelastic analysis. *Acta Mechanica*, v. 223, n. 5, p. 891–909, 2012.

158 NA, K. S.; KIM, J. H. Optimization of volume fractions for functionally graded panels considering stress and critical temperature. *Composite Structures*, v. 89, n. 4, p. 509–516, 2009.

159 HUSSEIN, O.; MULANI, S. B. Two-dimensional optimization of functionally graded material plates subjected to buckling constraints. *58th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, 2017.

160 HUSSEIN, O. S.; MULANI, S. B. Multi-dimensional optimization of functionally graded material composition using polynomial expansion of the volume fraction. *Structural and Multidisciplinary Optimization*, v. 56, n. 2, p. 271–284, 2017.

161 HUANG, J.; RAPOFF, A. J. Optimization design of plates with holes by mimicking bones through nonaxisymmetric functionally graded material. *Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications*, v. 217, n. 1, p. 23–27, 2003.

162 SILVA, T. A. N.; LOJA, M. A. R. Differential evolution on the minimization of thermal residual stresses in functionally graded structures. *Intelligent Systems, Control and Automation: Science and Engineering Computational Intelligence and Decision Making*, p. 289–299, 2012.

163 TORABI, K.; AFSHARI, H. Optimization of flutter boundaries of cantilevered trapezoidal functionally graded sandwich plates. *Journal of Sandwich Structures and Materials*, v. 21, n. 2, p. 503–531, 2017.

164 MOITA, J. S. et al. Material distribution and sizing optimization of functionally graded plate-shell structures. *Composites Part B: Engineering*, v. 142, p. 263–272, 2018.

165 MOLEIRO, F. et al. Thermo-mechanical design optimization of symmetric and non-symmetric sandwich plates with ceramic-metal-ceramic functionally graded core to minimize stress, deformation and mass. *Composite Structures*, p. 114496, 2021. ISSN 0263-8223. Available from Internet: https://www.sciencedirect.com/science/article/pii/S0263822321009582.

166 VEL, S. S.; PELLETIER, J. L. Multi-objective optimization of functionally graded thick shells for thermal loading. *Composite Structures*, v. 81, n. 3, p. 386–400, 2007.

167 SHABANA, Y. M. et al. Stresses minimization in functionally graded cylinders using particle swarm optimization technique. *International Journal of Pressure Vessels and Piping*, v. 154, p. 1–10, 2017.

168 TORNABENE, F.; CERUTI, A. Mixed static and dynamic optimization of four-parameter functionally graded completely doubly curved and degenerate shells and panels using gdq method. *Mathematical Problems in Engineering*, v. 2013, p. 1–33, 2013.

169 OOTAO, Y. et al. Neural network optimization of material composition of a functionally graded material plate at arbitrary temperature range and temperature rise. *Archive of Applied Mechanics (Ingenieur Archiv)*, v. 68, n. 10, p. 662–676, 1998.

170 AKBARI, M. et al. Multicriteria optimization of mechanical properties of aluminum composites reinforced with different reinforcing particles type. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, v. 232, n. 3, p. 323–337, 2017.

171 GARMSIRI, K.; JALAL, M. Multiobjective optimization of composite cylindrical shells for strength and frequency using genetic algorithm and neural networks. *Science and Engineering of Composite Materials*, v. 21, n. 4, p. 529–536, 2014.

172 BLOM, A. W.; STICKLER, P. B.; GüRDAL, Z. Optimization of a composite cylinder under bending by tailoring stiffness properties in circumferential direction. *Composites Part B: Engineering*, v. 41, n. 2, p. 157–165, 2010.

173 NJIM, E. K.; BAKHY, S. H.; AL-WAILY, M. Optimization design of vibration characterizations for functionally graded porous metal sandwich plate structure. *Materials Today: Proceedings*, 2021. ISSN 2214-7853. Available from Internet: https://www.sciencedirect.com/science/article/pii/S2214785321022641.

174 PITTON, S. F.; RICCI, S.; BISAGNI, C. Buckling optimization of variable stiffness cylindrical shells through artificial intelligence techniques. *Composite Structures*, v. 230. ISSN 0263-8223. Available from Internet: https://www.sciencedirect.com/science/article/pii/S0263822319320549.

175 YIN, H. et al. Multiobjective crashworthiness optimization design of functionally graded foam-filled tapered tube based on dynamic ensemble metamodel. *Materials and Design*, v. 55, p. 747–757, 2014.

176 HOLLAND, J. H. *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*. [S.l.: s.n.], 1975.

177 HOLLAND, J. H. Genetic algorithms and adaptation. *Adaptive Control of Ill-Defined Systems*, p. 317–333, 1984.

178 BRATTON, D.; KENNEDY, J. Defining a standard for particle swarm optimization. *2007 IEEE Swarm Intelligence Symposium*, 2007.

179 KENNEDY, J.; MENDES, R. Neighborhood topologies in fully informed and best-of-neighborhood particle swarms. *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, 2006.

180 STORN, R. On the usage of differential evolution for function optimization. *On the usage of differential evolution for function optimization*, 1996.

181 ARUNACHALAM, V. *Optimization using Differential Evolution*. [S.l.], 2008.

182 SILVA, E. K. da; BARBOSA, H. J.; LEMONGE, A. C. An adaptive constraint handling technique for differential evolution with dynamic use of variants in engineering optimization. *Optimization and Engineering*, v. 12, p. 31–54, 2011. ISSN 13894420.

183 MADUREIRA, A. et al. Computational Intelligence and Decision Making : Trends and Applications. *Computational Intelligence and Decision Making*, p. 289–299, 2013.

184 KREMPSER, E. et al. Differential evolution assisted by surrogate models for structural optimization problems. *Proceedings of the Eighth International Conference on Engineering Computational Technology*, v. 100, 2012.

185 MICHALEWICZ, Z.; SCHOENAUER, M. Evolutionary algorithms for constrained parameter optimization problems. *Evolutionary Computation*, v. 4, n. 1, p. 1–32, 1996.

186 LEMONGE, A. C. C.; BARBOSA, H. J. C. An adaptive penalty scheme for genetic algorithms in structural optimization. *International Journal for Numerical Methods in Engineering*, v. 59, n. 5, p. 703–736, 2003.

187 SCHMIT, L.; FARSHI, B. Some approximation concepts for structural synthesis. *AIAA Journal*, v. 12, n. 5, p. 692–699, 1974.

188 JACOBS, J. et al. Framework for sequential approximate optimization. *Structural and Multidisciplinary Optimization*, v. 27, n. 5, 2004.

189 GIUNTA, A. A.; ELDRED, M. S. Implementation of a trust region model management strategy in the DAKOTA optimization toolkit. *8th Symposium on Multidisciplinary Analysis and Optimization*, 2000.

190 VENKATARAMAN, S.; HAFTKA, R. Optimization of composite panels- a review. *American Society for Composites, Technical Conference, 14 th, Fairborn, OH*, p. 479–488, 01 1999.

191 WANG, G. G.; SHAN, S. Review of metamodeling techniques in support of engineering design optimization. *Journal of Mechanical Design*, v. 129, n. 4, p. 370, 2007. ISSN 10500472.

192 SIMPSON, T. W.; LIN, D. K. J.; CHEN, W. Sampling strategies for computer experiments: Design and analysis. *International Journal of Reliability and Applications*, 2002.

193 MATHEW, T. V. et al. Adaptive importance sampling based neural network framework for reliability and sensitivity prediction for variable stiffness composite laminates with hybrid uncertainties. *Composite Structures*, Elsevier, v. 245, 2020. ISSN 02638223. Available from Internet: https://doi.org/10.1016/j.compstruct.2020.112344.

194 TRAN-NGOC, H. et al. Efficient Artificial Neural Networks based on a hybrid metaheuristic optimization algorithm for damage detection in laminated composite structures. *Composite Structures*, Elsevier Ltd, 2020. Available from Internet: https://doi.org/10.1016/j.compstruct.2020.113339.

195 LIU, J. et al. Efficient aerodynamic shape optimization of transonic wings using a parallel infilling strategy and surrogate models. *Structural and Multidisciplinary Optimization*, Springer, v. 55, n. 3, p. 925–943, 2017. ISSN 16151488.

196 GOEL, T. et al. *Ensemble of surrogates*. Springer-Verlag, 2006. Available from Internet: https://link.springer.com/article/10.1007/s00158-006-0051-9.

197 ZHANG, J. et al. Adaptive hybrid surrogate modeling for complex systems. *AIAA Journal*, v. 51, n. 3, p. 643–656, 2013.

198 BHOSEKAR, A.; IERAPETRITOU, M. Advances in surrogate based modeling, feasibility analysis, and optimization: A review. *Computers and Chemical Engineering*, v. 108, p. 250–267, 2018.

199 BROCHU, E.; CORA, V. M.; FREITAS, N. de. *A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning*. 2010.

200 KHALFALLAH, S.; GHENAIET, A. Radial basis function-based shape optimization of centrifugal impeller using sequential sampling. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, v. 229, n. 4, p. 648–665, Oct 2014.

201 GHASSEMI, P.; MEHMANI, A.; CHOWDHURY, S. Adaptive in situ model refinement for surrogate-augmented population-based optimization. *Structural and Multidisciplinary Optimization*, Structural and Multidisciplinary Optimization, v. 62, n. 4, p. 2011–2034, 2020. ISSN 16151488.

202 FUHG, J. N.; FAU, A.; NACKENHORST, U. State-of-the-Art and Comparative Review of Adaptive Sampling Methods for Kriging. *Archives of Computational Methods in Engineering*, Springer Netherlands, v. 28, p. 2689–2747, 2021. ISSN 18861784.

203 MüLLER, J.; SHOEMAKER, C. A. Influence of ensemble surrogate models and sampling strategy on the solution quality of algorithms for computationally expensive black-box global optimization problems. *Journal of Global Optimization*, v. 60, n. 2, p. 123–144, 2014.

204 PALAR, P. S.; SHIMOYAMA, K. Efficient global optimization with ensemble and selection of kernel functions for engineering design. *Structural and Multidisciplinary Optimization*, Structural and Multidisciplinary Optimization, v. 59, n. 1, p. 93–116, 2019. ISSN 16151488.

205 OZCANAN, S.; ATAHAN, A. O. *Minimization of Accident Severity Index in concrete barrier designs using an ensemble of radial basis function metamodel-based optimization*. Springer US, 2020. ISSN 15732924. ISBN 0123456789. Available from Internet: https://doi.org/10.1007/s11081-020-09522-x.

206 RICHE, R. L.; PICHENY, V. Revisiting bayesian optimization in the light of the coco benchmark. *Structural and Multidisciplinary Optimization*, v. 64, n. 5, p. 3063–3087, 2021.

207 ZHU, W. G. et al. Optimization design for laminated composite structure based on kriging model. *Applied Mechanics and Materials*, v. 217-219, p. 179–183, 2012.

208 PASSOS, A. G.; LUERSEN, M. A. Multiobjective optimization of laminated composite parts with curvilinear fibers using Kriging-based approaches. *Structural and Multidisciplinary Optimization*, Structural and Multidisciplinary Optimization, v. 57, n. 3, p. 1115–1127, 2018.

209 JAISWAL, P.; PATEL, J.; RAI, R. Build orientation optimization for additive manufacturing of functionally graded material objects. *International Journal of Advanced Manufacturing Technology*, v. 96, n. 1-4, p. 223–235, 2018.

210 QUEIPO, N. V. et al. Surrogate-based analysis and optimization. *Progress in Aerospace Sciences*, v. 41, n. 1, p. 1–28, 2005. ISSN 03760421.

211 BLANK, J.; DEB, K. Psaf: A probabilistic surrogate-assisted framework for single-objective optimization. In: *Proceedings of the Genetic and Evolutionary Computation Conference*. New York, NY, USA: Association for Computing Machinery, 2021. (GECCO '21), p. 652–659. ISBN 9781450383509. Available from Internet: https://doi.org/10.1145/3449639.3459297.

212 TENNE, Y. Initial sampling methods in metamodel-assisted optimization. *Engineering with Computers*, v. 31, n. 4, p. 661–680, 2014.

213 KLEIJNEN, J. P. C. et al. State-of-the-art review: A user's guide to the brave new world of designing simulation experiments. *INFORMS Journal on Computing*, v. 17, n. 3, p. 263–289, 2005.

214 VU, K. K. et al. Surrogate-based methods for black-box optimization. *International Transactions in Operational Research*, v. 24, n. 3, p. 393–424, 2017. ISSN 14753995.

215 CHO, I. et al. Comparison study of sampling methods for computer experiments using various performance measures. *Structural and Multidisciplinary Optimization*, v. 55, n. 1, p. 221–235, 2016.

216 KIM, B. S.; LEE, Y. B.; CHOI, D. H. Comparison study on the accuracy of metamodeling technique for non-convex functions. *Journal of Mechanical Science and Technology*, v. 23, n. 4, p. 1175–1181, 2009. ISSN 1738494X.

217 TOAL, D. J. et al. The development of a hybridized particle swarm for kriging hyperparameter tuning. *Engineering Optimization*, v. 43, n. 6, p. 675–699, 2011. ISSN 0305215X.

218 SHAN, S.; WANG, G. G. Survey of modeling and optimization strategies to solve high-dimensional design problems with computationally-expensive black-box functions. *Structural and Multidisciplinary Optimization*, v. 41, n. 2, p. 219–241, 2010.

219 KANG, K.; LEE, I. Efficient high-dimensional metamodeling strategy using recursive decomposition coupled with sequential sampling method. *Structural and Multidisciplinary Optimization*, Structural and Multidisciplinary Optimization, 2020.

220 REHBACH, F.; GENTILE, L.; BARTZ-BEIELSTEIN, T. Variable reduction for surrogate-based optimization. *GECCO 2020 - Proceedings of the 2020 Genetic and Evolutionary Computation Conference*, p. 1177–1185, 2020.

221 AMOUZGAR, K.; STRöMBERG, N. Radial basis functions as surrogate models with a priori bias in comparison with a posteriori bias. *Structural and Multidisciplinary Optimization*, v. 55, n. 4, p. 1453–1469, 2016.

222 JIN, R.; CHEN, W.; SIMPSON, T. W. Comparative studies of metamodelling techniques under multiple modelling criteria. *Structural and Multidisciplinary Optimization*, v. 23, n. 1, p. 1–13, 2001. ISSN 1615147X.

223 WANG, C. et al. An evaluation of adaptive surrogate modeling based optimization with two benchmark problems. *Environmental Modelling and Software*, Elsevier Ltd, v. 60, p. 167–179, 2014. ISSN 13648152.

224 De Ath, G. et al. *Greed is Good: Exploration and Exploitation Trade-offs in Bayesian Optimisation*. 2019.

225 WILLIAMS, B.; CREMASCHI, S. Selection of surrogate modeling techniques for surface approximation and surrogate-based optimization. *Chemical Engineering Research and Design*, Institution of Chemical Engineers, v. 170, p. 76–89, 2021. ISSN 02638762. Available from Internet: https://doi.org/10.1016/j.cherd.2021.03.028.

226 HAMMERSLEY, J. M. Monte carlo methods for solving multivariable problems. *Annals of the New York Academy of Sciences*, v. 86, n. 3, p. 844–874, 1960. Available from Internet: https://nyaspubs.onlinelibrary.wiley.com/doi/abs/10.1111/j.1749-6632.1960.tb42846.x.

227 MASCHIO, C.; SCHIOZER, D. J. S. Probabilistic history matching using discrete latin hypercube sampling and nonparametric density estimation. *Journal of Petroleum Science and Engineering*, v. 147, p. 98 – 115, 2016. ISSN 0920-4105. Available from Internet: http://www.sciencedirect.com/science/article/pii/S0920410516301759.

228 MURPHY, K. P. *Machine Learning: A probabilistic perspective*. [S.l.]: MIT Press, 2012.

229 WOLPERT, D.; MACREADY, W. No free lunch theorems for optimization. In: *IEEE Transactions on Evolutionary Computation*. [S.l.: s.n.], 1997. v. 1, n. 1.

230 HUSSAIN, M. F.; BARTON, R. R.; JOSHI, S. B. Metamodeling: Radial basis functions, versus polynomials. *European Journal of Operational Research*, v. 138, n. 1, p. 142–154, 2002. ISSN 03772217.

231 ELSAYED, K.; LACOR, C. Robust parameter design optimization using kriging, rbf and rbfnn with gradient-based and evolutionary optimization techniques. *Applied Mathematics and Computation*, v. 236, p. 325 – 344, 2014. ISSN 0096-3003. Available from Internet: http://www.sciencedirect.com/science/article/pii/S0096300314004445.

232 MEHMANI, A. et al. Concurrent surrogate model selection (cosmos): optimizing model type, kernel function, and hyper-parameters. *Structural and Multidisciplinary Optimization*, v. 57, n. 3, p. 1093–1114, 2017.

233 FERREIRA, W. G. *Efficient Global Optimization Driven by Ensemble of Metamodels: New Directions Opened by Least Squares Approximation*. Dissertação (Mestrado) — Universidade Estadual de Campinas, 2016.

234 ZAEFFERER, M. et al. Efficient global optimization for combinatorial problems. *Proceedings of the 2014 Genetic and Evolutionary Computation Conference*, p. 871–878, 2014.

235 ZAEFFERER, M. *Surrogate models for discrete optimization problems*. Dissertação (Mestrado) — Technischen Universitat Dortmund, 2018.

236 HARDY, R. L. Multiquadric equations of topography and other irregular surfaces. *Journal of Geophysical Research*, v. 76, n. 8, p. 1905–1915, 1971.

237 KALITA, K. et al. Performance analysis of radial basis function metamodels for predictive modelling of laminated composites. *Materials*, v. 14, n. 12, 2021. ISSN 1996-1944. Available from Internet: https://www.mdpi.com/1996-1944/14/12/3306.

238 GARG, S.; PATRA, K.; PAL, S. K. Effect of different basis functions on a radial basis function network in prediction of drill flank wear from motor current signals. p. 777–787, 2008.

239 WU, Z. et al. Unified estimate of gaussian kernel width for surrogate models. *Neurocomputing*, Elsevier Science Publishers B. V., v. 203, p. 41–51, aug 2016. ISSN 0925-2312. Available from Internet: https://doi.org/10.1016/j.neucom.2016.03.039.

240 BENOUDJIT, N.; VERLEYSEN, M. On the kernel widths in radial-basis function networks. *Neural Processing Letters*, v. 18, n. 2, p. 139–154, 2003.

241 ACAR, E. Simultaneous optimization of shape parameters and weight factors in ensemble of radial basis functions. *Structural and Multidisciplinary Optimization*, v. 49, n. 6, p. 969–978, 2014. ISSN 16151488.

242 CHEN, Y. T.; XIANG, S.; ZHAO, W. P. Generalized multiquadrics with optimal shape parameter and exponent for deflection and stress of functionally graded plates. *Applied Mechanics and Materials*, v. 709, p. 121–124, 2014.

243 HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. [S.l.]: Macmillan Publishing, 1994.

244 NAKAYAMA, H.; ARAKAWA, M.; SASAKI, R. Simulation-based optimization using computational intelligence. *Optimization and Engineering*, v. 3, n. 2, p. 201–214, 2002.

245 BENGIO, Y.; GRANDVALET, Y. No unbiased estimator of the variance of k-fold cross-validation. *JOURNAL OF MACHINE LEARNING RESEARCH*, v. 5, p. 1089–1105, 2003.

246 ELISSEEFF, A.; PONTIL, M. Leave-one-out error and stability of learning algorithms with applications stability of randomized learning algorithms source. *International Journal of Systems Science - IJSySc*, v. 6, 01 2002.

247 KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1995. (IJCAI'95), p. 1137–1143. ISBN 1-55860-363-8. Available from Internet: http://dl.acm.org/citation.cfm?id=1643031.1643047.

248 MARTIN, J. D.; SIMPSON, T. W. Use of kriging models to approximate deterministic computer models. *AIAA Journal*, v. 43, n. 4, p. 853–863, 2005.

249 MICHELLI, C. A. Interpolation of Scattered Data: Distance Matrices and Conditionally Positive Definite Functions. *Constructive Approximations*, v. 2, p. 11– 22, 1986.

250 LOWE, D.; BROOMHEAD, D. S. Multivariable Functional Interpolation and Adaptive Networks. *Complex Systems*, v. 2, p. 321– 355, 1988.

251 KRIGE, D. G. A statistical approaches to some basic mine valuation problems on the witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, v. 52, p. 119–139, 1951.

252 SACKS, J. et al. Design and analysis of computer experiments. *Statistical Science*, v. 4, n. 4, p. 409–423, 1989.

253 CRESSIE, N. The origins of kriging. *Mathematical Geology*, v. 22, n. 3, p. 239–252, 1990.

254 SALEM, M. B. et al. Sequential dimension reduction for learning features of expensive black-box functions. In: . [s.n.], 2019. Available from Internet: https://hal.archives-ouvertes.fr/hal-01688329v2.

255 BACHOC, F. *Cross Validation and Maximum Likelihood estimations of hyper-parameters of Gaussian processes with model misspecification*. 2013.

256 LIU, H. et al. Cope with diverse data structures in multi-fidelity modeling: A Gaussian process method. *Engineering Applications of Artificial Intelligence*, Elsevier Ltd, v. 67, n. August 2017, p. 211–225, 2018. ISSN 09521976. Available from Internet: https://doi.org/10.1016/j.engappai.2017.10.008.

257 WANG, X. et al. Surrogate based multidisciplinary design optimization of lithium-ion battery thermal management system in electric vehicles. *Structural and Multidisciplinary Optimization*, v. 56, n. 6, p. 1555–1570, 2017.

258 ROUSTANT, O.; GINSBOURGER, D.; DEVILLE, Y. Dicekriging, diceoptim: Two r packages for the analysis of computer experiments by kriging-based metamodeling and optimization. *Journal of Statistical Software, Articles*, v. 51, n. 1, p. 1–55, 2012. ISSN 1548-7660. Available from Internet: https://www.jstatsoft.org/v051/i01.

259 PERDIKARIS, P. et al. Multi-fidelity modelling via recursive co-kriging and Gaussian-Markov random fields. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, v. 471, n. 2179, 2015. ISSN 14712946.

260 EMMERICH, M.; GIANNAKOGLOU, K.; NAUJOKS, B. Single- and multiobjective evolutionary optimization assisted by gaussian random field metamodels. *IEEE Transactions on Evolutionary Computation*, v. 10, n. 4, p. 421–439, 2006.

261 TOAL, D. J.; BRESSLOFF, N. W.; KEANE, A. J. Kriging hyperparameter tuning strategies. *AIAA Journal*, v. 46, n. 5, p. 1240–1252, 2008. ISSN 00011452.

262 LIU, H.; ONG, Y. S.; CAI, J. A survey of adaptive sampling for global metamodeling in support of simulation-based complex engineering design. *Structural and Multidisciplinary Optimization*, v. 57, n. 1, p. 393–416, 2018.

263 JONES, D. R. A taxonomy of global optimization methods based on response surfaces. *Journal of Global Optimization*, v. 21, n. 4, p. 345–383, 2001.

264 FORRESTER, A. I.; KEANE, A. J. Recent advances in surrogate-based optimization. *Progress in Aerospace Sciences*, v. 45, n. 1-3, p. 50–79, 2009. ISSN 03760421.

265 DANG, V.-T.; LABERGèRE, C.; LAFON, P. Adaptive metamodel-assisted shape optimization for springback in metal forming processes. *International Journal of Material Forming*, v. 12, n. 4, p. 535–552, Jun 2018.

266 SRINIVAS, N. et al. Gaussian process optimization in the bandit setting: No regret and experimental design. *ICML 2010 - Proceedings, 27th International Conference on Machine Learning*, p. 1015–1022, 2010.

267 MLAKAR, M. et al. GP-DEMO: Differential Evolution for Multiobjective Optimization based on Gaussian Process models. *European Journal of Operational Research*, v. 243, n. 2, p. 347–361, 2015. ISSN 03772217.

268 BOUHLEL, M. et al. Efficient global optimization for high-dimensional constrained problems by using the Kriging models combined with the partial least squares method. *Engineering Optimization*, v. 50, n. 12, p. 2038–2053, 2018. ISSN 10290273.

269 KITAYAMA, S. et al. Multi-objective optimization of blank shape for deep drawing with variable blank holder force via sequential approximate optimization. *Structural and Multidisciplinary Optimization*, v. 52, n. 5, p. 1001–1012, Jul 2015.

270 CHUNG, I.-B.; PARK, D.; CHOI, D.-H. Surrogate-based global optimization using an adaptive switching infill sampling criterion for expensive black-box functions. *Structural and Multidisciplinary Optimization*, v. 57, n. 4, p. 1443–1459, 2018.

271 PAN, G. et al. A sequential optimization sampling method for metamodels with radial basis functions. *The Scientific World Journal*, v. 2014, p. 1–17, 2014.

272 BALREIRA, D. S. *Otimização sequencial aproximada de estruturas laminadas de material compósito*. Dissertação (Mestrado) — Universidade Federal do Ceará, 2018.

273 WU, Y. et al. A rbf-based constrained global optimization algorithm for problems with computationally expensive objective and constraints. *Structural and Multidisciplinary Optimization*, Springer, v. 58, n. 4, p. 1633–1655, 2018. ISSN 16151488.

274 XING, J.; LUO, Y.; GAO, Z. A global optimization strategy based on the Kriging surrogate model and parallel computing. *Structural and Multidisciplinary Optimization*, Springer, v. 62, n. 1, p. 405–417, 2020.

275 XIA, W.; SHOEMAKER, C. GOPS: efficient RBF surrogate global optimization algorithm with high dimensions and many parallel processors including application to multimodal water quality PDE model calibration. *Optimization and Engineering*, Springer US, v. 22, 2020. ISSN 15732924. Available from Internet: https://doi.org/10.1007/s11081-020-09556-1.

276 GARCÍA-GARCÍA, J. C.; GARCÍA-RÓDENAS, R.; CODINA, E. *A surrogate-based cooperative optimization framework for computationally expensive black-box problems*. Springer US, 2020. v. 21. 1053–1093 p. ISSN 15732924. ISBN 0123456789. Available from Internet: https://doi.org/10.1007/s11081-020-09526-7.

277 HAFTKA, R. T.; VILLANUEVA, D.; CHAUDHURI, A. Parallel surrogate-assisted global optimization with expensive functions – a survey. *Structural and Multidisciplinary Optimization*, Structural and Multidisciplinary Optimization, v. 54, n. 1, p. 3–13, 2016. ISSN 16151488. Available from Internet: http://dx.doi.org/10.1007/s00158-016-1432-3.

278 REGIS, R. G.; SHOEMAKER, C. A. Improved strategies for radial basis function methods for global optimization. *Journal of Global Optimization*, v. 37, n. 1, p. 113–135, 2007. ISSN 09255001.

279 HAVINGA, J.; BOOGAARD, A. H. van den; KLASEBOER, G. Sequential improvement for robust optimization using an uncertainty measure for radial basis functions. *Structural and Multidisciplinary Optimization*, Structural and Multidisciplinary Optimization, v. 55, n. 4, p. 1345–1363, 2017. ISSN 16151488. Available from Internet: http://dx.doi.org/10.1007/s00158-016-1572-5.

280 YAO, W. et al. A surrogate-based optimization method with rbf neural network enhanced by linear interpolation and hybrid infill strategy. *Optimization Methods and Software*, v. 29, n. 2, p. 406–429, 2014.

281 RAHAT, A. A. M.; EVERSON, R. M.; FIELDSEND, J. E. Alternative infill strategies for expensive multi-objective optimisation. In: *Proceedings of the Genetic and Evolutionary Computation Conference*. New York, NY, USA: Association for Computing Machinery, 2017. p. 873–880. ISBN 9781450349208. Available from Internet: https://doi.org/10.1145/3071178.3071276.

282 MAIA, M. A.; Parente, E.; MELO, A. M. C. D. Kriging-based optimization of functionally graded structures. *Structural and Multidisciplinary Optimization*, 2021.

283 CHUNNA, L.; HAI, F.; CHUNLIN, G. Development of an efficient global optimization method based on adaptive infilling for structure optimization. *Structural and Multidisciplinary Optimization*, Structural and Multidisciplinary Optimization, 2020.

284 YI, J.; CHENG, Y.; LIU, J. An adaptive constraint-handling approach for optimization problems with expensive objective and constraints. *2020 IEEE Congress on Evolutionary Computation, CEC 2020 - Conference Proceedings*, 2020.

285 KUSHNER, H. J. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *Journal of Basic Engineering*, v. 86, n. 1, p. 97–106, 1964.

286 SCHOEMAKER, P. J. H. *Experiments on Decisions under Risk: The Expected Utility Hypothesis*. [S.l.: s.n.], 1980.

287 MOČKUS, J. On bayesian methods for seeking the extremum. In: MARCHUK, G. I. (Ed.). *Optimization Techniques IFIP Technical Conference Novosibirsk, July 1–7*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1974. p. 400–404. ISBN 978-3-540-37497-8.

288 HABIB, A.; SINGH, H. K.; RAY, T. A study on the effectiveness of constraint handling schemes within Efficient Global Optimization framework. *2016 IEEE Symposium Series on Computational Intelligence, SSCI 2016*, 2017.

289 WANG, H.; HU, W.; LI, E. Handling of constraints in efficient global optimization. *International Journal of Computational Methods*, v. 18, n. 2, 2021. ISSN 17936969.

290 MÜLLER, J. SOCEMO: Surrogate optimization of computationally expensive multiobjective problems. *INFORMS Journal on Computing*, v. 29, n. 4, p. 581–596, 2017. ISSN 15265528.

291 REGIS, R. G. A survey of surrogate approaches for expensive constrained black-box optimization. *Advances in Intelligent Systems and Computing Optimization of Complex Systems: Theory, Models, Algorithms and Applications*, p. 37–47, 2019.

292 PARR, J. M. et al. Review of efficient surrogate infill sampling criteria with constraint handling. *2nd International Conference on Engineering Optimization*, 2010.

293 DONG, H. et al. Kriging-assisted teaching-learning-based optimization (KTLBO) to solve computationally expensive constrained problems. *Information Sciences*, Elsevier Inc., v. 556, p. 404–435, 2021. ISSN 00200255. Available from Internet: https://doi.org/10.1016/j.ins.2020.09.073.

294 FERREIRA, W. G.; SERPA, A. L. Ensemble of metamodels: extensions of the least squares approach to efficient global optimization. *Structural and Multidisciplinary Optimization*, Structural and Multidisciplinary Optimization, v. 57, n. 1, p. 131–159, 2018. ISSN 16151488.

295 BASUDHAR, A. et al. Constrained efficient global optimization with support vector machines. *Structural and Multidisciplinary Optimization*, v. 46, n. 2, p. 201–221, 2012. ISSN 1615147X.

296 SCHONLAU, M. *Computer Experiments and Global Optimization*. Tese (Doutorado) — University of Waterloo, 1997.

297 SOHST, M.; AFONSO, F.; SULEMAN, A. Surrogate - based optimization based on the probability of feasibility. *Structural and Multidisciplinary Optimization*, Springer Berlin Heidelberg, v. 4, 2022. ISSN 1615-1488. Available from Internet: https://doi.org/10.1007/s00158-021-03134-4.

298 GRAMACY, R. B. et al. Modeling an augmented lagrangian for blackbox constrained optimization. *Technometrics*, Taylor and Francis, v. 58, n. 1, p. 1–11, 2016. Available from Internet: https://doi.org/10.1080/00401706.2015.1014065.

299 YUAN, B. et al. Efficient global optimization strategy considering expensive constraints. In: SCHUMACHER, A. et al. (Ed.). *Advances in Structural and Multidisciplinary Optimization*. Cham: Springer International Publishing, 2018. p. 133–142. ISBN 978-3-319-67988-4.

300 PARR, J. M. et al. Infill sampling criteria for surrogate-based optimization with constraint handling. *Engineering Optimization*, Taylor and Francis, v. 44, n. 10, p. 1147–1166, 2012. Available from Internet: https://doi.org/10.1080/0305215X.2011.637556.

301 REGIS, R. G. Large-Scale Discrete Constrained Black-Box Optimization Using Radial Basis Functions. *2020 IEEE Symposium Series on Computational Intelligence, SSCI 2020*, p. 2924–2931, 2020.

302 GóMEZ-BOMBARELLI, R. et al. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Central Science*, v. 4, n. 2, p. 268–276, 2018. PMID: 29532027. Available from Internet: https://doi.org/10.1021/acscentsci.7b00572.

303 VOUTCHKOV, I. et al. Weld sequence optimization: The use of surrogate models for solving sequential combinatorial problems. *Computer Methods in Applied Mechanics and Engineering*, v. 194, n. 30-33 SPEC. ISS., p. 3535–3551, 2005. ISSN 00457825.

304 HILDEBRANDT, T.; BRANKE, J. On using surrogates with genetic programming. *Evolutionary Computation*, v. 23, n. 3, p. 343–367, 2015.

305 MÜLLER, J.; SHOEMAKER, C. A.; PICHÉ, R. SO-MI: A surrogate model algorithm for computationally expensive nonlinear mixed-integer black-box global optimization problems. *Computers and Operations Research*, v. 40, n. 5, p. 1383–1400, 2013. ISSN 03050548.

306 MULLER, J. *MATSuMoTo: The MATLAB Surrogate Model Toolbox For Computationally Expensive Black-Box Global Optimization Problems*. 2014.

307 MÜLLER, J. MISO: mixed-integer surrogate optimization framework. *Optimization and Engineering*, v. 17, n. 1, p. 177–203, 2016. ISSN 15732924.

308 ZHOU, Q. et al. A sequential multi-fidelity metamodeling approach for data regression. *Knowledge-Based Systems*, Elsevier B.V., v. 134, p. 199–212, 2017. ISSN 09507051. Available from Internet: http://dx.doi.org/10.1016/j.knosys.2017.07.033.

309 DURANTIN, C. et al. Multifidelity surrogate modeling based on radial basis functions. *Structural and Multidisciplinary Optimization*, v. 56, n. 5, p. 1061–1075, 2017. ISSN 16151488.

310 XIAO, M. et al. Extended Co-Kriging interpolation method based on multi-fidelity data. *Applied Mathematics and Computation*, Elsevier Inc., v. 323, p. 120–131, 2018. ISSN 00963003. Available from Internet: https://doi.org/10.1016/j.amc.2017.10.055.

311 CHENG, K.; LU, Z.; ZHEN, Y. Multi-level multi-fidelity sparse polynomial chaos expansion based on Gaussian process regression. *Computer Methods in Applied Mechanics and Engineering*, Elsevier B.V., v. 349, p. 360–377, 2019. ISSN 00457825. Available from Internet: https://doi.org/10.1016/j.cma.2019.02.021.

312 BUNNELL, S.; GORRELL, S.; SALMON, J. Multi-fidelity surrogates from shared principal components: Application to structural design exploration and optimization. *Structural and Multidisciplinary Optimization*, Structural and Multidisciplinary Optimization, 2021. ISSN 16151488.

313 CAI, X. et al. Metamodeling for high dimensional design problems by multi-fidelity simulations. *Structural and Multidisciplinary Optimization*, Springer, v. 56, n. 1, p. 151–166, 2017. ISSN 16151488.

314 BERTRAM, A.; ZIMMERMANN, R. Theoretical investigations of the new Cokriging method for variable-fidelity surrogate modeling: Well-posedness and maximum likelihood training. *Advances in Computational Mathematics*, Advances in Computational Mathematics, v. 44, n. 6, p. 1693–1716, 2018. ISSN 15729044.

315 HEBBAL, A. et al. Multi-fidelity modeling with different input domain definitions using deep Gaussian processes. *Structural and Multidisciplinary Optimization*, n. 63, 2021. ISSN 16151488.

316 NACHAR, S. et al. Multi-fidelity bayesian optimization using model-order reduction for viscoplastic structures. *Finite Elements in Analysis and Design*, Elsevier B.V., v. 176, n. October 2019, 2020. ISSN 0168874X. Available from Internet: https://doi.org/10.1016/j.finel.2020.103400.

317 ROHIT, R. J.; GANGULI, R. Co-kriging based multi-fidelity uncertainty quantification of beam vibration using coarse and fine finite element meshes. *International Journal for Computational Methods in Engineering Science and Mechanics*, Taylor and Francis, 2021. Available from Internet: https://doi.org/10.1080/15502287.2021.1921883.

318 WANG, S. et al. A multi-fidelity surrogate model based on moving least squares: fusing different fidelity data for engineering design. *Structural and Multidisciplinary Optimization*, Springer Berlin Heidelberg, 2021. ISSN 16151488. Available from Internet: https://doi.org/10.1007/s00158-021-03044-5.

319 TOAL, D. J. Some considerations regarding the use of multi-fidelity Kriging in the construction of surrogate models. *Structural and Multidisciplinary Optimization*, v. 51, n. 6, p. 1223–1245, 2015. ISSN 16151488.

320 ROBINSON, T. D. et al. Surrogate-based optimization using multifidelity models with variable parameterization and corrected space mapping. *AIAA Journal*, v. 46, n. 11, p. 2814–2822, 2008. ISSN 00011452.

321 ZHANG, Y. et al. Multifidelity surrogate based on single linear regression. *AIAA Journal*, v. 56, n. 12, p. 4944–4952, 2018. ISSN 00011452.

322 COURRIER, N.; BOUCARD, P.-A.; SOULIER, B. Variable-fidelity modeling of structural analysis of assemblies. *Journal of Global Optimization*, v. 64, n. 3, p. 577–613, 2015.

323 KHATOURI, H. et al. Constrained multi-fidelity surrogate framework using Bayesian optimization with non-intrusive reduced-order basis. *Advanced Modeling and Simulation in Engineering Sciences*, Springer International Publishing, v. 7, n. 1, 2020. ISSN 22137467. Available from Internet: https://doi.org/10.1186/s40323-020-00176-z.

324 HUANG, D. et al. Sequential kriging optimization using multiple-fidelity evaluations. *Structural and Multidisciplinary Optimization*, v. 32, n. 5, p. 369–382, 2006. ISSN 1615147X.

325 KENNEDY, M. C.; O'HAGAN, A. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, v. 63, n. 3, p. 425–464, 2001. Available from Internet: https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9868.00294.

326 REISENTHEL, P. H.; ALLEN, T. T. Application of multifidelity expected improvement algorithms to aeroelastic design optimization. In: ____. *10th AIAA Multidisciplinary Design Optimization Conference*. [s.n.], 2014. Available from Internet: https://arc.aiaa.org/doi/abs/10.2514/6.2014-1490.

327 BARROSO, E. S. et al. Bios: An object-oriented framework for surrogate-based optimization using bio-inspired algorithms. *Structural and Multidisciplinary Optimization*, 2022. Submitted for publication.

328 HO-HUU, V. et al. Optimization of laminated composite plates for maximizing buckling load using improved differential evolution and smoothed finite element method. *Composite Structures*, v. 146, p. 132–147, 2016. ISSN 0263-8223. Available from Internet: https://www.sciencedirect.com/science/article/pii/S026382231630157X.