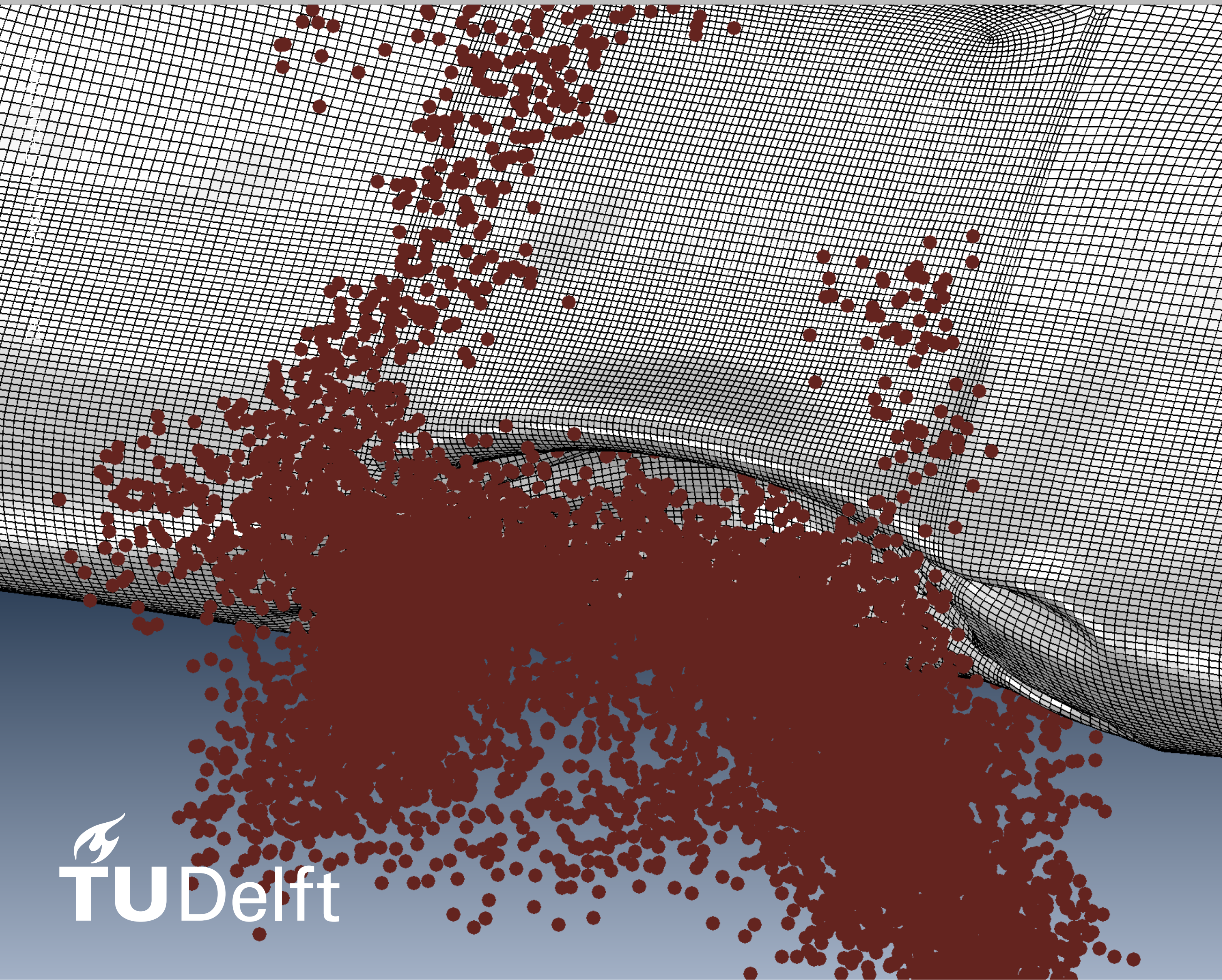
Incorporating Bird Strike Crashworthiness Requirements within an MDO Framework

Raluca-Ioana Ciobotia



Delft University of Technology

Incorporating Bird Strike Crashworthiness Requirements within an MDO Framework

by

Raluca-Ioana Ciobotia

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Friday January 12, 2024 at 9:00 AM.

Student number:	5626528	
Project duration:	March 1, 2023 – January 12, 2024	
Supervisors:	Dr.ir. S. Castro	TU Delft
	Dr.ir. D. Peeters	TU Delft
	Dr.ir. A.H. van der Laan	GKN Aerospace
Thesis committee:	Dr.ir. D. Zarouchas	TU Delft, chair
	Dr. B. Chen	TU Delft, examiner
	Dr.ir. S. Castro	TU Delft, supervisor
	Dr.ir. D. Peeters	TU Delft, supervisor

This thesis is confidential and cannot be made public until January 12, 2026.

Preface

This study marks the completion of my master's degree, a journey which was both fun and difficult. In the last two years I have evolved greatly not only from an educational point of view, but also on a personal level, and I will surely miss desperately working on an assignment while looking forward to meeting my colleagues in the Atmosfeer.

To the reader, I have started this project nearly a year ago without any clue regarding surrogate models or Bayesian optimization. Everything seemed confusing, especially since the terminology differed from field to field. Therefore, I wrote this report such that my former self could have read it without a problem. Thank you for taking the time to go through my work.

There are many people to whom I am indebted. I would like to first thank my supervisors, Dr.ir. Saullo Castro and Dr.ir. Daniël Peeters, for their guidance and patience. They have continuously provided me with valuable input whenever I had discovered something new and interesting, which also includes the times when I needed to tone down my excitement and start focusing on the end goal of my project, or when I was in deadlock for weeks due to modelling problems.

I would also like to express my gratitude towards my external supervisor, Dr.ir. Ton van der Laan, who checked on me from time to time as a friend, on top of all his invaluable advice. I have greatly appreciated the help of Dr.ir. Tobie van den Berg as well, although I have only got the chance to have him as a supervisor for the first few months of the project. I would also like to extend my gratitude to Ir. Wydo van de Waerdt and Ir. Luc Hootsmans for helping me with bird strike finite element modelling. I am especially thankful for the opportunity of working on my thesis from my home country, where, luckily, there is also another GKN Fokker office. Therefore, I want to also say thank you to my Romanian colleagues, who have welcomed me with open arms.

Last but not least, I am most indebted to my friends and family, who have continuously supported me and showed great interest in my work. Thank you for rooting for me!

Raluca-Ioana Ciobotia
Delft, December 2023

Abstract

The present study, which was carried out in collaboration with GKN Fokker, focuses on incorporating bird strike crashworthiness requirements within a multidisciplinary optimization (MDO) framework. During the preceding three-month internship in the same company, a pivotal contribution to this project was the development of an Abaqus interface for the Multidisciplinary Modeller, MDM, created within the Center of Competence in Design department. MDM is a Python/ParaPy-based automated generator of wings, moveables and flaps, starting from a set of user-specified parameters. The generation of ready-to-run input files thus lays the foundation for the subsequent optimization process, as any changes in materials or geometry can be easily accommodated.

The core objective of the research is to minimize the weight of an aircraft wing while taking into account additional requirements related to the extent of damage caused by bird strikes. Unfortunately, such events occur more frequently than one would be comfortable with, and stringent requirements are set in place to guarantee the safety of the passengers. Among these requirements, the aircraft must be capable of landing safely after such an event, being subject to loads associated with get-home conditions.

As a consequence, two critical constraints are formulated within the optimization framework, addressing the residual strength of the damaged front spar following a bird strike, coupled with a requirement based on a maximum penetration depth. The last constraint has also been included due to the rising popularity of the electric vertical take-off and landing aircraft, which not only fly at low altitudes, thus increasing the risk of bird strike, but may also contain battery packs in the leading edge, for instance, which can pose a significant risk if damaged. To tackle the complexity of this highly-dimensional optimization challenge, a methodology based on Bayesian optimization is proposed, employing surrogate models coupled with a preliminary variable ranking procedure.

The Kriging metamodel is identified as a suitable candidate, thanks to its error prediction capabilities, which are paramount in Bayesian optimization. A variance-based dimensionality reduction method is proposed, which makes use of an initial surrogate to estimate the main and interaction effects of the variables. The quantification of the significance of a variable is expressed as its percentage contribution to the total variance, thus allowing for an intuitive selection of the most important parameters. After the screening procedure is complete, the optimization procedure is carried out in the reduced design space, which uses the constrained expected improvement as an acquisition function. The proposed methodology is then applied on a case study problem, involving a five-bay metallic wing segment subject to the constraints aforementioned, involving 19 design variables representing the thicknesses of various components.

Remarkable weight savings have been achieved, the final result being 40% lighter than the lightest feasible design among the initial data points. A significant dimensional reduction has also been attained for the maximum depth constraint, which is expected due to the local nature of the impact. Not only did the number of variables greatly decrease from 19 to just 3, but a considerable increase in the accuracy of the corresponding metamodel has also been registered, thanks to an increase in sampling density in the reduced space. However, the variable screening procedure revealed intricate interaction effects with respect to the residual strength of the front spar, emphasizing the nuanced complexity inherent in crashworthiness considerations. Nevertheless, a moderate dimensional reduction has been achieved for this constraint as well, reducing the number of variables to 8, thus proving the efficacy of the proposed variable screening procedure.

In conclusion, the utilization of Kriging models, variable ranking procedures, and Bayesian optimization collectively contributed to the success of achieving remarkable weight savings, proving the efficiency of the proposed methodology. Moreover, it has been shown that the integration of a residual strength requirement is necessary, as many cases were uncovered where no significant penetration occurred, although the application of the considered load case, which is not from critical to an undamaged wing, resulted in high stresses to the front spar of the damaged structure.

Contents

List of Figures

List of Tables

Nomenclature

Abbreviations

Abbreviation	Definition
ALE	Arbitrary Lagrange Euler
ANOVA	Analysis of Variance
AMC	Acceptable Means of Compliance
CAR	Civil Air Regulations
CoC	Center of Competence
DACE	Design and Analysis of Computer Experiments
DoE	Design of Experiments
EASA	European Aviation Safety Agency
EGO	Efficient Global Optimization
EOS	Equation of State
ESE	Enhanced Stochastic Evolutionary
eVTOL	Electric Vertical Take-Off and Landing
FAA	Federal Aviation Administration
FE	Finite Element
FEA	Finite Element Analysis
FEM	Finite Element Method
GP	Gaussian Process
GSA	Global Sensitivity Analysis
IBIS	ICAO Bird Strike Information System
ICAO	International Civil Aviation Organization
KBE	Knowledge-Based Engineering
KPLS	Kriging Partial Least Squares
LHD	Latin Hypercube Design
LHS	Latin Hypercube Sampling
MARS	Multivariate Adaptive Regression Splines
MDM	Multidisciplinary Modeler
MDO	Multidisciplinary Optimization
MF-SRS	Multi-Fidelity Successive Response Surface
MLE	Maximum Likelihood Estimate
MSE	Mean Squared Error
PLS	Partial Least Squares
PRESS	Predicted Residual Error Sum of Squares
PRS	Polynomial Response Surface
RBF	Radial Basis Function
RMSE	Root Mean Square Error
SLHD	Sliced Latin Hypercube Design
SMT	Surrogate Modeling Toolbox
SPH	Smoothed Particle Hydrodynamics
SQP	Sequential Quadratic Programming
SRSM	Successive Response Surface Modelling
SVR	Support Vector Regression
VGP	Variational Gaussian Process

Symbols

Symbol	Definition	Unit
$[\cdot]$	Matrix or unit, by context	-
$\{\cdot\}$	vector	-
$c_1(\cdot)$	difference between the maximum Mises stress on the front spar and 295 MPa	MPa
$c_2(\cdot)$	difference between the penetration depth and 240 mm	mm
c_v	specific heat capacity at constant volume	$J/(kg \cdot K)$
$d(\cdot)$	distance function	-
$E[\cdot]$	expected value	-
E_m	specific internal energy	J/kg
G_f	Fracture energy	N/mm
$[\mathbf{C}]$	Damping matrix	$[C_{ij}] = kg/s, \forall C_{ij} \in [\mathbf{C}]$
f	function to be modelled/approximated	-
$\hat{f}$	surrogate model predictor	-
$\{\mathbf{F}\}$	Force vector	$[F_i] = N, \forall F_i \in \{\mathbf{F}\}$
k	number of dimensions	-
$[\mathbf{K}]$	Stiffness matrix	$[K_{ij}] = kg/s^2, \forall K_{ij} \in [\mathbf{K}]$
m	mass	kg
$[\mathbf{M}]$	Mass matrix	$[M_{ij}] = tonne, \forall M_{ij} \in [\mathbf{M}]$
$\mathcal{M}_{m,n}$	the set of $m \times n$-dimensional real matrices	-
n	number of sampling points	-
p	pressure or material constant in Cowper-Symonds law	N/mm^2 or $-$
$p_i, i \in \overline{1,k}$	the power of the distance in the Kriging exponential correlation function	-
$\boldsymbol{R}$	correlation matrix	-
s	standard deviation	-
T	temperature	K
$\{\mathbf{u}\}$	Displacement vector	$[u_i] = mm, \forall u_i \in \{\mathbf{u}\}$
$\{\dot{\mathbf{u}}\}$	Node velocity vector	$[\dot{u}_i] = mm/s, \forall \dot{u}_i \in \{\dot{\mathbf{u}}\}$
$\{\ddot{\mathbf{u}}\}$	Node acceleration vector	$[u_i] = mm/s^2, \forall \ddot{u}_i \in \{\ddot{\mathbf{u}}\}$
$\overline{u}_f^{pl}$	plastic displacement at failure	mm
V	volume or velocity	mm^3 or mm/s
$Var[\cdot]$	variance	-
V_c	cruise speed	m/s
$\boldsymbol{x}$	a k-dimensional sampling point	-
$\boldsymbol{X}$	vector consisting of sampling points	-
$\boldsymbol{x}^{(i)}$	the i$^{\text{th}}$ sampling point, $i \in \overline{1,n}$	-
x_j	the j$^{\text{th}}$ coordinate of a sampling point	-
Y	observational stochastic variable	-
$\boldsymbol{Y}$	observational stochastic vector	-
$\boldsymbol{y}$	response vector	-
ε	strain	-
ε_V	volumetric strain	-
$\dot{\varepsilon}$	strain rate	s^{-1}
$\overline{\varepsilon}_D^{pl}$	equivalent plastic strain at damage initiation	-
η	stress triaxiality	-
$\theta_l, l \in \overline{1,k}$	hyperparameter of the Kriging predictor	-
μ	mean of the Kriging stochastic term	-
$\hat{\mu}$	maximum likelihood predictor of μ	-
ρ	density	kg/m^3 or g/cm^3

Symbol	Definition	Unit
σ	stress or standard deviation of the Kriging stochastic term	MPa or -
σ_{y0}	ultimate stress	MPa
ω_D	plastic deformation variable; $\omega_D = 1$ if damage occurred	-

1

Introduction

On the 15th of January, 2009, US Airways Flight 1549 experienced a complete loss of engine power after hitting a flock of geese shortly after take-off. Fortunately, the experienced pilots managed to perform an emergency landing on the Hudson River, as seen in figure 1.1, hence the nickname of this incident, “the miracle on the Hudson”. This case was probably popularized by the movie “Sully” released in 2016, where Tom Hanks stars as Captain Chesley “Sully” Sullenberger.

Figure 1.1: US Airways Flight 1549, shortly after an emergency landing on the Hudson River. Image source: Greg Lam Pak Ng, 2009.

It is probably not an exaggeration, though, that there is a common conception that such events rarely happen. Unfortunately, bird strikes occur more frequently than one would expect, as will be presented in the first section of this chapter. Thus, the adage “prevention is better than cure” is proven true once again. Airlines and aviation authorities take steps to reduce the likelihood of severe damage caused by bird impacts, such as implementing bird control measures in the vicinity of airports, or enforcing regulations regarding the ability of an aircraft of landing safely after the occurrence of such an event. The goal of the present thesis, which has been carried out in collaboration with GKN Fokker, is to further these efforts by investigating how bird strike crashworthiness requirements can be taken into consideration early in the design process, in the context of multidisciplinary optimization. As will become apparent in section 1.3, such requirements are usually taken into account much later, and the resulting design may not be an optimal one in terms of weight.

The following aspects will be covered in this opening chapter: firstly, an overview of recent statistics with respect to bird strikes will be presented in the first section. Afterwards, crashworthiness requirements related to such events will be reviewed, followed by a brief discussion on why multidisciplinary design optimization is necessary in sections 1.2 and 1.3, respectively. The fourth and fifth sections are going to identify two key aspects which will constitute the main focus of the thesis, namely surrogate models and variable screening procedures. Section 1.6 will go over the tools available for this scope, as some Python packages which were used extensively throughout the project have been developed within GKN Fokker. Finally, the last section will present the outline of the report.

1.1. Bird strike statistics

As previously hinted, bird strikes are a rather frequent occurrence, and data indicates that they present a notable safety hazard. The International Civil Aviation Organization, ICAO, has put in place the ICAO Bird Strike Information System, IBIS, which has been collecting and analysing wildlife strike reports since 1980. The number of reporting states has steadily increased since then: for instance, a total of 91 states participated in the 2008-2015 survey compared to 136 in the most recent analysis (International Civil Aviation Organization, 2023).

There were 273,343 reported strikes between 2016 and 2021 alone, a figure which has dramatically risen from 2008-2015, when just over a third were declared, namely 97,751. It should also be noted that the last bulletin covered 6 years, a period which was also affected by the COVID-19 pandemic and subsequent travel bans, while the analysis done before it had spread over 8 years. As a consequence, the average number of reports per year increased from 12,219 to 45,557 between the last two surveys. However, it cannot be concluded that the frequency of such occurrences has increased, as the data is affected by the lack of participation of several states in the study. A summary of the total number of reports and reporting regions for the last three data collection periods can be found in table 1.1.

Reporting period	Total number of reports	Total number of reporting regions
2001-2007	42,508	51
2008-2015	97,751	91
2016-2021	273,343	136

Table 1.1: Total number of reports and reporting regions for the last three data collection periods.

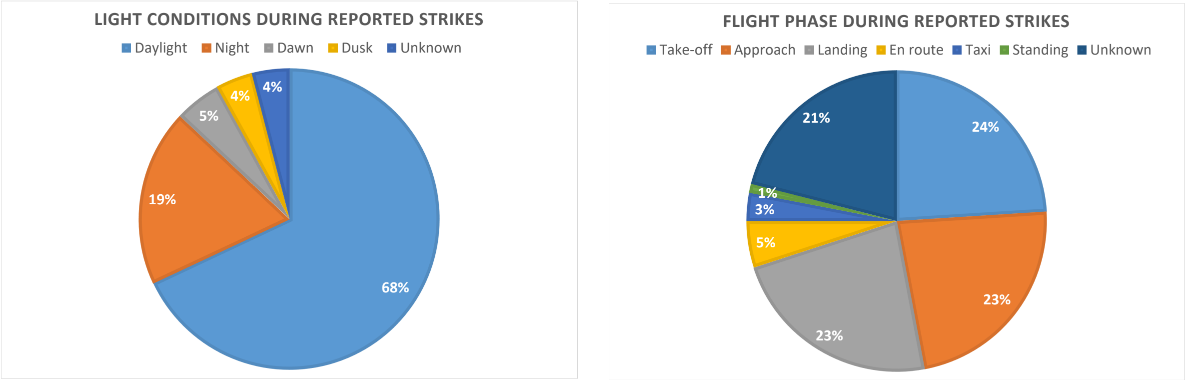
As impacts with various mammals take up only 2.8% of the cases, the vast majority of events are caused by birds. It is reiterated that bird strikes are not sufficiently declared in some regions, as 0.7% of the reported strikes occurred in African or Middle-Eastern regions. However, the efforts made in reporting and collecting the data are obvious: between 2008 and 2015, the North American, Central American, Caribbean, European and North Atlantic regions made up nearly 98% of the total reports, compared to 59.4% in the present. Table 1.2 presents the share of reports by region of occurrence over the last three data collection periods. While South America and Asia/Pacific have recently joined the study, wildlife strikes in African and Middle Eastern regions seem to be under-reported.

Reporting period Region	2001-2007	2008-2015	2016-2021
North American, Central American and Caribbean	69.9%	47.3%	33.3%
European and North Atlantic	26.3%	50.2%	26.1%
South America	1.1%	0.1%	25.9%
Asia/Pacific	0.2%	1.2%	14.0%
Africa	1.9%	1.1%	0.6%
Middle East	0.6%	0.1%	0.1%

Table 1.2: Regions of occurrence versus percentage of total impacts declared during the last three data collection periods.

In the same ICAO bulletin, both the light conditions and the flight phase during impact have been analyzed, as presented in figure 1.2. It is revealed that 68% of bird strikes occur during the day, and only 19% take place at night, the rest being either at dawn (5%), dusk (4%) or not stated (4%), figures which do not differ much from the last analyses conducted. Moreover, July and August are the months with the highest wildlife strike activity, while December, January and February seem to be the safest. There are some changes in this aspect since the last analysis, when most reports originated from the Northern Hemisphere.

As for the flight phase when the strikes occurred, 23% of the reports took place during approach, 24% during take-off, and 23% during landing, with only 5% en route. However, compared to 2001-2007, there is an increase in unknown cases. Nevertheless, it can be concluded based on the 2001-2007, 2008-2015 and 2016-2021 data that the majority of strikes occur in the vicinity of airports. Therefore, military flight missions at low altitudes have a high risk of collision as well, which is especially concerning in the case of time-sensitive rescue operations, for example. Moreover, the increasingly popular electric vertical take-off and landing (eVTOL) aircraft are also at high risk. For instance, the Joby S4 aircraft, which has recently been delivered to the Edwards Air Force Base (Joby Aviation, 2023) in September, 2023 for demonstration purposes, is targeting the electric aerial ridesharing market. As a consequence, the flight altitudes are expected to be low compared to conventional aircraft, thus increasing the risk of bird strikes.



(a) Light conditions versus share of reports.

(b) Flight phase versus share of reports.

Figure 1.2: Reports by light conditions and flight phase at the time of impact, between 2016-2021.

All forward-facing components are at risk. The engines bear the greatest vulnerability, with 17% of all impacts occurring in this area, primarily due to their large frontal area. Following closely are the windshield, radome, and wing, each accounting for 14%, succeeded by the fuselage (12%) and nose (10%). Out of the total number of reports in the last analysis, 3% resulted in damaged components. Engines and wings unfortunately face the highest likelihood of damage, as depicted in figure 1.3, accounting for 34% and 18% of total damages, respectively. Note that, in figure 1.3, the impact and damage shares are given with respect to the total number of reports and the total number of events in which damage occurred, respectively.

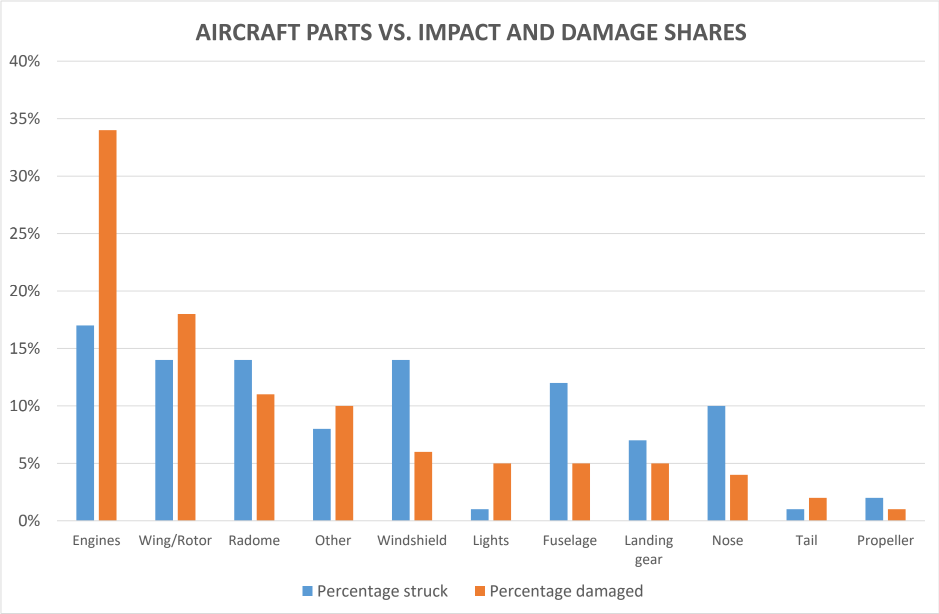


Figure 1.3: The shares of the total damage and impact reports of various components, between 2016-2021.

Nevertheless, even a small impact can disrupt flight operations, making it crucial for airlines and aviation authorities to take steps towards mitigating this risk. All in all, Goraj and Kustron, 2018, who conducted a review on research trends in both bird strike and hail impact simulations on wing leading edges, state that a bird strike occurs once every 2000 flights, which is not an insignificant figure. Airlines often struggle with the costs of maintaining their aircraft, especially when having to keep them grounded and losing revenue. The annual cost of such events is estimated by Heimbs, 2011 to range from 614 million to 1.28 billion US dollars globally.

1.2. Bird strike crashworthiness requirements

Because it is now settled that bird strikes are a rather common occurrence, it comes with no surprise that crashworthiness requirements have been formulated by aviation authorities. The review on bird impact simulations by Guida et al., 2022 begins with a summary of the US regulations related to bird strike crashworthiness. They provide a brief historical overview, highlighting that prior to 1970, the Civil Air Regulations (CAR) were the only authority which required the prevention of windshield penetration caused by a four-pound bird impact at cruise speed. Unfortunately, in 1962, a Vickers Viscount turboprop airplane had its horizontal stabilizer and elevator severely damaged following an impact with a swan. In response to this fatal accident, the regulations were changed in 1970, and the FAA then required transport airplanes to be capable of safe flight and landing after impact with birds weighing up to eight pounds. Regrettably, as is often the case, many air regulations are enacted after the occurrence of tragic incidents.

Nowadays, both the European Aviation Safety Agency, EASA, and the Federal Aviation Administration, FAA, have bird strike requirements in place. The regulations are the same regarding wing leading edge structures, which will constitute the focus of the present thesis. EASA CS 25.631 requires the following:

> The aeroplane must be designed to assure capability of continued safe flight and landing of the aeroplane after impact with a 4 lbs bird when the velocity of the aeroplane (relative to the bird along the aeroplane's flight path) is equal to V_c at sea-level or 0.85 V_c at 8000 ft, whichever is more critical. Compliance may be shown by analysis only when based on tests carried out on sufficiently representative structures of similar design.

In the EASA certification memorandum CM-S-001, there is a collection of additional criteria regarding the guarantee of "continued safe flight and landing" with respect to other requirements. Among these, the ones relating to the wing leading edge are the following:

- EASA CS 25.571(e) requires that "the aeroplane must be capable of successfully completing a flight during which likely structural damage occurs as a result of bird impact as specified in CS 25.631. The AMC to 25.571 (in paragraph 2.7.2) specifies the loads associated with get home conditions that have to be met for this case."
- EASA Acceptable Means of Compliance (AMC) 25.571, which concerns the damage tolerance and fatigue evaluation of structures, paragraph 2.7.2, states the following:

  > The maximum extent of immediately obvious damage from discrete sources should be determined and the remaining structure shown to have static strength for the maximum load (considered as ultimate load) expected during the completion of the flight. In the absence of a rational analysis the following ultimate loading conditions should be covered:
  >
  > a. *At the time of the incident*:
  >   i. The maximum normal operating differential pressure (including the expected external aerodynamic pressures during 1 g level flight) multiplied by a factor 1.1 combined with 1 g flight loads.
  >   ii. The aeroplane, assumed to be in 1g level flight should be shown to be able to survive the overswing condition due to engine thrust asymmetry and pilot corrective action taking into account any damage to the flight controls which it is presumed the aeroplane has survived.
  >
  > b. *Following the incident*: 70% limit flight manoeuvre loads and, separately, 40% of the limit gust velocity (vertical or lateral) as specified at V_c up to the maximum likely

> operational speed following failure, each combined with the maximum appropriate cabin differential pressure (including the expected external aerodynamic pressures). Further, any loss in structural stiffness which might arise should be shown to result in no dangerous reduction in freedom from flutter up to speed V_c/M_c.

- EASA CS 25.629 requires "freedom for aeroelastic instability throughout the envelope described in CS 25.629 (b)(2) for any damage or failure condition, required or selected for investigation by CS 25.571 and any damage, failure or malfunction, considered under CS 25.631"

Aeroelastic analyses are out of the scope of the present research, therefore related regulations will not be taken into account. Although the criteria aforementioned do provide some guidelines regarding which loads the damaged structure should withstand, a simplistic load case will be treated in the present report. During the structural optimization procedure, the weight of the wing will serve as the objective function to be minimized, and two constraint functions will be formulated concerning both the penetration depth of the impact, and the static strength of the damaged structure. The first constraint is formulated with eVTOLs in consideration, where battery packs could be placed in the leading edge of a wing, posing a potential risk of catastrophe in case of damage, while the second constraint addresses the previously mentioned requirements. Such simplified constraints associated with bird strike crashworthiness regulations are going to be discussed in chapter 4, section 4.2. However, it should be noted that EASA CM-S-001 specifies that no penetration is the *preferred* certification approach, although it is hypothesized that allowing for some penetration may result in more drastic weight savings.

1.3. The need for multidisciplinary design optimization

In a review of multidisciplinary optimization (MDO) architectures, Cramer et al., 1994 put forward the idea that MDO was expected to make product design more economical, particularly when applied in the early stages of the design process. This notion was reiterated almost two decades later in a separate survey by Martins and Lambe, 2013. This section will briefly address the necessity of MDO in the aviation industry and highlight the importance of integrating bird strike requirements into the optimization workflow as early as possible in the design process.

Schuhmacher et al., 2002, who used MDO techniques in the preliminary sizing of the wingboxes for the Fairchild Dornier regional jet family, emphasized that the typical design process often focuses on a limited number of dominating load cases. As a consequence, this approach carries the risk of not meeting all design criteria. When assessing the feasibility of a design with respect to the whole set of requirements later in the process, such setbacks may result in reworks and increased expenses. Moreover, due to resource and time constraints, the manual iterative process is generally terminated after attaining a design which is viable in terms of strength and is close enough to a preset goal weight. Multidisciplinary optimization aims not only to accelerate the design process, but also to improve the performance of the final solution solely thanks to the incorporation of as many requirements as possible. To illustrate the efficacy of MDO, Schuhmacher et al., 2002 took into account a variety of constraints, concerning stress, buckling, fatigue, manufacturing, lightning protection, and aeroelastic requirements.

As for bird strike crashworthiness requirements, Georgiadis et al., 2008, who proposed and validated a methodology for certifying the composite moveable trailing edge of the Boeing 787 Dreamliner against bird strike, emphasize the need for eliminating the conservatism associated with empirical formulae in order to save weight. For example, the RAE document (McNaughtan, 1972) defines a conservative method of designing a wing leading edge to be resistant to bird strike, and is usually employed for the preliminary sizing of metallic structures. Indeed, Guida et al., 2022, who conducted a review on bird impact processes, also outlined that empirical data exists, although it is only preparatory. Another conservative approach mentioned by Georgiadis et al., 2008 is the so-called "cannonball method", where the bird supposedly removes everything in its path during impact. While this approach does simplify the determination of the residual strength, simulations of a higher fidelity are needed to prevent oversized structures.

As a consequence, there is a need for accurate numerical models which would remove the conservatism aforementioned, while also keeping the number of experiments necessary for certification as low as possible, due to their exorbitant costs. However, Goraj and Kustron, 2018, who conducted a review of current research trends in bird strike and hail impact simulations, emphasized that more compliant structures alleviate impact resistance as more kinetic energy can be absorbed by transforming it to plastic deformation. Unfortunately, this contradicts the static load requirements, which usually promote rather rigid structures. Tackling both contradicting requirements simultaneously is thus necessary, which is possible in the context of multidisciplinary optimization.

1.4. The need for surrogate modelling

While accurate bird strike simulations are needed to keep the number of experiments as low as possible, the computational costs for such highly non-linear analyses are tremendous. For instance, a single bird strike analysis took 24 hours to complete in Georgiadis et al., 2008. However, it is acknowledged that the analysis involved composite structures and that the paper appeared 15 years ago, therefore it is expected that the actual computational time is smaller nowadays. Nevertheless, a single analysis of the case study detailed in section 4.1 requires approximately 10 hours to complete on a 32-processor node of the university's high performance computing cluster, which is far from a negligible duration.

One possible method for structural optimization is to employ genetic algorithms and particle swarm optimizers, as they do not make use of gradients. Indeed, as outlined by Naceur et al., 2006, who developed an optimization methodology for geometry tools used in sheet metal forming, the gradient information necessary for most minimization algorithms is not always available or accurate, particularly when black-box commercial software are employed. In the context of bird strike simulations, gradients are not readily available, and should be approximated using finite differences, for instance. However, the latter entails conducting even more analyses only to find a single gradient, which is prohibitive in terms of time. Therefore, gradients are truly not available. Unfortunately, though, gradient-free direct optimizers require a rather large number of sequential computations in order to locate the optimum, therefore neither this approach is feasible.

Surrogate models, also called metamodels, constitute a solution to this problem. Starting from a discrete set of points, the response of the underlying function is reconstructed, and the optimization task is carried out on this approximation instead. There are numerous metamodels which can be employed, but Gaussian processes will be identified as viable candidates in section 3.3. Succinctly, such approximations are based on the prior assumption that the observations are normally distributed, and posterior distributions of the response at an unknown data point can be derived. These surrogate models constitute the foundation of Bayesian optimization, as will be presented later in section 3.5, where additional data points are suggested such that the design space is explored and exploited in a balanced manner.

1.5. The need for variable screening procedures

Surrogate models can serve not only to accelerate the identification of optimal designs but also to evaluate the impact of each design variable on the studied response. For instance, while the rear spar plays a crucial role in the overall static strength of a wing, its contribution to the impact energy absorption capabilities is probably minimal. Hence, quantifying the importance of a variable with respect to a particular constraint does not only provide valuable information about the behaviour of the structure, but can also be utilized to reduce the dimensionality of the search space.

Undoubtedly, as the number of variables increases, analysis and optimization times tend to increase at a superlinear rate. This phenomenon has been coined as the “curse of dimensionality” by Koch et al., 1999, as an excessive number of design variables may impede the construction of accurate metamodels. This setback is due to the significant computational cost associated with generating a sufficiently large number of training points to ensure an acceptable sample density throughout the whole design space.

Indeed, Viana et al., 2021, who conducted a review on surrogate modelling in the context of multidisciplinary structural optimization, recommend reducing the dimensionality of the problem if there are more than 10 variables, which was given as an approximate figure based on the authors' experience. As a consequence, there is a strong need for variable screening procedures to be carried out before the optimization process starts, in order to reduce the size of the design space. A suitable variable screening method will be covered in section 3.4.

1.6. Available tools

The exploration of several structural concepts is often a time-consuming and repetitive process, complicated by the involvement of multiple disciplines such as stress, cost, weight, and manufacturability. This complexity is further exacerbated by the concurrent development process typically employed nowadays, which requires numerous iterations to arrive at a feasible design (van den Berg and van der Laan, 2021).

In order to respond to this challenge, a Knowledge-Based Engineering (KBE) multidisciplinary design system has been developed at GKN Fokker within the Center of Competence (CoC) in Design department, which is responsible for developing various tools in support of design process optimization. The Multidisciplinary Modeler, MDM, is a Python package that generates products such as flaps, wingboxes, and moveables starting from generic primitives such as skin panels, spars, ribs, stringers, and brackets. These products, which are easily defined by the user using Python dictionaries, are intended to enable design space exploration for trade studies and a rapid and consistent analysis model generation.

During the author's internship within the CoC department, two additional analysis modules have been developed to further the capabilities offered by MDM: a Mesh Module and an Abaqus Interface. The latter aims to generate ready-to-run input files containing all the necessary information for the creation of finite element (FE) models based on these products. The two analysis modules were developed using the pre-existing ParaPy Mesh and Abaqus Adaptor packages (ParaPy B.V., 2022), and are going to be extensively used for data collection, as they facilitate rapid input file generation based on user-supplied design variables.

Two additional open-source Python packages will be used throughout the present thesis: the Surrogate Modeling Toolbox, SMT 2.3.0 (Saves et al., 2023), and Trieste 1.2.0 (The Trieste Contributors, 2020a). Both of them contain continuous and mixed-integer surrogate models, benchmark functions, and optimization algorithms. SMT is going to be used for the variable ranking procedure, while Trieste is going to be used for the Bayesian optimization on the reduced design space. No commercial software besides Abaqus will be used for this project.

1.7. Simplifying assumptions

Due to the time constraints for completing a Master's thesis, several simplifying assumptions will be formulated. These simplifications, which are summarized in the following list, are needed due to the numerous aspects involved in the study, namely surrogate models, Bayesian optimization, and high-speed impact simulations:

1. Only metallic structures will be considered.
2. The metallic parts will be connected via tie mesh constraints, emulating a perfect bond between them.
3. Only ordinary Kriging surrogate models will be employed, although two different kernels will be covered.
4. The only design variables will be the thicknesses of various parts. As a consequence, the wing architecture will not change throughout the whole study, i.e. the geometrical surfaces of the wing are fixed.
5. The design variables will be treated as discrete, although the underlying surrogate model will have a continuous formulation.
6. During the variable screening procedure, only up to second-order effects will be considered.
7. It will be assumed that, once the variable screening procedure is conducted, the significant variables will not change during the optimization process, i.e. no recovery mechanism will be employed to safeguard against an erroneous identification.
8. The critical bird impact location is going to be assumed beforehand, although it is acknowledged that it may vary with each design.
9. The contact between elements in the model will be assumed to be frictionless.
10. A 10,000-particle bird SPH model will be assumed to provide sufficiently accurate results.
11. Temperature-dependent phenomena will not be included in the material model of the wing.

12. The responses of the constraint functions will be treated as normally distributed observational random variables, and the suitability of the prior distribution will not be assessed.
13. It will be assumed that the static strength of the damaged wing is given only by the wingbox, i.e. the leading edge will not be included in the model responsible for its evaluation.
14. It will be assumed that the damping ratio of the structure is not going to be severely affected by the bird impact and resulting damage.

1.8. Outline of the report

To conclude this introductory chapter, the focus of the thesis will be the formulation of constraint functions with respect to bird strike crashworthiness requirements and their approximations using surrogate models, as well as the employment of variable screening procedures. Two viable constraints will be identified, and figure 1.4 outlines the proposed steps for constructing their corresponding metamodels as well as their incorporation within a Bayesian optimization framework. Firstly, an initial sampling needs to be done in order to evaluate the importance of each variable via a screening study for each individual constraint. After the insignificant parameters are eliminated, the initial sampling is projected onto a new design space of reduced dimension. A surrogate model will then be trained in the new space, and additional data points will be suggested based on the Bayesian optimization acquisition function, until a suitable convergence criterion is met. As the weight of the wing can be derived analytically thanks to its architecture not changing throughout the whole design space, no surrogate model is needed for the objective function. Although figure 1.4 is valid for this case, both chapters 3 and 5 will tackle black-box objective functions as well, ensuring that the methodology can easily accommodate them.

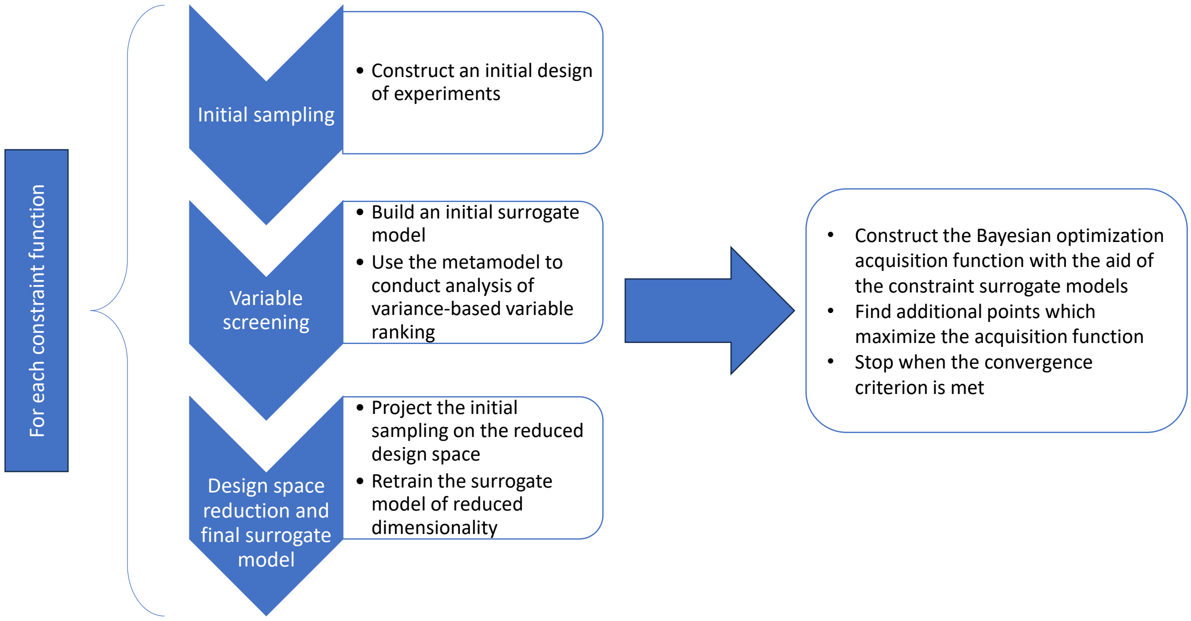


Figure 1.4: Proposed general steps for the structural optimization process.

Following these steps, the present thesis will be structured in the following manner:

- Chapter 2 will be focused on identifying structural optimization studies with respect to crashworthiness requirements, from both the automotive and aerospace engineering fields. This overview will then be followed by key literature on aspects related to bird impact modelling, sampling methods, surrogates, variable screening procedures, and Bayesian optimization. A number of key aspects which require further investigation will be pinpointed, and relevant research questions will be formulated.

- Chapter 3 will present all theoretical aspects relevant to the proposed methodology, which were identified in the literature review in chapter 2.
- Chapter 4 will present a case study problem of a simplified wing segment impacted by a bird, and two constraint functions will be formulated. Moreover, the Abaqus models used for data generation will be analyzed.
- Chapter 5 will propose a methodology for conducting structural optimization, and will cover initial data generation and model fitting, variable ranking, and the actual Bayesian optimization procedure.
- Chapter 6 will present the results obtained by applying the suggested methodology on the case study problem formulated in chapter 4.
- Chapter 7 will conclude the present thesis, discussing the answers to the proposed research questions, and recommending interesting directions for further research.

2

Literature review and proposed research questions

As presented earlier in chapter 1, the topic of structural optimization against bird strikes is going to cover a wide range of research fields, comprising high-speed impact modelling, surrogate models, and Bayesian optimization, to name a few. For this reason, these aspects will be covered in separate sections of the present review, although their theoretical basis will also be addressed as stand-alone topics later on, in chapter 3. As a consequence, the present chapter will first cover literature on crashworthiness structural optimization, in both the aerospace and automotive industries, followed by a review on bird impact modelling, design of experiments, surrogate modelling, variable ranking, Bayesian optimization, and variable ranking. Finally, some interesting research gaps are going to be identified, and research questions will be formulated accordingly.

2.1. Crashworthiness structural optimization

Bisagni et al., 2002 conducted surrogate model-based structural optimization on the subfloor structure of a helicopter. It was emphasized that crashworthiness requirements should be taken into consideration during the first phases of the project, as outlined earlier in the introductory chapter. The significant variables were identified via a sensitivity study, and Artificial Neural Networks were the chosen metamodel. The thicknesses of various parts were treated as continuous variables, and were rounded up to the nearest commercially available value. The number of rivets was also included. Unfortunately, there was no mention of the sampling method, only that it “guarantees a random and homogeneous allocation of the runs inside the optimization domain” (Bisagni et al., 2002).

The design of the subfloor structure was parametrized in order to enable automatic generation of PAM-CRASH files using MATLAB. Two different optimizers were compared, both of them being implemented in MATLAB, namely Sequential Quadratic Programming (SQP) and genetic algorithms. The multi-objective problem was converted to a single-objective one by assigning weights to the mean and maximum forces encountered and crash efficiency. The solutions obtained with the two optimizers were similar, and after conducting a full finite element analysis, it was discovered that the response differed from the prediction of the surrogate model by less than 5%, allowing for an increase of 25% in crash efficiency.

The generalization capability of the surrogate model was demonstrated by the fact that the optimum design had 7 rivets, while data points with only 4, 6 and 8 rivets were used for training. Moreover, it was emphasized that the neural network training processes and all optimization runs were comparable to a single finite element analysis in terms of computational time. It was estimated that a direct optimization (i.e. without employing a surrogate model) would have taken approximately 2.5 times more, thus proving the need for employing metamodels.

Lanzi et al., 2004 continued this study by decomposing the whole structure into smaller linked ones, the interactions being validated by a simplified model of the helicopter subfloor. The behaviour of each substructure was then approximated using Artificial Neural Networks as in the earlier study, and the optimization was carried out using only genetic algorithms. The thicknesses were considered yet again as continuous variables. While a smaller increase in crash efficiency was achieved compared to the earlier study (12 versus 25%), a significant CPU time saving was obtained.

Jumping to the automotive industry, Craig et al., 2005 were the first to use response surfaces to perform variable screening in the crashworthiness field, according to their own paper. They used analysis of variance (ANOVA) in conjunction with a linear response surface to conduct variable ranking in the context of successive response surface modelling, SRSM. While the latter will not be covered in the present thesis, it entails building surrogate models on subdomains of the design space, which are chosen by panning and zooming onto candidate optimal points in each iteration. By identifying the most important factors in an automotive structural optimization against frontal crash, Craig et al., 2005 found a reduction in computational time of more than 30 per cent when compared against the same problem with no variable screening performed.

Another optimization study on subfloor structures of helicopters was conducted by Astori and Impari, 2013, involving 4 variables using genetic algorithms. The main difference is that no high-fidelity simulations were made: a series of response surfaces were generated by conducting full multi-body analyses, which only took 20 seconds of CPU time for a 0.1 simulation time. The main advantage of using low-fidelity models is that a much larger number of training points can be generated. Although the responses are less accurate, the general trends may be uncovered much faster, and more accurate surrogate models can be built in the regions of interest.

Moving on to bird strike crashworthiness requirements, there are even fewer papers which take them into account. Pahange and Abolbashari, 2016 conducted a multi-objective mass and leading edge intrusion optimization of a riveted metallic wing structure against bird strike. The parameters taken into account were two different materials, three skin and rib thicknesses, three rib distances, three spar locations and three cut-out diameters of the nose ribs. Starting from a Taguchi design of experiments, which greatly reduces the number of simulations compared to a factorial sampling, for instance, the most important variables have first been identified by performing ANOVA. It was discovered that the most important variables were the skin thickness and material, while the spar location was insignificant. The multiple objectives, namely the mass and displacements obtained after a bird strike analysis, were converted to a single response function using grey relational analysis. However, this specific paper did not employ surrogate models. If metamodels were used, more designs could have been explored.

A research paper which encompasses most aspects presented in this chapter, though, is Ollar, Jones, and Toropov, 2017. Starting from the assumption that a bird strike is a local event, hence the significant components are only the leading edge skin, nose ribs, front spar, and the upper and lower skin panels behind the front spar, approximations are built in sub-spaces instead of the whole design space. Bird strike simulations have been made for multiple impact locations, which were assumed to be critical in the centre of each wing bay between consecutive nose ribs. Using a multiple-discipline feasible architecture, a surrogate model was built for the structural response of the wing impacted at each location, imposing a maximum intrusion penetration. A common feasible optimum design is sought after by the SQP optimizer.

The metamodel technique employed in this study was the moving least squares method, and the sampling was done with a method developed within Altair called “modified extensible lattice sequences”. However, Kriging and Latin Hypercubes are also cited as being suitable in terms of surrogate model and design of experiments, respectively. However, no variable screening study has been conducted, as the important variables were assumed based solely on engineering judgement. It should be noted, though, that this paper represents an application of an MDO methodology involving sub-space approximations, which is explained in greater detail in Ollar, Toropov, and Jones, 2017.

Jumping to more recently published papers, Lualdi et al., 2023 extended the SRSM method, which was also used by Craig et al., 2005, to incorporate multi-fidelity responses in a Gaussian process-based framework, resulting in the so-called Multi-Fidelity Successive Response Surface (MF-SRS) method. The latter has been tested on a 7-variable single-objective engineering use case of an aluminium crash-box for an urban electric vehicle, achieving a rather remarkable improvement over the traditional SRSM of approximately 14%. The low-fidelity models were obtained by the exclusion of an element failure criterion in the material definition, as well as the employment of a coarser mesh, thus considerably reducing the computational time from 20 minutes to 3.

2.2. Bird impact modelling

The present literature review will not go into any details regarding the material behaviour of metallic structures subject to high-speed impacts, as the difficulty is just a matter of finding those material laws which balance simplicity with the ability to capture the failure modes reasonably accurately. Moreover, each FE software has different laws implemented, therefore the formulations pertaining to Abaqus will be covered in section 3.1. What is more interesting, though, is how to include the bird itself in such a simulation.

2.2.1. Standard bird shapes used for analysis

Wilbeck, 1978 was the first to consider the bird to have a standard cylindrical shape. More recently, Heimbs, 2011 includes in his review a survey of the standard shapes used in the literature from 1979 to 2011 and revealed that, in practice, the wings, neck, or head of the bird are not modelled and that simple shapes are preferred, such as cylinders, cylinders with hemispherical ends or ellipsoids. In the Boeing 787 composite trailing edge certification, Georgiadis et al., 2008 modelled the bird as an oblate spheroid. Liu et al., 2017, who researched novel structures to improve the energy absorption capabilities against bird strike, have employed a cylindrical bird model, and the simulation has been validated against experimental results. However, in a follow-up study, Liu et al., 2018 have switched to hemispherical-ended cylinders.

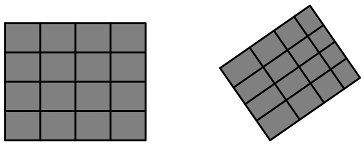
The consensus in the last years, though, seems to be that double hemispherical-ended cylinders are the preferred standard geometry for a bird. Investigations that employed this shape also include Di Caprio et al., 2019, who performed bird strike crashworthiness computer simulations and experiments on a vertical tail leading edge. Guida et al., 2022 settle the debate by conducting a comparative study between the following standard shapes: a right cylinder with the length being two times its diameter, a cylinder with hemispherical caps with a length to diameter ratio of two, and another one with the ratio approximately three. After validating the simulation results against experimental ones on a flat plate with a 4- and 8-pound bird, it was concluded that the cylinder with hemispherical caps and ratio 2 performed the best.

2.2.2. Bird modelling approaches

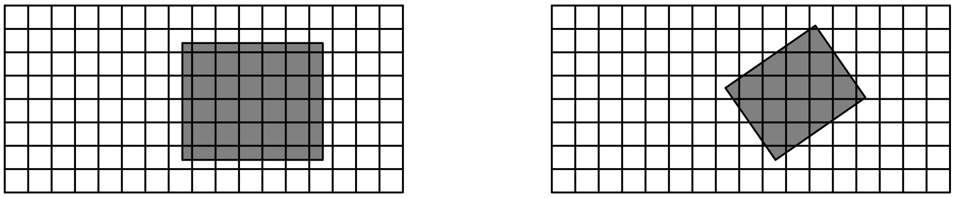
Now that the standard geometrical properties have been discussed, it is time to tackle the bird modelling aspect, and how it can be included in the finite element analysis. While there seem to be four possible ways to model the bird, only one is extensively used nowadays, especially since it has been included in validated methodologies such as Georgiadis et al., 2008. The first one is the classical method used in FEM, where the bird is modelled with Lagrangian elements. The second one is more appropriate for the hydrodynamic behaviour of the bird, and tries to employ a fully Eulerian approach. The third one is a mix of the last two, while the last one, Smoothed Particle Hydrodynamics (SPH) is a relatively newer, meshless method, which seems to be the panacea to bird modelling. Figure 2.1 presents these four approaches in the intuitive manner proposed by Heimbs, 2011, where the illustration is adapted from.

The first three methods have significant drawbacks. The Lagrangian finite element model may partially be able to depict the distortion and fragmentation of the bird, but it is not a feasible approach as the material cannot move relative to the mesh (Georgiadis et al., 2008). Additionally, the simulation time step must be significantly reduced as the bird distorts upon impact, which would greatly distort the mesh elements depicted in figure 2.1a, resulting in extensive computational time requirements, which in the end may not even provide the necessary accuracy. The Eulerian approach, where the mesh is completely fixed in space, and the material of the bird is free to move through it, as illustrated in figure 2.1b, can be used without the numerical problems associated with excessive distortion. Heimbs, 2011, who conducted a review of computational methods for bird strike simulations, raises attention to the fact that problems arise when both Eulerian and Lagrangian elements are used together. When the leading edge deforms, the Lagrangian mesh moves accordingly, and the Eulerian mesh should also start covering the region left behind by the protruded leading edge. It is then clear that the Eulerian computational domain needs to be rather large, thus being much more demanding, and it usually also requires a much smaller element size (Heimbs, 2011). Moreover, the contact between the fluid-like bird and the structure should be carefully defined.

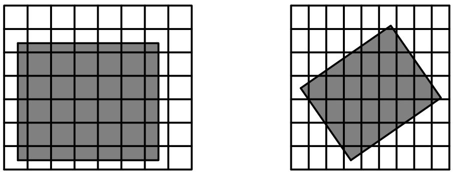
Another approach which aims at overcoming the disadvantages of both of the methods aforementioned is Arbitrary Lagrange Euler (ALE). The bird is modelled as a slug of fluid moving through an Eulerian box, which is no longer fixed in space, but can move freely to accommodate the expansion of the bird and the deformation of the leading edge, as depicted in figure 2.1c. From a computational point of view, it is much better than the fully Eulerian method, mainly because the mesh has a much smaller number of elements. While it does seem like the inherent problems with the Lagrangian method have been fixed by this approach, it still has problems with the material interface, as reported by Georgiadis et al., 2008. Moreover, it should be noted that while the bird spreads out upon impact, the Eulerian mesh deforms and expands in order to capture the material flowing through it. As stated by Heimbs, 2011, the mesh can encounter severe elongations, especially during the steady flow regime, which can jeopardize the accuracy of the model.



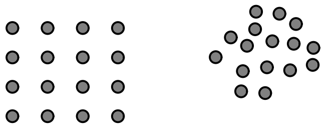
(a) The Lagrangian model, where the material cannot move relative to the mesh.



(b) The Eulerian model, where the mesh is fixed and the material moves freely through it.



(c) The Arbitrary Lagrange Euler model, where the mesh is not fixed, but the material moves freely through it.



(d) The Smoothed Particle Hydrodynamics model, where the material is a collection of particles, with no mesh involved.

Figure 2.1: Various approaches found in the literature to model the bird in a computer simulation.

Note: Illustration adapted from Heimbs, 2011.

In order to counteract the inherent deficiencies of the methods aforementioned, a technique called Smoothed Particle Hydrodynamics was devised in the late 1970s to address material fragmentation issues resulting from high-velocity impacts (Liu et al., 2017). The SPH approach is a meshless method, where the bird is modelled with particles which are not interconnected, as illustrated in figure 2.1d, thus allowing it to deform and split freely, while also being less computationally intensive than ALE, as it requires fewer elements (Liu et al., 2017). It is a widely used method, and Georgiadis et al., 2008 state that it is probably the most common method to investigate bird strike phenomena.

The bird is discretized into a number of particles, each one representing a portion of the bird's mass, and possessing fluid characteristics such as density, pressure, and velocity. If two particles are within a certain distance known as the smoothing length, they interact with one another based on a volume- and kernel-weighted interpolation technique (Siemann and Ritt, 2019). The details regarding these functions and the smoothing length will not be covered in the present report. Nonetheless, the SPH approach is based on a Lagrangian formulation, therefore material interface problems no longer happen (Heimbs, 2011), even if contact penalties are still needed (Siemann and Ritt, 2019).

Studies on the effect of particle distribution on bird strike simulations have not been conducted until Siemann and Ritt, 2019, at least to the knowledge of the authors of the article. Therefore, they suggest utilizing the Weighted Voronoi Tessellation algorithm as the basis for an iterative method to create initial particle distributions in arbitrarily shaped domains. The proposed technique has been proven to generate superior distributions when assessed on a flat plate example, compared to the approach relying on the conversion from an initial FE mesh, which is a readily available method in various FEA software. The latter presented some alignment, which was also demonstrated to decrease with the number of particles. The proposed bird model needs a smaller number of particles to offer the same accuracy, thus reducing the computational time.

2.2.3. Bird equation of state

While in the last two sections it was apparent that a consensus has been reached with respect to the standard bird shape and modelling approach, the equation of state (EOS) seems to be an ongoing debate. Hedayati and Sadighi, 2016 offer a thorough review on FE bird strike modelling, including a comprehensive survey on equations of state, together with their associated mathematical formulae. It is highlighted, however, that no analytical pressure-volume relationship has been found for liquids, contrary to gases, prompting all equations of state to be semi-empirical. The following formulae are included: linear and polynomial EOS, tabulated (where no analytical formulation is proposed), and Mie-Grüneisen.

Guida et al., 2022 have shown that the EOS of the mixture between 90% water and 10 % air gives accurate results, and implemented it in LS-Dyna in tabular form. Liu et al., 2018, on the other hand, utilize the Murnaghan EOS implemented in PAM-CRASH. Di Caprio et al., 2019 used the Mie-Grüneisen EOS, although they do not give any material constants. Finally, Marulo and Guida, 2014 utilize a tabular form of the equation, and also provide the data, which will be used in the present study.

2.3. Design of experiments

The simplest and most intuitive sampling plan is probably the full factorial design, where the number of training points is equal to the product between the number of levels for each factor. However, this number quickly becomes prohibitive for a high-dimensional problem. A way to circumvent the exponential growth of the number of training points with the number of variables is by using fractional factorial designs, which are simply a fraction of a full factorial design. According to Simpson et al., 2001, who conducted a comprehensive survey on metamodels for computer-based engineering design, the most popular ones are 2^{k-p}, which can be seen as a hypercube where only a fraction of the vertices are used for sampling, and which allow the identification of important variables under the assumption that the sparsity of effects principle holds. The latter states that, when there are many variables, the responses are dominated by main effects and low-order interactions, but Simpson et al., 2001 raise attention to the fact that it is not always true, and, moreover, that the effects cannot be estimated independently, as they are aliased with the two-factor interactions.

Placket-Burmann designs are a particularization of the fractional factorial designs where the effects of $k = n - 1$ variables can be studied with a design space of $n = 4m, m \in \mathbb{Z}$. The Central Composite design, on the other hand, is a 2-factor half-fractional or full factorial design, augmented by a center point and additional so-called “star” points positioned at $\pm\alpha$ for each factor. While this method seems to be very popular among aerospace engineers (Simpson et al., 2001) as quadratic effects can also be estimated, it is not suitable for high dimensional problems, as the number again increases exponentially. The Box-Behnken design, on the other hand, can be seen as a k-dimensional hypercube, where the training points are all in the midpoints of the edges, thus requiring less samples. Simpson et al., 2001 also mention that not sampling at the vertices of the hypercube is desirable especially if the responses at extreme points are not acquired easily,

Last but not least, orthogonal arrays and Taguchi designs are partial factorial designs where for each level of a variable, all the data points containing it are such that all the levels of the other variables appear at least once. Taguchi designs are rather simple to implement, as the orthogonal arrays can be readily found online. For instance, in their non-surrogate model-based mass and performance optimization of a wing leading edge against bird strike, Pahange and Abolbashari, 2016 employed a Taguchi DoE, while also studying the main effects and interactions between variables using analysis of variance. The latter will be covered in section 3.4, where the study of such effects is going to be discussed.

All the designs mentioned in this section are quite uniform, as seen in figure 2.2, but they possess a significant drawback: the projections of the data points on each axis have overlaps. As mentioned in chapter 1, a variable ranking study is usually conducted to reduce the number of variables in expensive, high-dimensional non-linear problems, as recommended by Viana et al., 2021. While it has been emphasized that these designs allow for the estimation of main effects and interactions between variables, after variable elimination, the next step would be to project the training points onto the remaining dimensions. Because of the overlap aforementioned, many training points would essentially be wasted.

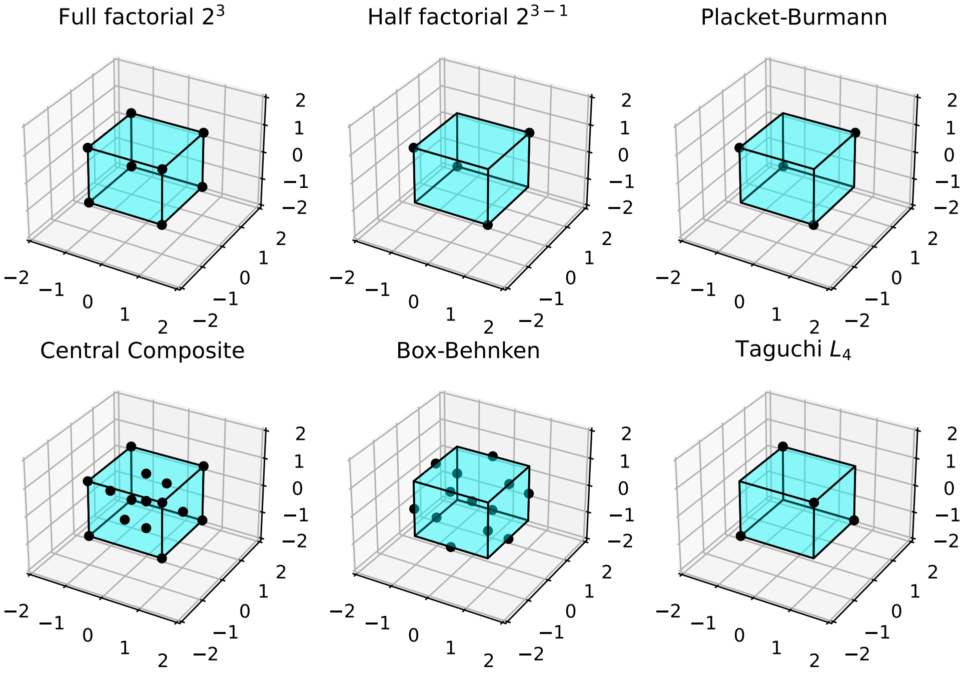


Figure 2.2: Various designs illustrated on a $[-1, 1]^3$ cube.

The problem is solved by employing the so-called Latin Hypercube design (LHD). Intuitively, if n samples are needed, each of the k dimensions is split into n intervals, and the samples are spread into the n^k resulting cells of the hypercube, such that each interval of each dimension is represented once. For a two-dimensional case, an LHD can be seen as an $n \times n$ Sudoku table, where each row and column contain exactly one point. This sampling strategy has been first put forward with respect to deterministic computer codes by McKay et al., 1979, and is also being employed in the surrogate model-based optimization algorithm of Jones et al., 1998, which will be discussed in section 3.5.

It is especially advisable to use Latin Hypercube samples during the initial phases of the design process, when the trends of the surrogate model are unknown and cannot be prespecified, thanks to their space-filling nature. Moreover, they are good for high-dimensional problems, and in conjunction with a variable screening procedure, they provide good results for problems with more than 10 variables (Viana et al., 2021). Traditionally, LHDs are only used with continuous variables. The Sliced Latin Hypercube Design (SLHD) put forward by Qian, 2012 integrates categorical variables, by introducing as many slices as categorical levels. Both the traditional LHD and its categorical variant will be discussed in more detail in section 3.2.

2.4. Surrogate models

Successive response surface modelling, SRSM, has been mentioned quite a few times in the section on crashworthiness optimization, and it does seem to have traction in the automotive industry. However, the focus of the present work will only be on Kriging/Gaussian processes (GPs), as they constitute the foundation of SRSM, and it is preferred to first gain familiarity with the basis before delving into more advanced models. The present section will first explore recently developed variants of the Kriging metamodel, followed by a discussion on ensembles of surrogates, which have seemingly built momentum during the last two decades.

2.4.1. Kriging variants

The Kriging model, also known as Gaussian process regression, is a statistical method used for spatial interpolation and regression analysis. By considering that the observations are following a multivariate Gaussian distribution, a posterior normal distribution of the response at an unsampled point can be derived. Therefore, the surrogate provides both predictions and error estimates. Traditionally, the model consists of a polynomial regression term, summed with the realization of a stochastic process, which will be detailed in section 3.3. Although the method was first put forward in 1951 by Danie Krige, it is widely used nowadays, and numerous variants have recently emerged.

Blind Kriging

Originally, the regression term of the Kriging model is assumed by the user to have a specific polynomial degree. On the other hand, Blind Kriging (Joseph et al., 2008) represents a modified Kriging where the regression functions are not assumed to be known anymore. The mean model is identified via a Bayesian forward variable selection technique. Joseph et al., 2008 also present an improvement in prediction when compared to the ordinary Kriging by testing the two models in an engine block optimization process. Although the accuracy is better, the computational cost seems to be double. However, the training of a surrogate model is probably much shorter than the evaluation of expensive functions such as the structural response against a bird strike; therefore, a lengthier training should not be seen as a setback if the accuracy would greatly increase and less training points would be required to obtain an acceptable predicted error of the model.

Cokriging

Cokriging is an extension of Kriging for functions with multidimensional responses (Viana et al., 2014), which also uses cross-correlations between the individual responses to build better predictions. Therefore, it might be a valuable tool to use if some of the constraints are expected to be correlated. Jones et al., 1998 also mention that Cokriging can be used for multi-fidelity surrogate modelling. The main idea behind it is that the responses given by low- and high-fidelity analyses can be seen as correlated variables. Essentially, the cheap outputs of the low-fidelity analyses can be used to quickly identify trends of the modelled functions in order to increase the accuracy of the high-fidelity response surface.

Kriging with integer or categorical variables

Traditionally, the Kriging model employs continuous variables. Garrido-Merchán and Hernández-Lobato, 2020 propose a modification to include mixed-integer variables. Integer variables are simply rounded to the closest value, while categorical variables are treated via the one-hot encoding procedure, which replaces them with as many variables as categories. A new covariance function based on this transformation is proposed, which allows for the underlying GPs to yield a constant objective function in the regions of the continuous input space. It is acknowledged that this approach leads to an increase in the number of dimensions, but prediction error estimates and all other attractive qualities of the traditional Kriging model are conserved.

2.4.2. Ensembles of surrogates

Bishop, 1995, pp. 364–371, first introduced the idea of constructing ensembles by performing a weighted sum of more surrogate models. However, it was only during the last two decades that this idea started to gain traction in the structural crashworthiness field. Acar and Solanki, 2009 have expressed the idea that no unique surrogate model is the best for all problems, and that it is also not always apparent why a specific metamodel is better than the other. Contrary to the practice of constructing multiple metamodels based on a common training dataset and selecting the best one while discarding the others, Acar and Rais-Rohani, 2009 contend that this approach is wasteful, and challenge the assumption that changes in the training set do not impact the choice of the best model.

Zerpa et al., 2005 constructed an ensemble of surrogates by assigning weights to each model inversely proportional to the point-wise estimate of their prediction variance. Being a local measure, the weights then vary over the design space. Goel et al., 2007, on the other hand, propose a global weighting scheme such that the factors are inversely proportional to the cross-validation errors of each model. The idea that regions of high uncertainty can be identified by ensembles has also been put forward, as a large disagreement among the surrogates in certain areas would indicate a high uncertainty.

Viana et al., 2009 also propose to minimize the cross-validation error. Firstly, they put forward a method of predicting the root mean square error (RMSE) from the cross-validation error. Afterwards, an analytical solution is proposed to determine the weights such that the approximated mean square error is minimized. Unfortunately, they found that the gain in accuracy decreases with the number of dimensions.

Acar and Solanki, 2009 and Acar and Rais-Rohani, 2009 suggest yet again to find the weights of the models such that the cross-validation error is minimized; however, they consider the weight factors as design variables in an optimization problem. One year later, Acar, 2010 proposed a more sophisticated weighting method, which varied over the design space. However, it was found that constant, global weights performed better for high-dimensional problems.

X. J. Zhou et al., 2011 propose a recursive scheme: the main idea is to calculate the root mean square error of all individual surrogates, and successively replace the worst-performing surrogate with the arithmetic average of all candidate surrogates, until the RMSE of the worst one is close to the best one within a user-specified tolerance. The advantage of this method is that it is much faster than the optimization problem proposed by Acar and Rais-Rohani, 2009, and it also ensures that the weights are between 0 and 1, which neither Acar and Rais-Rohani, 2009 nor Viana et al., 2009 do.

In conclusion, ensembles seem to only be useful for low-dimensional problems, and only if the surrogates are of similar accuracy but offer different predictions at unsampled points (Viana et al., 2009). Moreover, especially in the context of optimization, the global accuracy of a surrogate model is not as important as its ability to find the global optimum. Another problem arises as well: how would the prediction error estimate of an ensemble of surrogate models be computed, which are paramount to Bayesian optimization? And can they be used for variable screening? For this reason, the idea of employing ensembles is not going to be pursued in the present study.

2.5. Variable screening procedures

As mentioned in the introductory chapter, the “curse of dimensionality” outlined by Koch et al., 1999 poses a challenge in high-dimensional optimization. It involves the need for a substantial number of training points to maintain an adequate sample density across the entire design space. Due to this setback, Viana et al., 2021 propose to reduce the dimensionality of the problem before proceeding with the optimization process.

The simplest and probably most popular method is stepwise regression, which was first put forward by

Efroymson, 1960. There are three variants, all of which start with a linear model. Forward stepwise regression adds variables one by one according to a criterion, such as lowest prediction or leave-one-out cross validation error, as outlined by Desboulets, 2018. Backward stepwise regression starts with all variables included, and eliminates them one at a time, while the bidirectional variant is a combination of the two. All of these methods have a significant drawback, regardless of the chosen criterion: they are not exploratory. Indeed, the procedures move along only one path, as mentioned by Desboulets, 2018, who also highlights in his review that an abundance of papers proved inconsistent selection of stepwise regression. Moreover, the significance of a variable cannot be quantified.

More sophisticated techniques are variance-based, the simplest method relying on an initial fractional factorial or Taguchi design, followed by an analysis of variance procedure, as conducted by Pahange and Abolbashari, 2016. The main problem with this approach is related to the necessary design of experiments, where training points may become confounded after projecting the design space onto the reduced dimensions. For this reason, Saltelli and Sobol', 1995 proposed the Global Sensitivity Analysis (GSA), which computes the contribution to the total variance of a set of variables starting from an initial surrogate. This idea is attractive because not only are the contributions quantified, but main and interaction effects can also be estimated. Moreover, the method has been applied to the Kriging surrogate by Schonlau and Welch, 2006.

Another dimensionality reduction method is the Partial Least Squares (PLS), which establishes a linear correlation between input and output variables. In order to do so, input variables are projected into a reduced-dimensional space that is made up of principal components, the latter being a linear combination of the initial variables. An advantage of this method is that it does not require any initial surrogate model to be conducted. However, it does not quantify the main and interaction effects of various variables, and the resulting set of variables does not offer much intuitive information about their relationship with the initial one.

Nevertheless, it is a promising method if these quantitative effects are not necessarily desired. An interesting recent advancement has been made by Bouhlel et al., 2016, who propose the Kriging Partial Least Squares (KPLS) model. Firstly, the traditional PLS method is applied to define the coefficients of the variables in the resulting linear combination. Afterwards, a new covariance function is constructed based on these weights, thus combining the advantages of the Kriging metamodel with the PLS method. Not only is the dimensionality reduced, but the linear mapping is adjusted with each additional training point, thus providing an inherent recovery mechanism against erroneous identification of principal components.

2.6. Bayesian optimization

The expected improvement acquisition function is arguably the most widely used in the context of Bayesian optimization, which is apparent in the samples of papers related to optimization between 1998 and 2010 conducted by Viana et al., 2021. Not only does the method balance exploration and exploitation, but it can easily accommodate constraints as well. Practically, the expected improvement makes use of the posterior normal distribution given at a sampling point, the mean and variance being given by the Gaussian process predictor, and derives an improvement of the objective function which can be reasonably expected. Constraints can be easily taken into account by multiplying the expected improvement with the probability of each constraint being met. Based on these benefits, Jones et al., 1998 propose the Efficient Global Optimization (EGO) algorithm, which performs a single-objective unconstrained Bayesian optimization employing a Kriging surrogate. In the same year, Schonlau et al., 1998 adapt the EGO algorithm to take into account multiple unknown constraints as well, where each of them is modelled with a separate Kriging model.

A downside of EGO is that it can be easily tricked by deceiving functions or unfortunate samples, as outlined by Jones, 2001 in a comprehensive taxonomy of global optimization methods. For instance, consider a highly multi-modal function, whose surface has various hills and valleys. If the training dataset is unfortunate enough such that all observations are nearly the same in a region, the surrogate may predict low variances, and prompt the algorithm not to search there, although the region actually contains a valley. To counteract this deficiency, Forrester and Jones, 2008 propose a so-called one-stage modification of the unconstrained EGO algorithm. The infill points are no longer computed based solely on the prediction error estimates of the surrogate model; as Forrester and Keane, 2009 put it, in one-stage strategies, the infill criterion is used to calculate the surrogate model, not the other way around. Practically, both the surrogate itself and the next sampling point are included in an optimization problem, which finds the next sampling point and model hyperparamaters such that a likely goal of the objective function is achieved.

2.7. Research questions

The previous sections have shed light not only on research trends within the crashworthiness structural optimization field, but also on a number of gaps with respect to the incorporation of bird strike constraints.

The general trend is to employ surrogate models, mainly due to the high computational times associated with gathering data from crashworthiness analyses. Moreover, the metamodels must be able to capture the expected non-linear and multi-modal landscapes, while also offering error prediction capabilities to be used in conjunction with Bayesian optimization. As stated earlier, the Kriging metamodel will be used, which will be detailed in section 3.3.

Another recent trend is associated with multi-fidelity modelling, which is an interesting direction of research, especially since high-fidelity impact analyses are time-consuming. However, the development of low-fidelity bird impact analyses is an arduous task, which will not be tackled in the present thesis, as it deserves an in-depth study on its own. No other gaps have been identified with respect to bird strike finite-element analyses, as the most suitable standard bird shape and modelling approach have been found, together with an appropriate equation of state. All of the aspects aforementioned will be detailed in section 3.1.

Moving on to the construction of surrogate models, the Sliced Latin Hypercube sampling has been determined to be a suitable design of experiments, as it is space-filling, no two samples have any equal coordinate, and categorical variables can also be included. More details on this design are given in section 3.2. A number of interesting variants of the Kriging metamodel have been reviewed, which tackle different regression terms, cross-correlations among high-dimensional outputs, and mixed-integer variables. An interesting trend was related to ensembles of surrogates, which does not seem to give much better results than individual models in high-dimensional problems. Moreover, their use also raises numerous other problems, such as the evaluation of a Bayesian acquisition function, and their incorporation in a variable screening procedure. The ordinary Kriging metamodel will be covered in section 3.3.

Several dimensionality reduction methods of different complexities have been identified, with variance-based approaches and Partial Least Squares being the most promising. However, the former requires the construction of an initial surrogate, while the latter does not quantify the main and interaction effects between variables. As a consequence, the variance-based method put forward Schonlau and Welch, 2006 will be used in the present study, in order to allow a more intuitive interpretation of the reduced set of dimensions, and will be presented in section 3.4.

Finally, the constrained efficient global optimization algorithm will be used for the optimization process itself after variable reduction, which is based on the constrained expected improvement acquisition function, and will be covered in section 3.5. The one-step approach put forward by Forrester and Jones, 2008 is an interesting method to counteract the inherent deficiency of the EGO algorithm, as it can be easily tricked by deceiving functions and unfortunate samples. This approach is presented in appendix B. Nevertheless, because constraints are not taken into account in this method, it cannot be employed in the present study.

Moving back to bird strike structural optimization studies, a number of additional gaps have been identified. As stated earlier, both Pahange and Abolbashari, 2016 and Ollar, Jones, and Toropov, 2017 tackle only the leading edge indent as an objective and constraint function, respectively. However, it is hypothesized that both the penetration depth and the residual strength of the wing are important, especially since EASA CS 25.571(e) requires the damaged aircraft to be capable of completing a flight under the loads associated with get home conditions. To the knowledge of the author, no optimization study has been carried out which takes into account the residual strength of the damaged structure. Therefore, the following research questions are proposed:

- What constraint function can be formulated, such that it is sufficiently representative of the feasibility of a design to ensure compliance with bird strike crashworthiness regulations?
- What is a suitable approach to incorporate the proposed function in a multidisciplinary optimization framework?

To answer these questions, a constraint function will be proposed in chapter 4, coupled with another constraint based on the maximum penetration depth. As Pahange and Abolbashari, 2016 have also conducted a variance-based variable screening process, although without employing surrogate models, the following research topics are advanced:

- What are the limitations of the proposed variance-based variable screening method with respect to the constraint functions?
- Which additional criteria can be formulated to safeguard against an erroneous identification of significant variables?
- How can the optimization problem be formulated if the objective and constraint functions are defined on different reduced design spaces?
- How does the global accuracy of the surrogate model change with the reduction of the design space?

The variance-based variable screening procedure will be described in section 3.4. Moreover, the two constraint functions will have a different number of significant variables after screening, and the formulation of the optimization problem will be described in chapter 5, when presenting the proposing methodology.

Finally, as surrogate model-based optimization has been identified as a suitable approach for the incorporation of bird strike requirements and Gaussian processes are proposed to be employed, the following research questions are put forward:

- Given that global accuracy is not of great interest during the optimization process, what is a sufficient number of training points for the initial surrogate model?
- Due to the large dimensionality of the problem, coupled with a high computational intensity of impact simulations, what is a suitable measure to assess the quality of the surrogate model of the proposed constraint?
- What is a suitable convergence criterion for the optimization process?

These research questions will be explicitly addressed in the last chapter, which will provide an overview of the proposed methodology, along with any shortcomings it may demonstrate during its application.

3

Theoretical aspects

The present chapter will cover all relevant theoretical aspects of the proposed methodology. As outlined earlier in the introductory chapter, the following steps will be undertaken: firstly, an initial sampling needs to be done in order to construct a preliminary surrogate model. Afterwards, a variance-based screening study will be conducted in order to evaluate the importance of the design variables. After the insignificant parameters are eliminated, the initial sampling is projected onto a new design space of reduced dimension. A surrogate model will then be trained in the new space, and new data points will be obtained in regions of interest via Bayesian optimization. After a convergence criterion is met, the optimum will finally be discovered.

Therefore, the following areas will be covered in the present chapter: bird impact modelling, design of experiments, Gaussian processes, variable ranking, and Bayesian optimization. While these aspects will not be put forward in the order of their usage during the proposed methodology, this manner of presentation will hopefully provide a clear overview on the relevant theory. However, for aspects concerning material laws, certain parallels are going to be drawn with their implementations in Abaqus due to the numerous plasticity models available, thus somewhat deviating from a pure theoretical presentation.

3.1. Bird impact modelling

As specified in the introductory chapter, Abaqus will be used to evaluate the constraint responses during the optimization procedure. Before delving into the modelling aspects, though, it is important to swiftly go over what actually happens during a bird strike from a physical point of view. In a review of current research on simulating bird strikes and hail impacts on wing leading edges, Goraj and Kustron, 2018 provide an excellent and intuitive explanation of the behaviour of a bird during a high-velocity impact event: the projectile undergoes rapid deceleration, resulting in the generation of stress levels that significantly exceed the yield stress of the material. Such behaviour is characteristic to the hydrodynamic regime, where the projectile can be regarded as a fluid. As a result, the density of the bird dominates its behaviour, rather than its inherent material strength.

The hydrodynamic behaviour has been proven by Wilbeck, 1978, whose paper can be considered as one of the most influential works concerning the typical response of a bird upon impact with a structure. In order to standardize the material of the bird, an extensive experimental campaign was conducted utilizing both real birds and artificial projectiles composed of gelatin or room-temperature-vulcanizing rubber. Remarkably, it was demonstrated that gelatin replicas were able to faithfully capture the observed behaviour, substantiating the notion that the impact of a low-strength projectile, such as a bird, can be modelled as an unsteady and hydrodynamic phenomenon, where birds exhibit fluid-like behaviour at typical flight velocities.

Guida et al., 2022, who also conducted a review on bird impact processes, offer a nice addition to the explanation of Goraj and Kustron, 2018: during the propagation of the shock wave through a bird due to the high velocity of the impact, the material at the edges of the bird experiences a sharp pressure gradient. This results in the radial acceleration of the material away from the centre of impact and the consequent generation of a release wave. The primary function of the latter is to alleviate the radial pressure caused by the impact. However, the shear stresses experienced by the bird greatly exceed its shear strength; thus, a condition of steady flow is finally established, where the bird has a hydrodynamic behaviour. Figure 3.1 illustrates the behaviour aforementioned, while also depicting schematically the shock wave travelling through the bird, and the radial displacement of the material when it comes into contact with the target. The bird is represented by a cylinder, as suggested by Wilbeck, 1978. However, this shape is only used for illustration purposes, as it was revealed in chapter 2 that is suboptimal in terms of simulation fidelity.

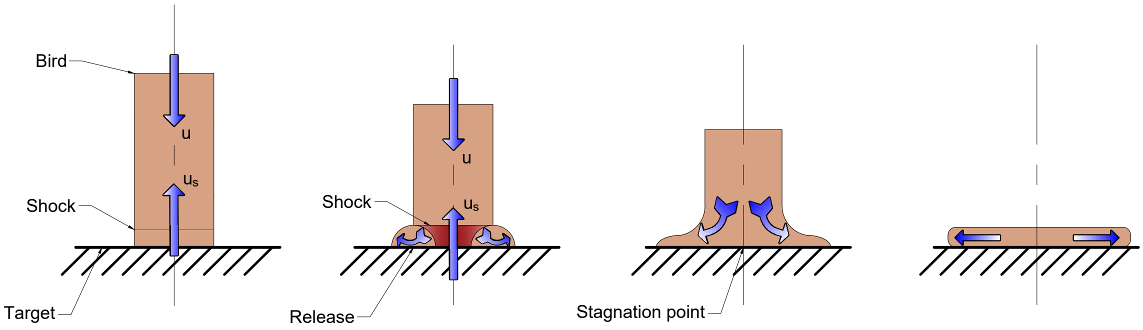


Figure 3.1: The hydrodynamic behaviour of a bird during impact with a rigid target.

Note: Illustration adapted from Wilbeck, 1978.

3.1.1. Standard shape

During experiments and computer simulations alike, the bird strike process is seen as (approximately) equivalent to a stationary wing leading edge being impacted by the bird flying at the aeroplane speed mentioned by the EASA CS 25 certification requirements, which were discussed in section 1.2. As the present paper is focused mainly on structural optimization, no further details will be given with respect to the hydrodynamic behaviour of the bird, apart from what is needed to conduct an accurate computer simulation. It was mentioned in the literature review of chapter 2 that a double hemispherical-ended cylindrical bird shape with a length-to-diameter ratio of 2 performs best during impact simulations, as proven by Guida et al., 2022 in their comparative study.

Having the length-to-diameter ratio, the geometrical parameters can be easily computed. Both Barber et al., 1975 and Wilbeck, 1978 give an average bird density of 950 kg/m^3, and many other authors use it, such as Liu et al., 2017. It should be noted, though, that the study of Budgey, 2000, who aids in establishing a standardized bird throughout the aerospace community and surveyed thirty bird species, arrived at an average of 943 kg/m^3. Nevertheless, the former value will be used in the present study, a higher density providing more conservative results due to the increase in kinetic energy. The mass of the bird is known, as well, from the certification requirements, with a value of 1.81 kg. As the volume of the bird with the shape aforementioned is:

$$\begin{aligned} V_{bird} &= V_{cylinder} + V_{sphere} \\ &= L\frac{\pi D^2}{4} + \frac{\pi D^3}{6} \\ (L = 2D) &= \frac{2\pi D^3}{3} \\ &= \frac{m_{bird}}{\rho_{bird}} \end{aligned} \tag{3.1}$$

With the above formulae, a diameter of approximately 96.89 mm is found, giving a length of 193.78 mm.

3.1.2. The Smoothed Particle Hydrodynamics bird model

As mentioned in the literature review in chapter 2, the SPH method is by far the most widely used technique to model the bird, having been employed in validated methodologies such as Georgiadis et al., 2008. The SPH method is readily available in Abaqus, and the particle nodes can either be defined directly, or a conventional finite element model can be converted to particle elements.

However, the former approach will be undertaken. GKN Fokker will supply the bird model, which is generated based on the method proposed by Siemann and Ritt, 2019. As outlined in the literature review, the latter is based on an iterative method based on the Weighted Voronoi Tessellation algorithm, which first creates an initial particle distribution in arbitrarily shaped domains and then refines it until alignment issues are alleviated. A significant benefit of the method is the reduction of the computational time as a smaller number of particles is needed for a similar accuracy.

Concerning the number of particles, Guida et al., 2022, for instance, used 16,086 particles for the 4-pound bird, while Liu et al., 2017 used 15,360. Nevertheless, for a faster computational time, 10,000 particles will be employed in the present study, although there may be a slight decrease in accuracy. The bird model, as rendered by Abaqus, can be observed in figure 3.2.

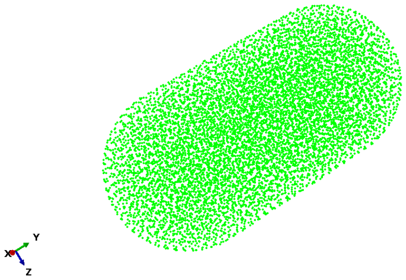


Figure 3.2: The 10,000-particle SPH bird model used in the present study.

3.1.3. Equation of state, EOS

Because the bird has a hydrodynamic or fluid-like behaviour during impact, its material needs an equation of state associated with it. There seems to be a debate, though, regarding which relationships are the best, as mentioned in chapter 2. As an example, both Liu et al., 2018 and McCarthy et al., 2004, who conducted their analyses in PAM-CRASH, use the Murnaghan EOS. However, this type of equation is not available in Abaqus, unless a user-defined EOS is defined; as there are alternatives, this option will be dismissed.

Di Caprio et al., 2019, who did employ Abaqus Explicit, has chosen the Mie–Grüneisen EOS, but no material constants were provided in the paper. The only information given is that the EOS of the bird has been calibrated using experimental pressure data.

Finally, the easiest implementation would be the tabulated EOS. Guida et al., 2022 have shown that the EOS of the mixture between 90% water and 10 % air gives accurate results, and implemented it in LS-Dyna in tabular form. Both LS-DYNA and Abaqus relate the pressure p to the volumetric strain ε_V and the specific internal energy, E_m. The following equation, as given by Dassault Systèmes, 2022g, presents the pressure-volume relationship:

$$p = f_1(\varepsilon_V) + \rho_0 f_2(\varepsilon_V) E_m, \varepsilon_V = ln\left(\frac{\rho_0}{\rho}\right) \tag{3.2}$$

In equation 3.2, the tabulated input requires the specification of the $f_{1,2}$ quantities as functions of the volumetric strain ε_V, which will then be linearly interpolated by Abaqus. Guida et al., 2022 also explain that the temperature does not contribute significantly to the pressure and it is hence negligible. But the temperature is responsible for the specific internal energy, as the latter is characterized by the thermal energy state of the fluid, $E_m \approx c_v \Delta T$, c_v being the specific heat capacity at constant volume and ΔT being the change in temperature versus a reference state. Therefore, the last term in equation 3.2 is negligible as a whole, and f_2 can be given a constant null value.

The remaining parameters are then easy to define. The reference density simply takes the assumed value of 950 kg/m^3 given earlier, and the tabular data relating the pressure to the volumetric strain is readily given by Marulo and Guida, 2014. The actual values will be given in the next chapter, section 4.3.1, when discussing the Abaqus bird strike model in more detail for a specific case study.

3.1.4. Wing leading edge structure finite element modelling

When a projectile strikes an object, its kinetic energy is converted into heat and sound energy through the induced deformations and vibrations. However, these processes cannot take place instantaneously, especially in high-velocity collisions where there is insufficient time for them to occur. This causes the struck material to behave as if it were more brittle than it would be under normal circumstances. The dynamic structural analysis of a leading edge during a bird strike involves inelastic strains, high strain rates and large deformations (Georgiadis et al., 2008). Therefore, even if it has been mentioned that the analysis will be simplified and only metallic parts will be considered, the material law to be used still needs to be carefully considered.

The Cowper-Symonds constitutive law has been employed to model the plastic response for metallic parts in many studies, such as Liu et al., 2017, Liu et al., 2018, Di Caprio et al., 2019 and Guida et al., 2022, to name a few. Liu et al., 2018 also provide the material parameters for two types of aluminium, therefore the Cowper-Symonds law will be given here using their notations:

$$\sigma(\varepsilon, \dot{\varepsilon}) = [a + b(\varepsilon_p)^n] \left[1 + (\dot{\varepsilon}/D)^{1/p}\right] \tag{3.3}$$

For aluminium 2024-T3, the parameters are readily given in Liu et al., 2018: $a = 350$ MPa, $b = 426$ MPa, $n = 0.34$, $D = 6500 s^{-1}$, $p = 1$. Beware of confusing notations, as p is just a material constant, not a pressure. In Abaqus, though, the Cowper-Symonds law is not readily implemented, an alternative being the Johnson-Cook model (Dassault Systèmes, 2022a). Both models are empirical, the only difference laying in the strain rate dependency. As a consequence, the Johnson-Cook law is going to be used in the present thesis. This model entails the following:

$$\sigma(\varepsilon, \dot{\varepsilon}) = [A + B(\varepsilon_p)^n] [1 + C \ln \dot{\varepsilon}] \tag{3.4}$$

It is to be noted, though, that equation 3.4 does not include any temperature dependency. For simplification purposes, it will be assumed that temperature does not play a major role in the plasticity model, although it is acknowledged that high strain rates may result in an increase in temperature. However, papers such as Liu et al., 2017, Di Caprio et al., 2019 or Georgiadis et al., 2008 do not include temperature in their plasticity models either.

Moving back to equation 3.4, fortunately, for this rate-dependent plasticity model, the material constants of the aluminium employed in the case study are readily given in Lesuer, 2009. The actual values will, once more, be given in chapter 4, along with a more detailed description of its implementation, from a practical point of view.

Regarding the failure model, though, the Johnson-Cook criterion is available only in Abaqus Explicit. As will become apparent later in the next chapter, an Abaqus Standard analysis will be conducted onto the damaged wing structure. As the material state is transferred along with the material definition, the implementation of the Johnson-Cook failure model is, unfortunately, not an option.

Vershinin, 2015, who validated various metal plasticity and fracture models on a high velocity perforation application, has investigated both the Johnson-Cook and ductile fracture criterion, leading to good agreement with the experimental results. What is fortunate is that the cited paper employs Abaqus Explicit for its numerical simulations, and provides tabular data for the ductile fracture criterion, which is dependent on the plastic strain rate, stress triaxiality, and the so-called Lode angle, which appears in the normalized third deviatoric stress invariant. However, the last parameter may or may not be included in the model. Therefore, as a simplification, it will be disregarded, as in the case of Bao, 2003. After all, the utilized plasticity model is not elaborate enough to entail an equally complex fracture model. Without said dependency, the ductile criterion becomes, with the notation of Dassault Systèmes, 2022f:

$$\omega_D = \int \frac{d\bar{\varepsilon}^{pl}}{\bar{\varepsilon}^{pl}_D(\eta, \dot{\bar{\varepsilon}}^{pl})} = 1 \tag{3.5}$$

ω_D can be seen as a variable which increases as more plastic deformation occurs, until reaching the value of 1, when the onset of damage occurs. The equivalent plastic strain corresponding to damage initiation, $\bar{\varepsilon}_D^{pl}$, is a function of the stress triaxiality, $\eta = -p/q$, and the equivalent plastic strain rate, $\dot{\bar{\varepsilon}}^{pl}$. The stress triaxiality is defined as the ratio between the pressure stress $-p$ and Mises equivalent stress q. In Abaqus, the parameters necessary for this criterion are given in tabular form, giving data points for $\bar{\varepsilon}_D^{pl}$, η and $\dot{\bar{\varepsilon}}^{pl}$.

According to Bao, 2003, the stress triaxiality and the equivalent plastic strain are the most significant contributors to crack formation. Different complex mechanisms, though, are observed for high and low stress triaxiality factors. For the former range, void nucleation is the first step to crack initiation, accompanied by growth and linkage. In the negative range, though, shear fracture occurs. While fracture mechanics is out of the scope of the present thesis, an important takeaway is that the stress triaxiality is a governing factor in crack initiation which should not be disregarded. Another substantiating reason is that, in the case study following in chapter 4, the nose rib crushing failure mode can occur, which was not present in the study of Liu et al., 2017. Indeed, Liu et al., 2017 mention that there are two possible failure modes of the novel leading edge design they proposed. The first one is related to tensile stretching, while the second one concerns the crushing of the material. Most of the material in the leading edge will fail due to the large displacements during impact towards the front spar, while the crushing mode can be neglected, as it would not significantly decrease the accuracy of the model, according to Liu et al., 2017. Therefore, the only failure mode defined in their study is a critical failure strain, which is an overly simplistic model for the expected failure modes.

Now that the damage initiation has been discussed, the only remaining aspect is its evolution. In Abaqus, it can be either based on the total displacement after damage initiation, or on the energy required for failure, which is also known as fracture energy. The two evolution mechanisms are discussed from a practical point of view in the next chapter, where they are shown to be equivalent. This numerical implementation is necessary as the typical stress-strain relationship is strongly mesh-dependent, according to Dassault Systèmes, 2022e.

Without going into any more numerical details, the plastic behaviour of a metal is the following, from a simplistic point of view: before yielding, the material follows a linear stress-strain relationship, as exhibited in figure 3.3, which is taken from the Abaqus documentation (Dassault Systèmes, 2022e). Afterwards, the material softens, according to the Johnson-Cook law, until damage initiation. The latter occurs when the corresponding equivalent plastic strain, $\bar{\varepsilon}_D^{pl}$, is encountered. The damage then evolves according to an exponential or linear relationship, until the equivalent plastic strain at failure is attained. It is this quantity that is mesh-dependent, and some numerical artifices are made in order to relate it to the fracture energy or plastic displacement at failure, in order to alleviate the dependency aforementioned. In the present thesis, a linear relationship will be employed, because it does not require any additional parameter to describe the variation, contrary to the exponential law depicted in figure 3.3.

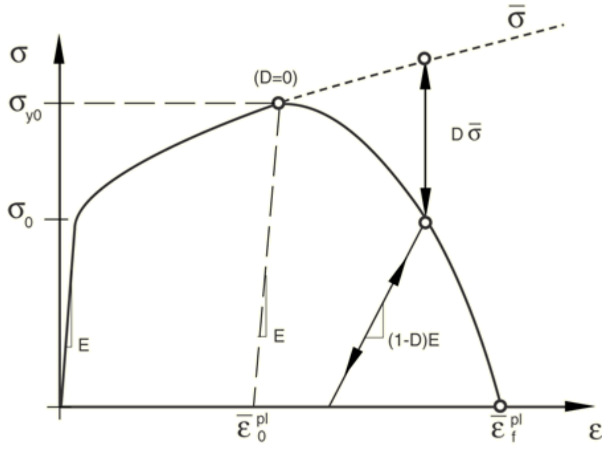


Figure 3.3: Stress-strain curve with progressive damage degradation.

Note: illustration taken from Dassault Systèmes, 2022e, figure 1.

In order to further simplify the problem description, it has been mentioned that no joints would be modelled, therefore no other failure mechanism is needed. As a final note, achieving a balance between factors like computational load, the critical time step of the explicit simulation, and minimum element size can be challenging. According to Georgiadis et al., 2008, the time step must be selected to be smaller than the duration of shock wave propagation through the smallest element in the model, which is typically in the order of microseconds. A time step of such magnitude is necessary to precisely capture all the energies related to the propagation of the shock wave. In order to also give an order of magnitude of the element size in the FE model, Liu et al., 2018 used an element size of approximately 10 mm to simulate an impact with an 8-pound bird.

It will be seen in chapter 4 that, unfortunately, this element size is too large, as it introduces a significant amount of artificial strain energy when nose rib crushing occurs, which further increases the computational time. As for the time step during simulations, Abaqus chooses an appropriate value by itself, and, with the mesh selected after the convergence study in chapter 4, the stable increment size is approximately 0.2 microseconds. It is then clear why many simplifications are needed, such as the decrease of the bird particle number, for instance.

3.2. Design of experiments

Building an initial set of training points is probably an all-or-nothing step in surrogate model-based optimization. As will become apparent later, a good design of experiments is the foundation for both training and building a surrogate model and for dimensionality reduction, where variables of little importance are removed from the optimization process. The actual Bayesian optimization procedure, where infill points are added such that an acquisition function is maximized based on the weight improvement and the probability of feasibility of the design against the constraints, also strongly depends on the initial accuracy.

Gaussian processes are inherently more accurate in the vicinity of a training point: the further a prediction point, the less reliable. Indeed, this will become obvious in section 3.3, when the predictions of surrogate models are going to be analytically described. For now, it is intuitive that for a reasonable accuracy in the whole design space, the training points should be distributed such that there are no regions with sparse sampling. Simpson et al., 2001 recommend the use of space-filling sets in the early design stages, when the trends of the function are not known. There are various designs which address this issue, such as factorial, Placket-Burmann, Central Composite, Box-Behnken designs, or orthogonal and Taguchi arrays.

However, as presented back in the literature review conducted in chapter 2, there is one more aspect needed from a good design of experiments in the proposed methodology, aside from being space-filling. Due to the rather large dimensionality of the problem, a variable ranking procedure is needed. Although the latter has not been discussed yet, it is clear that, after the number of dimensions is reduced, the initial set of training points needs to be projected onto the new design space. If any two of the initial points overlap after projection on the reduced design space, at least one analysis would be a waste of computational effort. Therefore, it can be concluded that a design of experiments where the projections of any two points are not the same is needed. Such a design is the Latin Hypercube, which will be discussed in the present section, and will be used in the proposed methodology.

3.2.1. Latin Hypercube designs

This sampling strategy has been first put forward with respect to deterministic computer codes by McKay et al., 1979, and is also being employed in the surrogate model-based optimization algorithm of Jones et al., 1998, which will be discussed later in this chapter, in section 3.5 regarding Bayesian optimization. It is especially advisable to use Latin Hypercube samplings during the initial phases of the design process, when the trends of the surrogate model are unknown and cannot be prespecified, as outlined by O'Hagan, 2006, who put together a tutorial about Bayesian analysis of responses generated by deterministic computer codes. Moreover, they are good for high-dimensional problems, and, in conjunction with a variable screening procedure, they provide good results for problems with more than 10 variables, as advised by Viana et al., 2021.

Park, 1994, who researched optimal Latin Hypercube designs for computer experiments, offers a comprehensive definition of random Latin Hypercube samplings. Considering a k-dimensional design space, let $F_1, \ldots, F_k$ denote the distribution functions of the independent input variables $X_1, \ldots, X_k$. The j$^{\text{th}}$ coordinate of the i$^{\text{th}}$ sampling point will be symbolized by $x_j^{(i)}$. Supposing n sampling points are sought after, let $P = p_{ij} \in \mathcal{M}_{n \times k}$ represent a matrix such that each column is a random permutation of $\{1, \ldots, n\}$, and let r_{ij} be independent random values distributed in the $[0, 1]$ range. A Latin Hypercube sampling then consists of the $\boldsymbol{x}^{(i)}, \forall i \in \overline{1, n}$ points such that:

$$x_j^{(i)} = F_j^{-1}\left[\frac{1}{n}(p_{ij} - r_{ij})\right] \tag{3.6}$$

To provide a more illustrative representation of a Latin hypercube design, one may consider the design space of $[-1, 1] \times [-1, 1]$, and suppose that 10 sampling points need to be generated. If all F_j are uniform distribution functions and all $r_{ij} = 0.5$, then each of the two one-dimensional axes is split into 10 equal intervals, and the resulting domain can be seen as a 10×10 grid similar to a Sudoku table. Akin to the game aforementioned, 10 points will be placed in some of the cells, such that each row or column contains exactly one point. As Wang, 2003 stated, the values of p_{ij} determine which cells are filled with sampling points, and the r_{ij} term shows where the data point is located within its cell. The resulting sampling is depicted in figure 3.4, which was created using the LHS function of the SMT toolbox (Saves et al., 2023), with the "center" criterion. The latter simply states that each sampling point is found in the center of its corresponding cell, such that all r_{ij} terms are equal to 0.5 in equation 3.6. However, as the number of dimensions and sampling points increase, the number of potential Latin Hypercube designs rises dramatically. As a general rule, for a k-dimensional space, there are $(n!)^{k-1}$ possibilities for a center LHS; for the simple two-dimensional case presented here, $10!$ sampling plans can be constructed.

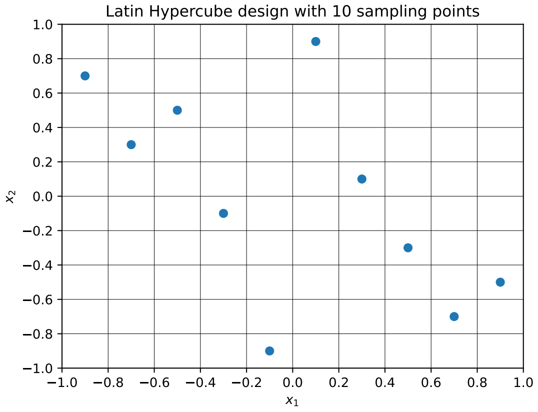


Figure 3.4: Latin Hypercube design with 10 sampling points, distributed in the center of each cell.

Moreover, it can be seen in figure 3.4 that there are regions where the sampling points are rather sparse. There are a number of optimality criteria which aim to provide a more space-filling Latin Hypercube design, such as the maximin, minimax, and entropy criteria. Moreover, a recently developed optimization algorithm for constructing optimal DoEs is the Enhanced Stochastic Evolutionary method put forward by Jin et al., 2005, which is also available in the SMT toolbox. While more optimality criteria have been studied in the paper, such as the entropy or ϕ_p criteria, the ESE algorithm is applied to maximin in the SMT toolbox, which will be the only one discussed in the present section. It should be noted that Trieste has no DoE capability on its own.

The maximin criterion

Johnson et al., 1990 have first put forward the notions of minimax and maximin distance designs. The latter criterion is widely used to construct initial designs for deterministic computer codes, and entails maximizing the minimum distance among any two of the n sampling points:

$$\max_{\boldsymbol{x}} \min_{1 \leq i < j \leq n} d(x_i, x_j) \tag{3.7}$$

In the last equation, the distance is usually taken as the Euclidian distance (Forrester et al., 2008, p. 18). It is reminded that in a k-dimensional space, the Euclidian distance between points $\boldsymbol{x}^{(i)}$ and $\boldsymbol{x}^{(j)}$ is simply $\sqrt{\sum_{p=1}^{k}\left(x_p^{(i)} - x_p^{(j)}\right)^2}$. An example of a maximin Latin Hypercube design obtained with the ESE algorithm is presented in figure 3.5.

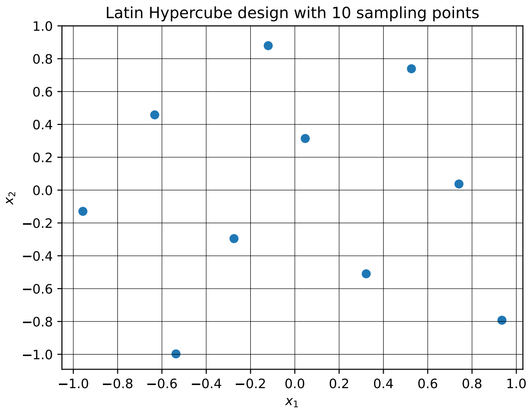
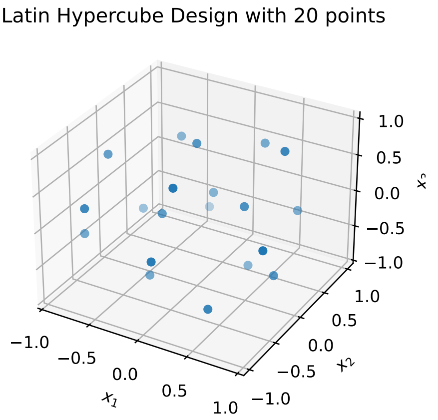
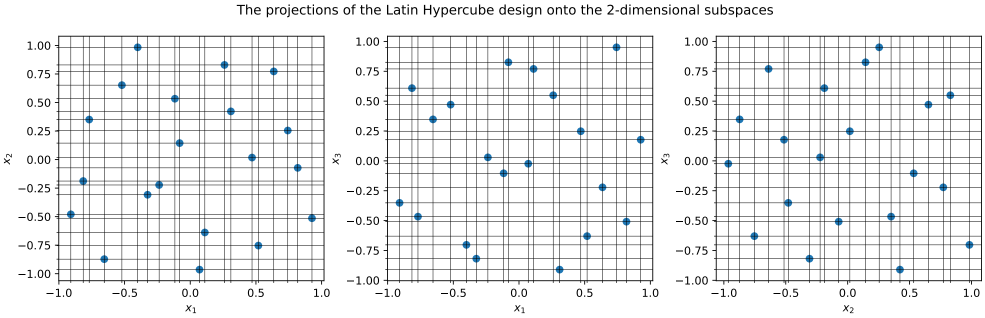


Figure 3.5: Latin Hypercube designs with 10 sampling points, with the ESE criterion.

3.2.2. Projecting the Latin Hypercube sampling points on sub-spaces of reduced dimensions



(a) A three-dimensional Latin Hypercube design with 20 sampling points, using the maximin criterion.



(b) The projections of the 3D sampling onto the 2D subspaces. Note: the grid lines are centered in each data point.

Figure 3.6: A 3D Latin Hypercube design, along with the projections on the lower-dimensional spaces.

A significant benefit of the Latin Hypercube design is that the projections on each axis are uniform, such that no two points would have the same coordinates. To illustrate this idea, suppose that a Latin Hypercube design with 20 samples is built in the $[-1,1]^3$ space, as depicted in figure 3.6a. The resulting points are then projected onto the two-dimensional subspaces, as illustrated in figure 3.6b. The grid lines in the latter have been plotted to also show the projections on each axis.

Fortunately, in this case, there are no points which have dangerously close coordinates. For instance, if the variable x_3 turned out to be of little importance after a variable screening procedure, all the initial sampling points would be projected onto the $x_1 - x_2$ subspace, as shown on the left side of figure 3.6b. The projection rule is rather simple: $(x_1, x_2, x_3) \rightarrow (x_1, x_2)$. In the reduced dimensional space, it is obvious that no two points have been confounded, therefore no training points have been wasted.

3.2.3. Dealing with mixed integer variables

Throughout the whole section, only continuous variables have been covered. As it has been stated that integer and categorical variables should also be easily incorporated in the proposed methodology, an extension of the Latin Hypercube will be presented.

Firstly, consider an integer variable. If there are a large number of possible values compared to the number of requested samples, as in the case of metal sheet thicknesses, for instance, the corresponding interval may be split similarly to the continuous case presented earlier. Supposing that all variables are distributed uniformly, then there would be at least one value in each interval. Therefore, the output of the continuous LHD can simply be rounded off to the nearest available value, and no actual change is required in the construction algorithm. However, a dilemma arises when the integer variable has a much smaller number of possible values than the number of requested sampling points, as this simple solution is no longer applicable. Because categorical variables may also be assigned integer representations, the Latin Hypercube sampling needs to be extended to accommodate such variables.

The Sliced Latin Hypercube Design (SLHD) put forward by Qian, 2012 is, by itself, an LHD, but it possesses a unique property: it can be partitioned into smaller sub-designs which are also LHD, given that the number of training points is a multiple of the requested “slices”. Other appealing aspects include that each partition has a maximum uniformity in every one-dimensional projection, and that, when collapsed over all partitions, i.e. when the sampling is projected onto the reduced design space composed solely of the continuous dimensions, no two points are confounded. Moreover, when compared against other designs on an analytical benchmark problem, the SLHD achieved a much smaller root mean square error than the reunion of t LHD samples, where t is the number of slices.

The SLHD generation algorithm will not be detailed in the present section. Both MATLAB and R packages have been published by the author. Three years later, Shan Ba and Brenneman, 2015 enhance the SLHD algorithm by guaranteeing that the resulting SLHD is optimal with respect to the maximin criterion. This method has also been published as an R package by the author at Ba, 2015.

As R is open-source, this package is proposed to be used in the case of categorical variables, although for continuous variables the SMT package suffices. The only inputs needed are the number of slices, the number of dimensions, and the number of samples in each slice. The design computed by the *maximinSLHD()* function can easily be written in a CSV format which can then be read by Python, by example. Therefore, the design built by the R package can be easily incorporated in the proposed methodology.

To offer an example, consider a design space with 2 continuous dimensions and one categorical variable with 2 levels, which correspond to two slices in the SLHD. Suppose 5 training points are needed in each slice, totalling up to 10 samples. The two continuous dimensions are assumed to be in the $[0, 1)$ range; it is advised, nevertheless, to standardize the variables to this range in both SMT and Trieste. The resulting SLHD is depicted below, in figure 3.7. The two colors represent different categorical levels. Each dimension has been split into 10 intervals, i.e. the total number of samples, and the training points were distributed such that both slices are optimal LHDs, according to the maximin criterion. It can be observed that the samples are centered in each cell of the 2D design space, though. If an integer variable is treated as continuous, then a workaround would be to randomly pick an integer value for each cell with a sampling point. Although this aspect will not be illustrated, the main takeaway is that LHDs can be expanded to account for mixed integer variables as well.

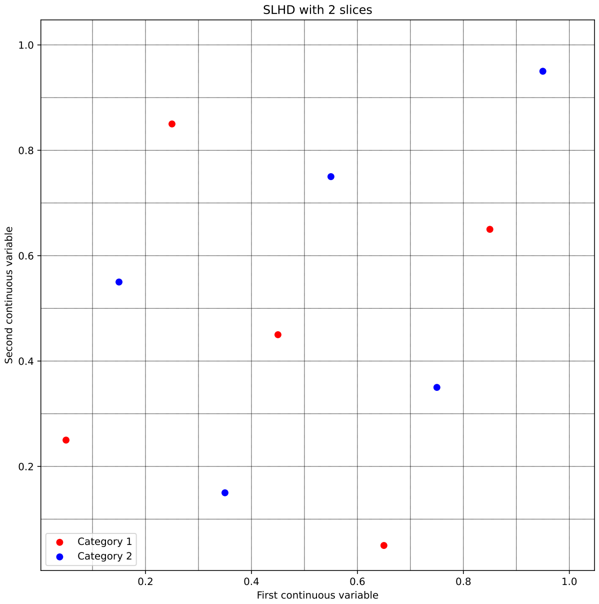


Figure 3.7: A two-dimensional Sliced Latin Hypercube Design with two slices and 5 samples per dimension.

3.3. Surrogate modelling

Suppose that it is now possible to evaluate one of the constraint functions, which will be detailed in the next chapter, section 4.2, with the Abaqus model built with the recommendations mentioned in section 3.1. As foreshadowed in the introductory chapter, in the present thesis, two constraint functions will be investigated: one is going to be related to the maximum penetration depth of the leading edge after impact, while the other one will concern the ability of the front spar to withstand a simplistic load case. Unfortunately, these function evaluations are computationally intensive, and time is often a limiting aspect, especially in preliminary design phases. This is where surrogate models turn out to be tremendously useful.

3.3.1. Necessary characteristics of the surrogate model

As mentioned in chapter 2, Gaussian processes or Kriging surrogates will be used in the present thesis. This choice will be argued once more by pinpointing the necessary characteristics of the metamodel, which concern the following: to be able to capture highly non-linear trends, to interpolate the sampled data, to work without any input regarding gradients, and to offer some form of prediction error estimates.

The need for non-linear modelling capabilities

Due to the various failure modes that can be encountered during a bird strike analysis, such as rib crushing or leading edge penetration, the constraint function may be highly non-linear. As one of the constraints identified in the next chapter concerns the penetration depth, for instance, the surrogate model needs to capture regions where sudden surges are expected, such as when the leading edge skin gets sheared off from a rigid nose rib. Such cases will be illustrated in chapter 6, but they can be anticipated before conducting the optimization study nevertheless. This is especially true for crashworthiness problems, where various phenomena due to high deformation rates take place: joints may fail (although they are not taken into account in the present study), various components may come into contact (e.g. leading edge skin and front spar), materials may start exhibiting elastic-plastic behaviour, and various elements may be deleted as the damage progresses, as outlined by Bisagni et al., 2002.

The need for interpolation

Traditionally, the investigation of a large design space with a limited number of data points was done by polynomial regression. While polynomial response surface (PRS) models are used in most applications (Simpson et al., 2001), they are regression techniques, as opposed to interpolating. Even if it is recognized that PRS is unsuitable anyway for multi-modal and non-linear design spaces unless one uses them only in smaller regions such as in trust-region methods, as outlined in both Forrester and Keane, 2009 and Sóbester et al., 2004, these models are going to be used in this section as an example to convey the idea that interpolating techniques should be used for the type of optimization treated in the report.

While it is true that the output from finite element analyses may seem noisy due to the discretization, for instance, computer simulations are entirely deterministic in the sense that the same results are obtained if the evaluation is carried out twice. Uncertainty quantification will not be treated in the present report. Therefore, there is no random error, and it is completely unjustifiable to smooth the response across data points, as mentioned in Forrester and Keane, 2009. PRS, for example, minimizes the sum of the squares of the residuals. It has been persistently highlighted by Sacks et al., 1989, Jones et al., 1998 and Simpson et al., 2001 that the usual measures of uncertainty derived from this least-squares regression have no obvious meaning, such as the root mean square error, which is a measure of the goodness of the fit. Indeed, any lack of fit simply shows that the set of regression terms is incomplete. Jones, 2001 also emphasizes that the minimum of an ill-fitted quadratic surface may not even lie close to the true minimum.

Other regression techniques include (but are not limited to) the Multivariate Adaptive Regression Splines (MARS) and Support Vector Regression (SVR). MARS was first introduced by Friedman, 1991 and is based on recursive partitioning and spline fitting. SVR, on the other hand, is a weighted sum of basis functions added to a base term, which allows for an error margin to be specified (Forrester et al., 2008, p. 63). This error margin represents a relaxation of the training data that the user is comfortable with; more precisely, the values at the data points are allowed to deviate from the true computed response by this margin. These techniques, however, will not be covered because of the underlying assumption of random errors which is not true for deterministic functions. However, there is a special case of SVR where the margin is null and the approximation is interpolating, called Radial Basis Functions. However, they will not be covered on their own, as Gaussian processes or Kriging metamodels are based on RBFs, and they are widely used in the context of Bayesian optimization, satisfying all necessary criteria.

The unavailability of gradients

Some response surfaces may be more accurately modelled if gradient information is available; an example of such an approximation is gradient-enhanced Kriging. Unfortunately, crashworthiness analyses rarely, if ever, offer gradient information. In the case of bird strike impact simulations, no gradients can be extracted due to changes in contact definitions or deleted elements, for instance.

Sóbester et al., 2004 mention that gradient-enhanced models should not be used if gradients are not readily available. Indeed, if not offered directly by the FEA software, one would need to use, for instance, the finite difference method to compute them. However, to evaluate the gradients at a single point, another k more points would be needed, where k is the total number of variables. This number is obviously prohibitively large, especially since a single analysis takes hours to finish.

The need for prediction error estimates

As repeatedly stated, bird strike simulations are extremely costly in terms of CPU time. Therefore, it is desirable to have as few data points as possible in order to reduce the time spent on evaluating the constraint functions. When the minimization of the objective function is the goal of the optimization process, the employed surrogate needs to possess one distinctive trait: to provide some form of prediction error estimation.

Indeed, as illustrated in section 3.5 of the present chapter, these estimates constitute the foundation of Bayesian optimization, where infill points are sought after in under-explored or under-exploited regions. Relying solely on the surrogate model constructed with an initial set of points is dangerous, no matter how well they are chosen, mainly because regions with local or global minimums may not be accurately captured. Jones, 2001 offers a superb taxonomy of surrogate-based global optimization methods which illustrates why prediction error estimates allow for a balanced exploration and exploitation of a design space, converging more rapidly towards a global optimum. Therefore, as mentioned also in Forrester and Keane, 2009, error estimates for the predictions made by the response surfaces are needed for global optimization, which are available in Gaussian processes.

3.3.2. Gaussian processes

Gaussian processes (GPs) are stochastic or random processes which assume a prior normal distribution of the responses of the modelled functions. Because the design spaces are usually multidimensional, GPs use multivariate normal distributions, i.e. any linear combination of the responses is also normally distributed. The prior assumption that the responses are normally distributed stems from the uncertainty regarding the response of the function at a point which was not previously observed.

Of course, there are some more general Bayesian methods which assume different prior distributions. For instance, the Weibull distribution is widely used in reliability engineering, and would be a sounder choice when modelling the age to failure of various aircraft components. However, in the case of black-box, expensive functions, normal distributions are widely used, and are frequently cited in literature regarding crashworthiness, such as the reviews of Forrester and Keane, 2009 or Simpson et al., 2001. Therefore, the focus of this section will be on normal distributions, and the choice of the most fitting prior is outside the scope of the present report.

Being essentially a Bayesian method due to the assumption of a prior distribution of the responses, a GP thus offers information about the uncertainty of the surrogate model at a certain point. This uncertainty is also closely related to the utility of evaluating the function at a specific location: if the uncertainty is large, then one could try to put a new sample in that region in order to increase the overall accuracy of the model. This is precisely the goal of Bayesian optimization: to make educated, statistical guesses about where one should evaluate an expensive function such that the optimization process is done with as few data points as possible.

Gaussian process regression, which is also called Kriging, was first developed by the South African mining engineer Danie Krige, to aid in mining and geostatical applications (Forrester et al., 2008, p. 50). However, it was Sacks et al., 1989 who introduced the Kriging technique into engineering design, under the name of *Design and Analysis of Computer Experiments*, DACE. Traditionally, the Kriging method used an exponential covariance function of a varying degree, which is a radial basis function on its own, to make the connection aforementioned with RBFs.

Sacks et al., 1989 state that Kriging “treats the deterministic response $y(\boldsymbol{x})$ as a realization of a random function (stochastic process), $Y(\boldsymbol{x})$, that includes a regression model”. If t denotes the number of terms $\beta_i g_i$ in the regression model, the general Kriging model becomes:

$$\hat{f}(\boldsymbol{x}) \doteq Y(\boldsymbol{x}) = \sum_{i=1}^{t} \beta_i g_i(\boldsymbol{x}) + Z(\boldsymbol{x}) \tag{3.8}$$

Note that $\hat{f}$ was used in lieu of y in the last equation, in order to reduce the confusion associated with various y notations which will be employed in this section. In the present report, the rather simple case of the so-called ordinary Kriging will be considered, where the regression term is simply a constant. A note should be made on the term *random*, as it has been repeatedly stated that considering random errors in a deterministic process does not make sense. This assumption simply implies that one is uncertain about the response at an unobserved point and should not be mistaken for any type of uncertainty quantification. Intuitively speaking, the closer an unobserved point is to a known data point, the more confident one can be that the response at the former approaches the response at the latter. While Sacks et al., 1989 do offer the standard derivations of the formulae used in Kriging, a more intuitive approach will be taken here, based on the wonderful introduction presented by Jones, 2001.

Firstly, suppose that the surrogate modelling process starts with an already acquired training set consisting of n data points, $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n)}$. Each data point is a k-dimensional vector, where k is the number of variables considered in the optimization process. Moreover, let $\boldsymbol{y}$ denote the vector of the responses at each of the n points.

Continuing the discussion about the stochastic process, from a statistical point of view, if one considers two points $\boldsymbol{x}^{(i)}$ and $\boldsymbol{x}^{(j)}$, the observational random variables $Y(\boldsymbol{x}^{(i)})$ and $Y(\boldsymbol{x}^{(j)})$ will be more correlated the smaller the distance $\left\|\boldsymbol{x}^{(i)} - \boldsymbol{x}^{(j)}\right\|$ is. Kriging assumes that the random variable $Y(\boldsymbol{x})$ is normally distributed with (unknown) mean and variance which will be the predictor and the error estimate themselves, and that the correlation aforementioned is given by:

$$Corr\left[Y(\boldsymbol{x}^{(i)}), Y(\boldsymbol{x}^{(j)})\right] = \exp\left(-\sum_{l=1}^{k}\theta_l \left|x_l^{(i)} - x_l^{(j)}\right|^{p_l}\right), \theta_l > 0, p_l \in [1,2] \tag{3.9}$$

Equation 3.9 yields a correlation of 1 if $\boldsymbol{x}^{(i)} = \boldsymbol{x}^{(j)}$, and tends to 0 if the points are very far away from each other. Moreover, it can be seen that the bigger θ_l is, the faster the correlation plummets. Therefore, high values of θ_l show that the variable in the l$^{\text{th}}$ direction is very active. Moreover, Jones, 2001 mentions that p_l determines the smoothness in the l$^{\text{th}}$ direction; the bigger it is, the smoother the function. While p_l is typically between 1 and 2 (as presented in Jones et al., 1998 or Forrester and Keane, 2009), a value tending to 0 would theoretically model a function which is completely rough and non-differentiable.

Recall that the observational random variable $Y(\boldsymbol{x})$ was assumed to be normally distributed, and let μ and σ^2 denote the unknown mean and variance, $Y(\boldsymbol{x}) \sim \mathcal{N}(\mu, \sigma^2)$. Intuitively, it means that, if the response at $\boldsymbol{x}$ was sought after, a good guess would be the mean μ, with a 99.7% chance of it lying within three standard deviations, i.e. somewhere between $\mu - 3\sigma$ and $\mu + 3\sigma$. Therefore, all the random variables $Y(\boldsymbol{x}^{(i)}), \forall i \in \overline{1,n}$ will follow this distribution, and the correlations between any two observations can be arranged in the following covariance matrix:

$$Cov(\boldsymbol{Y}) = \sigma^2\boldsymbol{R}, R_{ij} = Corr\left[Y(\boldsymbol{x}^{(i)}), Y(\boldsymbol{x}^{(j)})\right], i,j \in \overline{1,n} \tag{3.10}$$

Denoting by $\mathbf{1}$ the $\{1,\ldots,1\}^T$ column vector of dimension $n \times 1$, equation 3.10 will serve as the covariance matrix of the multivariate Gaussian distribution of the observational random *vector* $\boldsymbol{Y}$, where $\boldsymbol{Y} \sim \mathcal{N}(\mathbf{1}\mu, Cov(\boldsymbol{Y}))$. This multivariate normal distribution will be used to compute the probability of the prediction passing through the responses $\boldsymbol{y}$ as a function of the hyperparameters $\mu, \sigma^2, \theta_l, p_l, l \in \overline{1,n}$. The next step would be to find those parameters which give the best fit of the data, which is quantified by the likelihood of the observational random variable $\boldsymbol{Y}$ yielding the responses $\boldsymbol{y}$ based on the normal distribution aforementioned:

$$\mathcal{L}\left(\mu, \sigma^2, \theta_l, p_l, l \in \overline{1,k}\middle|\boldsymbol{y}\right) = \frac{1}{(2\pi)^{n/2}\sigma^n|\boldsymbol{R}|^{1/2}}\exp\left(-\frac{(\boldsymbol{y}-\mathbf{1}\mu)^T\boldsymbol{R}^{-1}(\boldsymbol{y}-\mathbf{1}\mu)}{2\sigma^2}\right) \tag{3.11}$$

It becomes apparent that it is desired to find the $2k+2$ parameters to maximize this likelihood. In practice, as outlined by both Forrester et al., 2008, p. 55 and Jones, 2001, it is more convenient to work with the natural logarithm likelihood:

$$\ln(\mathcal{L}) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2}\ln(|\boldsymbol{R}|) - \frac{(\boldsymbol{y}-\mathbf{1}\mu)^T\boldsymbol{R}^{-1}(\boldsymbol{y}-\mathbf{1}\mu)}{2\sigma^2} \tag{3.12}$$

Finding the mean μ and variance σ^2 to maximize equation 3.12 is straightforward, thus fully determining the prior normal distribution. Indeed, one only has to derive expression 3.12 with respect to σ^2 and μ, remembering that the correlation matrix $\boldsymbol{R}$ does not depend on these parameters. Moreover, $\boldsymbol{R}$ is symmetric positive definite, a proof of which can be found in Chuong, 2008, which means that $\mathbf{1}^T\boldsymbol{R}^{-1}\boldsymbol{y} = \left(\mathbf{1}^T\boldsymbol{R}^{-1}\boldsymbol{y}\right)^T = \boldsymbol{y}^T\boldsymbol{R}^{-1}\mathbf{1}$, where the first equality is true simply because the quantity is a scalar. The partial derivatives are then:

$$\frac{\partial\ln(\mathcal{L})}{\partial\mu} = \frac{1}{\sigma^2}(\mathbf{1}^T\boldsymbol{R}^{-1}\boldsymbol{y} - \mathbf{1}^T\boldsymbol{R}^{-1}\mathbf{1}\mu) \tag{3.13}$$

$$\frac{\partial\ln(\mathcal{L})}{\partial\sigma^2} = -\frac{n}{2\sigma^2} + \frac{(\boldsymbol{y}-\mathbf{1}\mu)^T\boldsymbol{R}^{-1}(\boldsymbol{y}-\mathbf{1}\mu)}{2\sigma^4} \tag{3.14}$$

The optimal values, represented by the hat symbols, are obtained by setting the derivatives to be null:

$$\hat{\mu} = \frac{\mathbf{1}^T\boldsymbol{R}^{-1}\boldsymbol{y}}{\mathbf{1}^T\boldsymbol{R}^{-1}\mathbf{1}} \tag{3.15}$$

$$\hat{\sigma}^2 = \frac{(\boldsymbol{y} - \mathbf{1}\hat{\mu})^T \boldsymbol{R}^{-1}(\boldsymbol{y} - \mathbf{1}\hat{\mu})}{n} \tag{3.16}$$

Going back to the optimal parameters, these terms are coined as *maximum likelihood estimates* (MLEs) for μ and σ^2 (Forrester et al., 2008, p. 55). As a side note, because $\boldsymbol{R}$ is symmetric positive definite, its inverse is quite easy to compute. Indeed, it can therefore be decomposed into the product of a lower triangular matrix and its transpose via Cholesky decomposition. The inverse of such matrices is cheap to evaluate. Unfortunately, though, it is problematic if data points are too close to each other, leading to numerical issues. Indeed, if $\boldsymbol{x}^{(i)}$ and $\boldsymbol{x}^{(j)}$ are close, the terms R_{ij}, R_{ji} will tend to 1. If many such pairs exist, there may be entire rows or columns which are nearly identical, yielding a null determinant. Generally, though, this should not be the case if a space-filling sample was chosen, thus demonstrating once again its necessity. However, one can imagine that it may become problematic if many infill points were added in a small region, such as in the proximity of a local or global optimum, although this challenge only arises if the ratio of infill to training points is very large, which rarely occurs due to time constraints. Of course, this is the case once $\boldsymbol{R}$ is known; however, the other $2k$ hyperparameters, $\boldsymbol{\theta}$ and $\boldsymbol{p}$, have not been determined yet.

If the optimal values from equations 3.15 and 3.16 are replaced in the natural logarithm likelihood equation 3.12 and the constant terms are removed, the *concentrated ln-likelihood function* (Forrester et al., 2008, p. 55) is obtained:

$$\ln(\mathcal{L}) \approx -\frac{n}{2}\ln(\hat{\sigma}^2) - \frac{1}{2}\ln(|\boldsymbol{R}|) \tag{3.17}$$

Unfortunately, the ln-likelihood equation is not as straightforward anymore to maximize, and is not analytically tractable. Forrester et al., 2008, p. 55, mention that one could use a global search algorithm to directly find the optimum parameters, mainly because equation 3.17 is quite cheap to compute. For instance, in the SMT toolbox (Saves et al., 2023), the terms $\boldsymbol{p}$ are fixed, and $\boldsymbol{\theta}$ can be found by optimization either via COBYLA (Constrained Optimization BY Linear Approximation) or TNC (Truncated Newton). These optimizers will not be detailed further, as they are not the scope of the present report.

Supposing the optimal hyperparameters are now found, consider a new, unobserved point $\boldsymbol{x}^*$. It turns out that the response y^* which is the likeliest to take place is actually the Kriging predictor itself. The proof is going to be reiterated here, following the demonstration of Sacks et al., 1989, who proved that Kriging gives the best linear unbiased prediction at unsampled locations.

If $\boldsymbol{r}$ denotes the correlation vector between the new point $\boldsymbol{x}^*$ and the training points:

$$\boldsymbol{r} = \begin{pmatrix} Corr\left[Y(\boldsymbol{x}^*), Y(\boldsymbol{x}^{(1)})\right] \\ \vdots \\ Corr\left[Y(\boldsymbol{x}^*), Y(\boldsymbol{x}^{(n)})\right] \end{pmatrix} \tag{3.18}$$

$$\tilde{\boldsymbol{y}} = \{y^{(1)}, \dots, y^{(n)}, y^*\}^T \tag{3.19}$$

Taking into account that the correlation of two equal variables is 1, which results in a diagonal consisting only of values of 1, the augmented correlation matrix, which simply takes into account the new unsampled location as well, will be the following:

$$\tilde{\boldsymbol{R}} = \begin{pmatrix} \boldsymbol{R} & \boldsymbol{r} \\ \boldsymbol{r}^T & 1 \end{pmatrix} \tag{3.20}$$

Going back to the equation 3.12, which is true even if the augmented matrices are considered due to the multivariate normal distribution of dimension $n+1$ with the new point, the only term which actually comprises $\boldsymbol{y^*}$ is the one on the far right:

$$-\frac{(\tilde{\boldsymbol{y}} - \mathbf{1}\hat{\mu})^T \tilde{\boldsymbol{R}}^{-1}(\tilde{\boldsymbol{y}} - \mathbf{1}\hat{\mu})}{2\hat{\sigma}^2} \tag{3.21}$$

The equation above can also be written in block-matrix form, which has the advantage of being easy to inverse and of highlighting the original matrix $\boldsymbol{R}$ and the vector $\boldsymbol{r}$:

$$-\frac{\begin{pmatrix}\boldsymbol{y}-\mathbf{1}\hat{\mu}\\ y^*-\hat{\mu}\end{pmatrix}^T\begin{pmatrix}\boldsymbol{R} & \boldsymbol{r}\\ \boldsymbol{r}^T & 1\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{y}-\mathbf{1}\hat{\mu}\\ y^*-\hat{\mu}\end{pmatrix}}{2\hat{\sigma}^2} \tag{3.22}$$

Indeed, the inverse of the augmented correlation matrix is:

$$\tilde{\boldsymbol{R}}^{-1}=\begin{pmatrix}\boldsymbol{R^{-1}}+\boldsymbol{R^{-1}r}(1-\boldsymbol{r^TR^{-1}r})^{-1}\boldsymbol{r^TR^{-1}} & -\boldsymbol{R^{-1}r}(1-\boldsymbol{r^TR^{-1}r})^{-1}\\ -(1-\boldsymbol{r^TR^{-1}r})^{-1}\boldsymbol{r^TR^{-1}} & (1-\boldsymbol{r^TR^{-1}r})^{-1}\end{pmatrix} \tag{3.23}$$

The goal is to find the predictor y^* such that the ln-likelihood is maximized. The next steps are to derive it with respect to y^* and to find when the derivative is null. It should be noted again that, because $\boldsymbol{R}$ is symmetric positive definite, $(\boldsymbol{y}-\mathbf{1}\hat{\mu})^T\boldsymbol{R^{-1}r}=\left[\boldsymbol{r^TR^{-1}}(\boldsymbol{y}-\mathbf{1}\hat{\mu})\right]^T$ is a scalar and has been used for the derivation of the term on the left side in the equation below. The only terms that actually count in equation 3.22 are then:

$$-(y^*-\hat{\mu})\frac{\boldsymbol{r^TR^{-1}}(\boldsymbol{y}-\mathbf{1}\hat{\mu})}{\hat{\sigma}^2(1-\boldsymbol{r^TR^{-1}r})}+\frac{(y^*-\hat{\mu})^2}{2\hat{\sigma}^2(1-\boldsymbol{r^TR^{-1}r})} \tag{3.24}$$

Deriving the equation above with respect to y^* and equating it to 0 yields:

$$\frac{y^*-\hat{\mu}}{\hat{\sigma}^2(1-\boldsymbol{r^TR^{-1}r})}=\frac{\boldsymbol{r^TR^{-1}}(\boldsymbol{y}-\mathbf{1}\hat{\mu})}{\hat{\sigma}^2(1-\boldsymbol{r^TR^{-1}r})} \tag{3.25}$$

Finally, the Kriging predictor becomes, where y^* is replaced by a more general $\hat{f}(\boldsymbol{x})$:

$$\hat{f}(\boldsymbol{x})=\hat{\mu}+\boldsymbol{r}^T\boldsymbol{R}^{-1}(\boldsymbol{y}-\mathbf{1}\hat{\mu}) \tag{3.26}$$

Denoting the correlation function from equation 3.9 by $\psi(\boldsymbol{x}-\boldsymbol{x}^{(i)})=\exp\left(-\sum_{l=1}^{k}\theta_l\left|x_l-x_l^{(i)}\right|^{p_l}\right)$ into expression 3.26, and denoting by b_i the i$^{\text{th}}$ element of the column vector $\boldsymbol{R}^{-1}(\boldsymbol{y}-\mathbf{1}\hat{\mu})$, the Kriging predictor in summation format becomes:

$$\hat{f}(\boldsymbol{x})=\hat{\mu}+\sum_{i=1}^{n}b_i\psi(\boldsymbol{x}-\boldsymbol{x}^{(i)}) \tag{3.27}$$

In the equation above, ψ is known as the Kriging basis function. The Kriging predictor, as a whole, can be seen as a linear combination of basis functions and polynomial terms. The mean $\hat{\mu}$ can also be replaced by a polynomial regression term, which is based on the assumption that the expected mean of the observational random variable $Y(\boldsymbol{x})$ also varies across the design space. For instance, in the SMT toolbox, constant, linear, or quadratic polynomials have been implemented. However, O'Hagan, 2006 mentions that because of the stochastic modelling, the choice of trend function does not affect the prediction much, unless the sampling is sparse. This aspect is also confirmed by Viana et al., 2009, who state that for a high point density, the regression term is not important; conversely, for a low point density, the correlation function loses its significance. This can be explained by the fact that GPs essentially depend on the distance between points; the higher it is, the less their influence, the stochastic term losing its importance.

Moreover, Kriging does interpolate the data; indeed, if one takes $\boldsymbol{x}=\boldsymbol{x}^{(i)}$ in equation 3.26, then $\boldsymbol{r}$ is going to be the i$^{\text{th}}$ column of the matrix $\boldsymbol{R}$. Therefore, if $\boldsymbol{e}_i$ denotes the i$^{\text{th}}$ unit vector, then $\boldsymbol{R}\boldsymbol{e}_i=\boldsymbol{r}\Rightarrow\boldsymbol{R}^{-1}\boldsymbol{r}=\boldsymbol{e}_i$. Moreover, because $\boldsymbol{R}$ is symmetric, then one also has $\boldsymbol{r}^T\boldsymbol{R}^{-1}=\boldsymbol{e}_i^T$. Plugging this back in expression 3.26 simply yields that $\hat{f}(\boldsymbol{x}^{(i)})=y^{(i)}$.

Prediction error estimate for Kriging

Without going into any derivation details, the mean-squared error (MSE) of the predictor derived using the standard stochastic-process approach (which will not be covered here) is, according to Sacks et al., 1989:

$$MSE_{KRG}(\boldsymbol{x}^*) \doteq s^2(\boldsymbol{x}^*) = \hat{\sigma}^2 \left[1 - \boldsymbol{r^T R^{-1} r} + \frac{(1 - \boldsymbol{r^T R^{-1} r})^2}{\boldsymbol{1^T R^{-1} 1}}\right] \tag{3.28}$$

The MSE of Kriging is also called the Kriging variance in literature (den Hertog et al., 2006). In practice, it is usual to consider the root of the expression above, which can be seen as a standard deviation. The Kriging predictor, $\hat{f}$, along with the Kriging variance, s^2, give the *posterior* distribution at the unobserved point $\boldsymbol{x}^*$: $f(\boldsymbol{x}^*) \sim \mathcal{N}(\hat{f}(\boldsymbol{x}^*), s^2(\boldsymbol{x}^*))$. It should be mentioned that den Hertog et al., 2006 state that this classical variance formula is wrong and that it is actually an underestimator, and propose a new variance prediction based on bootstrapping. However, the present variance is widely used in Kriging metamodels, and the proposed one will not be covered.

Other correlation functions

The exponential correlation function with a constant power of 1 or 2 is the most popular choice for Gaussian processes, i.e. $p_l = 1, \forall l \in \overline{1,k}$ in equation 3.9. However, many other types of correlation functions have been developed to work with the Kriging surrogate model instead of the one found in equation 3.9. For instance, in both the SMT toolbox (Saves et al., 2023) and the Trieste package (The Trieste Contributors, 2020a), the Matérn functions have been implemented, where $\theta_l > 0, l \in \overline{1,k}, i, j \in \overline{1,n}$, the most popular of them being:

1. Matérn 3/2 correlation function

$$\prod_{l=1}^{k} \left(1 + \sqrt{3}\theta_l |x_l^{(i)} - x_l^{(j)}|\right) \exp\left(-\sqrt{3}\theta_l |x_l^{(i)} - x_l^{(j)}|\right) \tag{3.29}$$

2. Matérn 5/2 correlation function

$$\prod_{l=1}^{k} \left(1 + \sqrt{5}\theta_l |x_l^{(i)} - x_l^{(j)}| + \frac{5}{3}\theta_l^2 \left(x_l^{(i)} - x_l^{(j)}\right)^2\right) \exp\left(-\sqrt{5}\theta_l |x_l^{(i)} - x_l^{(j)}|\right) \tag{3.30}$$

Because the training of the surrogate model is much faster than the actual evaluation of the responses at the training points, many different surrogate models can be easily compared to each other. However, in the present report, only the exponential correlation function with a constant power of 1 or 2 will be employed, due to its easier implementation in the variance-based variable screening procedure which will be exposed in the next section. Moreover, although it has been mentioned that the mean function can also vary, in the present study, a constant mean function will be employed for simplification purposes, like the one used in the derivation in this section.

Mixed integer Kriging

A special type of Kriging which should be discussed is the categorical Kriging. Although the design variables of choice in the case study to follow in chapter 4 will be treated as discrete and a continuous ordinary Kriging model will be employed, where the infill points suggested by the optimization procedure will simply be rounded off, the proposed methodology will need to be improved to accommodate categorical or integer variables, such as the material of various components or the number of fasteners. Fortunately, Garrido-Merchán and Hernández-Lobato, 2020 propose a way to overcome this difficulty in Bayesian optimization with GPs.

Integer variables are easily accommodated by simply rounding them to the closest value. Categorical variables, on the other hand, are treated via the one-hot encoding procedure, which replaces them with as many variables as categories; the following example is given: the set {red, green, blue} is replaced by {(1,0,0),(0,1,0),(0,0,1)}. The latter can then be viewed as three variables ranging from 0 to 1; during the optimization/surrogate modelling procedure, the regions corresponding to the largest value being, for instance, the 3rd variable, blue, will lead to the configuration (0,0,1). A new covariance function based on this transformation is proposed, which allows for the underlying GPs to yield a constant objective function in the regions of the continuous input space. It is acknowledged that this approach leads to an increase in the number of dimensions, but, as all other qualities of the Kriging model are preserved, such as prediction error estimates, the inherent computational burden of the method is justified.

Another valid method would be to assume there is some sort of correlation between the categorical variables and to model the responses using Cokriging, which, yet again, would lead to a larger number of smaller-dimensional metamodels.

Although no categorical variables will be included in the case study presented in chapter 4, it is important to stress the fact that the present methodology can accommodate them. In the next section, which details the proposed variable screening procedure, the treatment of categorical variables will be touched upon, based on the Kriging variant developed by Garrido-Merchán and Hernández-Lobato, 2020. As for the Bayesian optimization procedure itself, no further adjustments are needed.

3.4. Variable ranking

The optimization of a wing structure is a complex task, as it typically consists of numerous components, including spars, stringers, fasteners, and skin panels, each with varying geometrical properties. Given the numerous factors to take into account, achieving an optimal design requires a careful analysis and consideration of the contribution of each component to the overall performance of the wing. Especially in the context of expensive, black-box functions, where the contribution of some variables cannot be directly assessed solely via engineering judgment, it becomes increasingly apparent that the number of variables may become prohibitive. Indeed, now that the design of experiments and surrogate modelling topics have been covered, it is obvious that the number of training points dictates the accuracy of the prediction, while also depending on the number of dimensions.

In the field of engineering design optimization, the challenge of dealing with an overwhelming number of design variables has been referred to as the “curse of dimensionality”. Koch et al., 1999 specifically highlight the difficulty of constructing accurate response surface models when the number of design variables is prohibitively large due to the computational cost of obtaining sufficient training data. In addition, identifying robust regions of feasible solutions is particularly challenging during the preliminary stages of design, where reducing computational costs is imperative. Thus, managing the curse of dimensionality has become a fundamental aspect of research regarding engineering design optimization.

To mitigate the computational expenses linked to a large number of design factors, Koch et al., 1999 and Viana et al., 2021 suggest to conduct variable ranking or screening procedures if there are more than 10 design variables. Koch et al., 1999 mentioned this threshold because the focus of their study was on a second-order polynomial response surface model, where the number of training points needs to be at least $(k+1)(k+2)/2$ in order to determine the polynomial coefficients, with k being the number of variables. For a 10-dimensional problem, 66 sampling points are needed to fit the PRS model, which could be a prohibitive number. On the other hand, Viana et al., 2021 made this practical recommendation based on the authors’ experience. The ordinary Kriging surrogate model, which was presented in the last section, does not have any stringent threshold regarding the number of training points. For instance, Welch et al., 1992 achieved a reasonable accuracy on a 20-variable problem with only 60 sampling points. However, the necessary size of the sampling set may increase dramatically if the problem is highly non-linear, therefore the limit of 10 variables should only be taken as an approximate figure to convey the need for variable screening.

Some design variables have a more intuitive impact than others, and can be eliminated from the optimization process altogether. For instance, the rear spar of a wing probably has a negligible contribution to the penetration depth of a bird impact. However, identifying the specific skin panels which contribute to damage absorption, for instance, is a more nuanced issue. For instance, determining whether only the leading edge skin or also the first skin panel behind the front spar should be included is not immediately apparent. Incorrectly identifying irrelevant variables as important may increase computational time but is unlikely to cause significant harm, whereas erroneously neglecting important variables could result in suboptimal designs. Moreover, the interaction of many design variables might also not be immediately apparent, requiring a systematic way to assess the importance of variables and their joint effects.

This is precisely what the analysis of variance, ANOVA, entails. It captures not only the main effects of the variables, but also the higher-order interactions between them. For instance, Craig et al., 2005 used ANOVA on a linear response surface in the context of successive response surface modelling, SRSM. While the latter has not been covered in the present report, it entails building surrogate models on subdomains of the design space, which are chosen by panning and zooming onto candidate optimal points in each iteration. By identifying the most important factors in an automotive structural optimization against frontal crash, Craig et al., 2005 found a reduction in computational time of more than 30 per cent when compared against the same problem with no variable screening performed.

However, it was established in the last section that Gaussian process-based surrogate models are sought after. Therefore, the analysis of variance on the ordinary Kriging metamodels will be presented, following the steps provided by Schonlau and Welch, 2006. The methodology they propose has also been demonstrated on the "Wonderland" model, which is a 41-variable problem that predicts the global human development index based on various demographic and economic inputs. While the Universal Kriging has been employed in the study of Schonlau and Welch, 2006, the derivation exposed in the present chapter will only cover the ordinary Kriging which was detailed in the last section.

It should be noted that the same methodology has been presented in Saltelli et al., 2007, although not applied specifically to the Kriging metamodel. The Russian mathematician Sobol' is cited in the previous book to have been the first to generalize the so-called "high-dimensional model representation", where the function to be modelled is decomposed into terms of increasing dimensions which are related to the main or joint effects of specific variables and shows that, under certain assumptions, these terms can be unequivocally computed using the conditional expectations of the model output. The same method is also coined as Global Sensitivity Analysis, GSA. The variance of the effects aforementioned is then computed, as to allow a quantification of the contribution of each individual variable.

The advantage, though, is that the contribution of a set of variables is also possible to be determined. Therefore, not only are pairs allowed to have their joint effect approximated, but also sets of an even higher number of dimensions. Saltelli and Sobol', 1995 is the paper in question, but, unfortunately, the author of the present thesis was not successful in finding an English translation. However, the theory presented in Saltelli et al., 2007 does cover the aspects relevant to the variance-based variable screening procedure employing surrogate models, which is the scope of the present section, where the steps provided by Schonlau and Welch, 2006 are going to be covered. Three reasons are offered for this choice. Firstly, the methodology has been applied directly on the Kriging surrogate model, and has been validated against a high-dimensional 41-variable computer model. Secondly, the proposed approach does not entail the scaling of the variable inputs on the unit hypercube, as presented in Saltelli et al., 2007, although normalization is recommended by Forrester et al., 2008. Last but not least, the methodology can be easily extended to take into account uncertainties in the inputs. While it is acknowledged that uncertainty quantification is not the focus of future research, it is desirable to have a data model which allows naturally such improvements with minimum alterations. The present section will start with the derivation of the variance-based screening method applied to the ordinary Kriging metamodel with a squared exponential kernel, assuming that all variables are continuous. Afterwards, a modification to the approach will be proposed in order to accommodate integer and categorical variables.

3.4.1. Computing main and interaction effects of variables in a high-dimensional problem

Let $\hat{f}(x)$ denote the ordinary Kriging predictor derived the last section, which is defined on a continuous design space. It is assumed that the predictor has already been trained on a Latin Hypercube initial sampling, as suggested by Schonlau and Welch, 2006, which is an appropriate design of experiments for a non-linear, high-dimensional function. Recall that one major advantage of the Latin Hypercube Sampling (LHS) is that, after the variable ranking procedure, all the initial sampling points can be recycled and reused when retraining the surrogate model in the new design space, which is of lower dimension. Indeed, it has been mentioned in section 3.2 that in LHD the projections of any two points on any axis do not coincide; therefore, when projecting the initial sampling onto the reduced design space, no two data points will be confounded. During the literature review presented in chapter 2, it was hypothesized that, because the training points are more densely packed into the new design space, the accuracy of the metamodel will increase in the reduced design space. However, it will be seen in chapter 6 that this assumption does not hold for both of the constraints proposed for the case study in chapter 4.

Continuing with the ordinary Kriging predictor, the squared exponential kernel will be employed for simplicity, with $p_l = 2, \forall p \in \overline{1,k}$. Moreover, in the present derivation, it is assumed that the function to be modelled is continuous. A brief derivation of the Ornstein-Uhlenbeck kernel, i.e. $p_l = 1, \forall p \in \overline{1,k}$, can be found in appendix A. As mentioned earlier in section 3.3, the power of the distance function in the exponential function, p_l, is directly proportional to the smoothenss of the function. Therefore, it may well be the case that some constraint functions might exhibit a much larger degree of discontinuity, and that the squared exponential kernel may not model them correctly. Considering yet again an initial training set of n data points, the ordinary Kriging predictor will be reiterated here for convenience:

$$\hat{f}(\boldsymbol{x}) = \hat{\mu} + \sum_{i=1}^{n} b_i \prod_{j=1}^{k} \text{exp}\left(-\theta_j \left(x_j - x_j^{(i)}\right)^2\right) \tag{3.31}$$

In equation 3.31, $\boldsymbol{x}^{(i)}$ is the i^{th} of the n training points, x_j is the j^{th} coordinate of $\boldsymbol{x}$, and $\hat{\mu}, b_i, \forall i \in \overline{1,n}, \theta_j, \forall j \in \overline{1,k}$ are all hyperparameters which are computed as outlined in section 3.3. Keeping the notations of Schonlau and Welch, 2006, let χ denote the whole domain; it is then assumed that χ can be decomposed by the direct product of the one-dimensional domains of the individual variables:

$$\chi = \otimes_{j=1}^{k} \chi_j \tag{3.32}$$

One more supposition is made in Schonlau and Welch, 2006: that integrals, which will be omnipresent throughout this section, are computed with respect to a weight function $\boldsymbol{w}$ which can be also decomposed in the product of one-dimensional weights. These weight functions are nothing more than probability density functions, which relate the value of a specific variable to the probability of that value being encountered. Because it is assumed that no uncertainties are present with respect to the variables, a uniform distribution is chosen:

$$\begin{gathered}\boldsymbol{w}(\boldsymbol{x}) = \prod_{j=1}^{k} w_j(x_j) = \prod_{j=1}^{k} \frac{1}{x_j^u - x_j^l} \\ \text{where } x_j^u = \text{max}(\chi_j), x_j^l = \text{min}(\chi_j), \forall j \in \overline{1,k}\end{gathered} \tag{3.33}$$

These weights do not appear in Saltelli et al., 2007. However, because in the book aforementioned it is assumed that the variables range from 0 to 1 (i.e. the input domain is the unit hypercube), the uniform weight distribution function can be seen simply as a mapping from the $[0,1]$ interval to the actual range of the variable.

If $\boldsymbol{e}$ is a subset of $\overline{1,k}$, the notation $\boldsymbol{x_e}$ will be used to define the variables whose effect should be studied, while $\boldsymbol{x_{-e}}$ will denote the variables in $\{1,\ldots,k\} - \boldsymbol{e}$; thus, the input vector can be split in $\boldsymbol{x} = (\boldsymbol{x_e}, \boldsymbol{x_{-e}})$. For instance, if the joint effect of variables x_1 and x_2 is to be studied in a three-dimensional problem, $\boldsymbol{e} = \{1,2\}$. The notation $\boldsymbol{x_e}$ is essentially the canonical projection of $\boldsymbol{x} \in \chi$ onto the subspace $\otimes_{j \in \boldsymbol{e}} \chi_j$. Note that, for the study of main effects, $\boldsymbol{e}$ will only contain one element.

Now that the notations and assumptions have been discussed, the next focus is the definition of the effect aforementioned. Firstly, the so-called marginal effect $\hat{f}_{\boldsymbol{e}}$ of $\boldsymbol{x_e}$ should be computed by integrating out all other variables, which essentially yields a mean value given only some dimensions of the whole set. From a statistical point of view, this effect would be the expected value of the predictor $\hat{f}$, given variables $x_i, i \in \boldsymbol{e}$:

$$\hat{f}_{\boldsymbol{e}}(\boldsymbol{x_e}) = E[\hat{f}|X_i, i \in \boldsymbol{e}] = \int_{\underset{j \notin \boldsymbol{e}}{\otimes} \chi_j} \hat{f}(\boldsymbol{x_e}, \boldsymbol{x_{-e}}) \prod_{j \notin \boldsymbol{e}} w_j(x_j) dx_j \tag{3.34}$$

Based on the assumptions aforementioned and the formula of the Kriging predictor in equation 3.31, the marginal effect becomes:

$$\hat{f}_{\boldsymbol{e}}(\boldsymbol{x_e}) = \hat{\mu} + \sum_{i=1}^{n} b_i \prod_{j \notin \boldsymbol{e}} \frac{1}{x_j^u - x_j^l} \int_{\chi_j} \text{exp}\left(-\theta_j \left(x_j - x_j^{(i)}\right)^2\right) dx_j \prod_{j \in \boldsymbol{e}} \text{exp}\left(-\theta_j \left(x_j - x_j^{(i)}\right)^2\right) \tag{3.35}$$

The integral can also be computed by using the error function erf$(z) = \frac{2}{\sqrt{\pi}} \int_0^z \text{exp}(-t^2)dt$. In practice, during the implementation in Python, the error function computation is much faster than the direct numerical integration using, for instance, SciPy's *quad* function (The SciPy community, 2023). Splitting the integral domain $\chi_j = [x_j^l, x_j^{(i)}] \cup [x_j^{(i)}, x_j^u]$ and performing the substitutions $t = \sqrt{\theta_j}(x_j^{(i)} - x_j)$ and $t = \sqrt{\theta_j}(x_j - x_j^{(i)})$, the integral can be written as:

$$\int_{\chi_j} \text{exp}\left(-\theta_j \left(x_j - x_j^{(i)}\right)^2\right) dx_j = \frac{\sqrt{\pi}}{2\sqrt{\theta_j}} \left\{\text{erf}\left[\sqrt{\theta_j}(x_j^{(i)} - xj^l)\right] + \text{erf}\left[\sqrt{\theta_j}(x_j^u - x_j^{(i)})\right]\right\} \tag{3.36}$$

The main idea behind ANOVA is to compute the variance of the effect with respect to a mean value, which will be taken as the expected value of the ordinary Kriging predictor:

$$\begin{aligned} \mu_0 &= E[\hat{f}] \\ &= \int_{\chi} \hat{f}(\boldsymbol{x})\boldsymbol{w}(\boldsymbol{x})d\boldsymbol{x} \\ &= \hat{\mu} + \sum_{i=1}^{n} b_i \prod_{j=1}^{k} \frac{1}{x_j^u - x_j^l} \int_{\chi_j} \text{exp}\left(-\theta_j \left(x_j - x_j^{(i)}\right)^2\right) dx_j \end{aligned} \tag{3.37}$$

The marginal effect defined in equation 3.35 is not yet corrected with respect to this average. In the following, the discussion will be limited to the study of the effects of at most two variables; only the second-order interactions will be captured. The corrected effects, which can essentially be seen as how much each set of variables makes the marginal effect deviate from the overall mean, are defined as:

$$\mu_j(x_j) = \hat{f}_j(x_j) - \mu_0, \forall j \in \overline{1,k} \tag{3.38a}$$

$$\mu_{jl}(x_j, x_l) = \hat{f}_{jl}(x_j, x_l) - \mu_j(x_j) - \mu_l(x_l) - \mu_0, \forall j, l \in \overline{1,k} \tag{3.38b}$$

Equation 3.38a represents the corrected main effect of variable x_j, while equation 3.38b shows the corrected joint effect of variables x_j and x_l. Finally, the ANOVA decomposition states that the overall variance of the predictor is the sum of the variances of the main and joint effects of all possible sets of variables:

$$\begin{aligned} \int_{\chi} [\hat{f}(\boldsymbol{x}) - \mu_0]^2 \boldsymbol{w}(\boldsymbol{x})d\boldsymbol{x} = Var[\hat{f}] \\ &= \sum_{j=1}^{k} \int_{\chi_j} \mu_j^2(x_j)w_j(x_j)dx_j \\ &+ \sum_{j=1}^{k-1} \sum_{l=j+1}^{k} \int_{\chi_l} \int_{\chi_j} \mu_{jl}^2 w_j(x_j)w_l(x_l)dx_j x_l \\ &+ \ldots + \int_{\chi} \mu_{1\ldots k}^2(x_1, \ldots, x_k) \prod_{j=1}^{k} w_j(x_j)dx_j \\ &= \sum_{\boldsymbol{e} \subseteq \{1,2,\ldots,k\}} Var[\hat{f}|\boldsymbol{X_e}] \end{aligned} \tag{3.39}$$

The contribution of the main or joint effects to the total variance thus represents the importance of the variable(s). The latter is quantified by expressing the terms on the right side as percentages of the total variance. Thus, the importance of variable x_j will be $Var[\hat{f}|X_j]/Var[\hat{f}] \cdot 100\% = \int_{\chi_j} \mu_j^2(x_j)w_j(x_j)dx_j / \int_{\chi} [\hat{f}(\boldsymbol{x}) - \mu_0]^2 \boldsymbol{w}(\boldsymbol{x})d\boldsymbol{x} \cdot 100\%$; the importance of the interaction between two variables is expressed similarly. Moreover, the introduction of the weights in the integral help alleviate the differences in the ranges of the variables: if they were not present, the variables in the direction of the largest uni-dimensional domain would artificially have a larger effect. This methodology also allows for an extension to take into account uncertainties as well.

The computation of the left term in equation 3.39, which represents the total variance of the predictor, should also be briefly presented:

$$\begin{aligned}
\int_{\chi}[\hat{f}(\boldsymbol{x})-\mu_0]^2\boldsymbol{w}(\boldsymbol{x})d\boldsymbol{x} &= \int_{\chi}\left[\hat{f}^2(\boldsymbol{x})-2\mu_0\hat{f}(\boldsymbol{x})+\mu_0^2\right]\boldsymbol{w}(\boldsymbol{x})d\boldsymbol{x} \\
&= \int_{\chi}\left[\hat{\mu}+\hat{f}(\boldsymbol{x})-\hat{\mu}\right]^2\boldsymbol{w}(\boldsymbol{x})d\boldsymbol{x}-2\mu_0^2+\mu_0^2 \\
&= \hat{\mu}^2-\mu_0^2+2\hat{\mu}\left(\int_{\chi}\hat{f}(\boldsymbol{x})\boldsymbol{w}(\boldsymbol{x})d\boldsymbol{x}-\hat{\mu}\right)+ \\
&+\int_{\chi}\left[\sum_{i=1}^{n}b_i\prod_{j=1}^{k}\text{exp}\left(-\theta_j\left(x_j-x_j^{(i)}\right)^2\right)\right]^2\boldsymbol{w}(\boldsymbol{x})d\boldsymbol{x} \\
&= \hat{\mu}^2-\mu_0^2+2\hat{\mu}\mu_0-2\hat{\mu}^2+ \\
&+\sum_{i,p=1}^{n}b_ib_p\prod_{j=1}^{k}\frac{1}{x_j^u-x_j^l}\cdot\int_{\chi_j}\text{exp}\left(-\theta_j\left[\left(x_j-x_j^{(i)}\right)^2+\left(x_j-x_j^{(p)}\right)^2\right]\right)dx_j \\
&= -(\hat{\mu}-\mu_0)^2+\sum_{i=1}^{n}b_i^2\prod_{j=1}^{k}\frac{1}{x_j^u-x_j^l}\int_{\chi_j}\text{exp}\left(-2\theta_j\left(x_j-x_j^{(i)}\right)^2\right)dx_j+ \\
&+2\sum_{i=1}^{n-1}\sum_{p=i+1}^{n}b_ib_p\prod_{j=1}^{k}\frac{1}{x_j^u-x_j^l}\int_{\chi_j}\text{exp}\left(-\theta_j\left[\left(x_j-x_j^{(i)}\right)^2+\left(x_j-x_j^{(p)}\right)^2\right]\right)dx_j
\end{aligned} \tag{3.40}$$

In the last equation, the first integral can again be easily computed with the error function, which is more stable from a numerical point of view, as presented earlier. Moreover, the second integral can also be expressed in a similar manner. Firstly, the term inside the exponential function will be rewritten in the form:

$$\begin{aligned}
\left(x_j-x_j^{(i)}\right)^2+\left(x_j-x_j^{(p)}\right)^2 &= 2x_j^2-2x_j(x_j^{(i)}+x_j^{(p)})+x_j^{(i)2}+x_j^{(p)2} \\
&= 2\left(x_j-\frac{x_j^{(i)}+x_j^{(p)}}{2}\right)^2+\frac{1}{2}\left(x_j^{(i)}-x_j^{(p)}\right)^2
\end{aligned} \tag{3.41}$$

The integral then becomes:

$$\begin{aligned}
&\int_{\chi_j}\text{exp}\left(-\theta_j\left[\left(x_j-x_j^{(i)}\right)^2+\left(x_j-x_j^{(p)}\right)^2\right]\right)dx_j = \text{exp}\left[-\frac{\theta_j}{2}\left(x_j^{(i)}-x_j^{(p)}\right)^2\right]\cdot \\
&\cdot\int_{x_j^l}^{x_j^u}\text{exp}\left[-2\theta_j\left(x_j-\frac{x_j^{(i)}+x_j^{(p)}}{2}\right)^2\right]dx_j = \frac{\sqrt{\pi}\,\text{exp}\left[-\frac{\theta_j}{2}\left(x_j^{(i)}-x_j^{(p)}\right)^2\right]}{2\sqrt{2\theta_j}}\cdot \\
&\cdot\left\{\text{erf}\left[\sqrt{2\theta_j}\left(\frac{x_j^{(i)}+x_j^{(p)}}{2}-xj^{l}\right)\right]+\text{erf}\left[\sqrt{2\theta_j}\left(x_j^u-\frac{x_j^{(i)}+x_j^{(p)}}{2}\right)\right]\right\}
\end{aligned} \tag{3.42}$$

The last integral in the equation above was computed similarly to the one present in the marginal effect. The implementation in Python is also discussed because it is acknowledged that numerical integrations should be avoided as much as possible because they are based on the computation at a predefined finite number of points (e.g. Gaussian quadrature), which may not be always suitable for the problem at hand. Therefore, if there is no other alternative, when using SciPy's quad function, for instance, it is advised to check the predicted errors of the integrals, which are also readily provided.

The ANOVA decomposition, though, still relies on training the surrogate model with an initial sampling. In the context of expensive black-box functions such as the constraints employed in the present report, validation points are expensive to compute, therefore the model should be checked against cross-validation metrics such as the predicted residual error sum of squares, PRESS, or against a validation set to compute the root mean square error, RMSE, and more points should be added until a suitable accuracy is obtained.

It was mentioned in section 3.3 that the $\boldsymbol{\theta}$ hyperparameters in Kriging can also be used to assess the importance of a variable. Indeed, they appear in the basis functions in equation 3.31: the biggest θ_j is, the more rapidly the exponential function decays in the j^{th} direction, thus having a stronger influence on the function to be modelled. However, it does not provide any sort of information on how the variables interact with each other. It should be noted, though, that if variable x_j interacts strongly with variable x_i, the latter should not be discarded, even if its main effect may show that it is not important, a consequence which may not be apparent from the Kriging hyperparameter only.

It should be noted that the variable ranking will be done only for the bird strike constraint functions, whereas the mass of the wing is computed in the whole design space. Hence, the constraint and objective functions are defined in different domains. If a suitable threshold for the variance contribution is chosen, the low variance of the function in a direction yields negligible changes of the function with respect to the corresponding variable. Therefore, the surrogate model can be extended back into the whole design space such that the responses are constant along the directions of the insignificant variables.

3.4.2. Proposed modification for mixed-integer variables

As mentioned earlier in section 3.3, Garrido-Merchán and Hernández-Lobato, 2020 propose an extension to Kriging to accommodate integer and categorical variables, the method having been already implemented in the SMT toolbox. Its treatment of integer variables does not pose any sort of problem to the proposed variable ranking based on ANOVA, as the response is simply rounded off to the nearest value, the underlying surrogate being continuous. Therefore, it is proposed that, for integer variables, the variable screening procedure remains the same, applied to the underlying continuous Kriging. Of course, as mentioned in the last section, a different covariance function, which will be denoted by ψ, is employed in the proposed variant, therefore the implementation in Python will be slightly different.

What is more problematic, though, is the treatment of categorical variables. Garrido-Merchán and Hernández-Lobato, 2020 use the so-called "one-hot encoding", which assigns a continuous dimension for each level of a given categorical variable. For notation purposes, let $\mathcal{M}_{cat}$ denote the set of all one-dimensional categorical domains, and $\mathcal{M}_{cont}$ the set of all continuous/integer domains. Then for each $\phi \in \mathcal{M}_{cat}$, there are L_ϕ categorical levels such that $\phi = \otimes_{L=1}^{L_\phi} \phi_L$, i.e. the categorical domain ϕ can be decomposed as the direct product of L_ϕ continuous domains. Splitting the whole design space χ in the direct product of the categorical and the continuous/integer domains, $\chi_{cat} \otimes \chi_{cont}$, the following decomposition is assumed to take place:

$$\chi = \chi_{cat} \otimes \chi_{cont} = (\otimes_{\phi \in \mathcal{M}_{cat}} \phi) \otimes (\otimes_{\xi \in \mathcal{M}_{cont}} \xi) = \left[\otimes_{\phi \in \mathcal{M}_{cat}} \left(\otimes_{L=1}^{L_\phi} \phi_L\right)\right] \otimes (\otimes_{\xi \in \mathcal{M}_{cont}} \xi) \tag{3.43}$$

Equation 3.43 simply states that, with the one-hot encoding, the design space is split into the direct product of continuous one-dimensional domains, where a relationship exists between a categorical one-dimensional domain and the corresponding continuous domains, $\phi \to \{\phi_1, \ldots, \phi_{L_\phi}\}$. Let x_ϕ denote the variable corresponding to the ϕ domain, regardless of the latter being continuous or categorical. The ordinary Kriging predictor, with the new covariance function, can then be written as:

$$\hat{f}(\boldsymbol{x}) = \hat{\mu} + \sum_{i=1}^{n} b_i \left\{ \prod_{\xi \in \mathcal{M}_{cont}} \psi\left(x_\xi, x_\xi^{(i)}\right) \right\} \cdot \left\{ \prod_{\phi \in \mathcal{M}_{cat}} \left[\prod_{\phi_L, L \in \overline{1, L_\phi}} \psi\left(x_{\phi_L}, x_{\phi_L}^{(i)}\right) \right] \right\} \tag{3.44}$$

As the variance calculations entail the computation of various integrals, instead of computing the integral on the categorical χ_C domain, an L-dimensional integral or an L-fold multiplication should be tackled, as shown in the marginal effect 3.35, the overall average 3.37, and the total variance of the predictor 3.40. It is proposed to compute these integrals in the following manner:

$$\int_{\chi_j} \psi\left(x_j, x_j^{(i)}\right) dx_j = \begin{cases} \int_{x_j^l}^{x_j^u} \psi\left(x_j, x_j^{(i)}\right) dx_j, \text{if } \chi_j \in \mathcal{M}_{cont} \\ \int_0^1 \cdots \int_0^1 \prod_{\alpha \in \overline{1,L}} \psi\left(x_{j_\alpha}, x_{j_\alpha}^{(i)}\right) dx_{j_1} \dots dx_{j_L}, \text{if } \chi_j \in \mathcal{M}_{cat} \end{cases} \tag{3.45}$$

In the last equation, L denotes the number of possible levels corresponding to the categorical domain χ_j. The L-fold multiplication is carried out similarly to equation 3.44. It should be noted, though, that this proposed solution has not been implemented yet, as the case study presented in chapter 4 will not tackle any categorical variables.

Moreover, it is acknowledged that a relatively good initial fit of the model is needed for this method, at least from a qualitative point of view (i.e. the right trends are predicted), such that the proper significant variables can be found. The assessment of the accuracy of the initial surrogate model will be discussed in further detail in chapter 5. However, the variance-based variable screening method has been selected as it is able to also quantify the significance and interaction effects of the variables, being more valuable overall from an educational point of view. Nevertheless, as mentioned in the literature review in chapter 2, the PLS method is also a suitable candidate, particularly when used in conjunction with the Kriging model variant proposed by Bouhlel et al., 2016.

3.5. Bayesian optimization

When the goal is to find an optimal design constrained by expensive, black-box functions, there is usually not enough time to generate as many training points as needed for a globally accurate surrogate model. Therefore, the surrogate modelling process usually starts with a reduced number of sampling points, adding subsequent training points based on a carefully chosen criterion. The general term for this adaptive or updating sampling with the target of finding an optimal design is *Bayesian optimization*, which is the subject of this section, mainly due to the employment of Gaussian processes, which assume prior distributions of the observations and derive the posterior prediction based on Bayes' theorem. However, this term is mostly used in the computer science field, while engineering literature prefers the term *infill criteria*, as found in Forrester et al., 2008, for instance.

Two strategies can be distinguished:

1. **Exploration of the design space**
   This strategy is focused on increasing the global accuracy of a surrogate model, and supplements with training points the sparse regions of the design space. As an example, the points with the largest Kriging variance, equation 3.28, would be good candidates.
2. **Exploitation of the design space**
   Conversely, this strategy is focused on increasing the accuracy of a local region where the optimum is thought to lie. With this increase in local accuracy, the local minimum can be determined more precisely. An example would be to add points iteratively in the optimum of the response surface until that region of interest is accurately captured.

While both strategies have their benefits, there is a strong need for an approach that balances both exploration and exploitation. If the chosen method prioritizes the exploration of the design space, too many infill points may be added in regions which do not contain the global minimum, thus wasting precious time on modelling suboptimal regions. On the other hand, if the method is biased towards exploitation, the global optimum may never be found, as the local optimum determined from a response surface may not even be in the same region as the global optimum. As Forrester and Keane, 2009 put it, the choice of infill points constitutes the "heart" of surrogate-based optimization processes.

The discussion then falls back on how to find this balance. It has been hinted in section 3.3 that prediction error estimates are useful in this scope. To reiterate this idea, if the Kriging metamodel is employed, the surrogate is modelled as the realization of a stochastic process, i.e. the prediction at a point $\boldsymbol{x}$ is a normally distributed variable with mean $\hat{f}(\boldsymbol{x})$ and variance $s^2(\boldsymbol{x})$. Forrester et al., 2008, p. 86 state that by considering the possibility that the response at an unobserved point could take values different from the mean because of a large variance, an infill criterion could be constructed to balance the discrepancy between the mean and the variance. The present section will start with unconstrained problems, and their incorporation will be tackled in section 3.5.2. Three such criteria are distinguished in Forrester and Keane, 2009 to be used with Gaussian process-based surrogate models: statistical lower bound, probability of improvement, and expected improvement. These are also called *acquisition functions* in the context of Bayesian optimization.

However, only the expected improvement approach is widely used in surrogate model-based optimization. Indeed, the statistical lower bound is based on minimizing the following quantity, where A is a user-supplied value:

$$LB(\boldsymbol{x}) = \hat{f}(\boldsymbol{x}) - As(\boldsymbol{x}), A \in (0, \infty) \tag{3.46}$$

The goal of the optimization process would be to minimize the statistical lower bound, such that the optimal design point which minimizes the objective function is eventually discovered:

$$\min_{\boldsymbol{x}} \quad LB(\boldsymbol{x}) \tag{3.47}$$

The user-defined value of A, which is the multiplying coefficient of the standard deviation s, can be seen as the “weight” of exploration. As A tends to 0, the bound to be minimized is the predictor itself, leading to pure exploitation. Conversely, as A increases, the second term, which is proportional to $-s(\boldsymbol{x})$, becomes more and more important, therefore the variance is sought to be maximized. The latter corresponds to finding those regions in the design space which have the largest uncertainty, falling back to exploration. However, choosing the value of A is a challenging, problem-dependent task, and there is the risk of promoting too much exploration or exploitation.

To counteract this deficiency, the probability of improvement does not require any input, and is based on maximizing the quantity in equation 3.48. The minimum observed response value is denoted by $y_{min} = \min(y^{(1)}, \ldots, y^{(n)})$. Integrating the probability density function of the normal distribution of a point $\boldsymbol{x}$, $Y(\boldsymbol{x}) \sim \mathcal{N}(\hat{f}(\boldsymbol{x}), s^2(\boldsymbol{x}))$, yields the likelihood of obtaining a value lower than y_{min}, also called *the probability of improvement*, which will be:

$$\begin{aligned} P[I(\boldsymbol{x})] &\doteq P[Y(\boldsymbol{x}) \leq y_{min}] \\ &= \int_{-\infty}^{y_{min}} \frac{1}{\sqrt{2\pi}s(\boldsymbol{x})} \exp\left[-\frac{1}{2}\left(\frac{t - \hat{f}(\boldsymbol{x})}{s(\boldsymbol{x})}\right)^2\right] dt \end{aligned} \tag{3.48}$$

The goal is to find those infill points which maximize the probability of improvement, in order to find the design with the greatest probability of being lower than the minimum observed value of the objective function:

$$\max_{\boldsymbol{x}} \quad P[I(\boldsymbol{x})] \tag{3.49}$$

As the minimum response y_{min} is smaller than the mean $\hat{f}(\boldsymbol{x})$, and the standard deviation $s(\boldsymbol{x})$ eventually decreases with additional training points and an increase in accuracy, the term $\frac{y_{min} - \hat{f}(\boldsymbol{x})}{s(\boldsymbol{x})\sqrt{2}}$ becomes more negative. Thus, as $\frac{y_{min} - \hat{f}(\boldsymbol{x})}{s(\boldsymbol{x})\sqrt{2}}$ approaches $-\infty$, convergence is assured, as $P[I(\boldsymbol{x})]$ tends to 0 and no more infill points are added.

However, the convergence rate may differ from problem to problem. As the variance of the Kriging predictor is just an estimate, it may experience brief surges with the addition of training points in unexplored regions, especially if those areas have a non-linear trend which was not previously captured. Another variant of this method proposed by Jones, 2001 is to substitute y_{min} in equation 3.48 with another variable $T < y_{min}$ which would have a more stringent target value for improvement. The smaller T is, the more global the search. However, the choice of T again represents a setback, but at least the algorithm would converge sooner. Jones, 2001 suggests that one could use several values of T of different magnitudes, providing several search points in each iteration. Moreover, the probability of improvement does not offer any information regarding how much improvement is expected to take place.

Indeed, it may be the case that many infill points are suggested which have a high probability, but the actual improvement of the objective function is negligible. Nevertheless, there are better options, and the next acquisition function aims to counteract all the deficiencies of the last two.

3.5.1. Expected Improvement

The approach based on expected improvement follows the same trend as the last strategy, but aims at quantifying *how much* improvement is expected. Assuming yet again that $Y(\boldsymbol{x})$ is a random variable, then an improvement of I will be achieved if $Y(\boldsymbol{x}) = y_{min} - I$. Under these assumptions, the expected value of the improvement will then be:

$$E[I(\boldsymbol{x})] = \int_0^{\infty} I \frac{1}{\sqrt{2\pi} s(\boldsymbol{x})} \exp\left[-\frac{1}{2} \left(\frac{y_{min} - I - \hat{f}(\boldsymbol{x})}{s(\boldsymbol{x})} \right)^2 \right] dI \tag{3.50}$$

The goal of the optimization process is to find the infill points which maximize the expected improvement. Fortunately, as mentioned before, the improvement to be expected during the optimization process is now quantified, thus providing a more intuitive understanding on the convergence of the approach:

$$\max_{\boldsymbol{x}} \qquad E[I(\boldsymbol{x})] \tag{3.51}$$

For conciseness, the symbol $\boldsymbol{x}$ will be dropped. Considering the substitution $\frac{y_{min} - I - \hat{f}}{s} = u \Rightarrow dI = -sdu$, the integral becomes:

$$\begin{aligned} E[I] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{y_{min}-\hat{f}}{s}} (y_{min} - \hat{f} - su) \exp\left(-\frac{1}{2}u^2\right) du \\ &= (y_{min} - \hat{f}) \int_{-\infty}^{\frac{y_{min}-\hat{f}}{s}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right) du + \frac{s}{\sqrt{2\pi}} \int_{-\infty}^{\frac{y_{min}-\hat{f}}{s}} (-u) \exp\left(-\frac{1}{2}u^2\right) du \\ &= (y_{min} - \hat{f}) \Phi\left(\frac{y_{min} - \hat{f}}{s}\right) + \frac{s}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right) \Bigg|_{u=-\infty}^{\frac{y_{min}-\hat{f}}{s}} \\ &= (y_{min} - \hat{f}) \Phi\left(\frac{y_{min} - \hat{f}}{s}\right) + s\phi\left(\frac{y_{min} - \hat{f}}{s}\right) \end{aligned} \tag{3.52}$$

In equation 3.52, $\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right) du$ and $\phi(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right)$ denote the standard Gaussian cumulative distribution and probability density functions, respectively. The infill criterion will then constitute of finding those points in the design space which maximize the expected improvement. Compared to the probability of improvement, the "gain" is now quantified, as the acquisition function also computes how much the modelled function is expected to diminish, not only that there is some chance of obtaining a lower response.

Following this strategy, Jones et al., 1998 proposed the Efficient Global Optimization (EGO) method to add new sampling points iteratively to improve global optimization based on the expected improvement approach. This algorithm is also implemented in both the SMT toolbox and Trieste. While traditionally EGO adds infill points one by one, there is also a parallel version called the *Kriging believer*, which can suggest more infill points at the same time, in order to take advantage of parallel computing facilities. Practically, the algorithm works in the following manner: the point of maximum expected improvement is found. Afterwards, the value predicted by the surrogate model is imputed in the same point. The next point is then found on the surrogate model trained with the additional imputed point, and so on. Of course, because the surrogate model is not retrained after each infill, the expected improvement at each point is computed unknowingly of the true response, therefore it could be erroneous, and Fang et al., 2017 mention that this parallel algorithm has not been sufficiently tested. Nevertheless, if parallel computation facilities are available, the optimization wall time can be significantly decreased in this manner.

Forrester and Keane, 2009 do mention that EGO is the most popular method based on the sheer volume of citations that Jones et al., 1998 received. It is also more robust than the probability of improvement strategy and converges faster. However, an inherent disadvantage of two-stage approaches like the ones presented earlier is that they can be easily tricked by deceiving functions or unfortunate DoEs, leading to the algorithm completely ignoring certain regions, thus locating local, not global optima. The main reason is that the variance can be severely underestimated, rendering the infill criterion useless. Forrester and Jones, 2008 propose a one-stage approach, but, unfortunately, only unconstrained optimization problems are tackled. Nevertheless, the method can be found in appendix B, as the results presented by Forrester and Jones, 2008 seem promising.

3.5.2. Incorporating constraints

Last but not least, constraints can be easily taken into account in the frameworks of all three approaches exposed in the present section. For instance, for the expected improvement acquisition function which will be used in the present thesis, if there are N_c inequality constraints denoted by $c_i(\boldsymbol{x}) \leq 0, \forall i \in \overline{1, N_c}$, Schonlau et al., 1998 propose that the constrained EGO will be:

$$E[I(\boldsymbol{x})]_{constrained} = E[I(\boldsymbol{x})] \prod_{i=1}^{N_c} P[c_i(\boldsymbol{x}) \leq 0] \tag{3.53}$$

If the constraint function is known, then the probability is either 0 or 1, based solely on the feasibility of the design. However, if the constraint function is also approximated with surrogate models, then the probability of the constraint function being satisfied is similar to the probability of improvement back in equation 3.48:

$$P[c_i(\boldsymbol{x}) \leq 0] = \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi} s_i(\boldsymbol{x})} \exp\left[-\frac{1}{2}\left(\frac{t - \hat{f}_i(\boldsymbol{x})}{s_i(\boldsymbol{x})}\right)^2\right] dt, \forall i \in \overline{1, N_c} \tag{3.54}$$

In the equation above, $\hat{f}_i(\boldsymbol{x})$ is the Kriging predictor of the i$^{\text{th}}$ constraint, while $s_i^2(\boldsymbol{x})$ is the corresponding variance at a point $\boldsymbol{x}$. The expression evaluates the probability of feasibility based on the treatment of the constraint function as a stochastic process, which is simply the cumulative normal distribution function of a random variable with mean $\hat{f}_i(\boldsymbol{x})$ and variance $s_i^2(\boldsymbol{x})$.

Finally, the goal of the optimization process for constrained problems will be to find the infill point which maximizes the constrained expected improvement:

$$\max_{\boldsymbol{x}} \qquad E[I(\boldsymbol{x})] \prod_{i=1}^{N_c} P[c_i(\boldsymbol{x}) \leq 0] \tag{3.55}$$

4

Problem description

As the focus of this thesis involves the development of a methodology, the present chapter will construct a case study to demonstrate its application. As stated in the literature review conducted in chapter 2, two constraints will be formulated, pertaining to the maximum penetration depth, and the residual strength of the front spar after bird impact. The objective of this use case is to prioritize simplicity and reduce susceptibility to implementation errors, while also being representative of the crashworthiness requirements detailed back in chapter 1. The damaged structure will be subjected to a simple yet illustrative load case, in order to emulate the get-home flight conditions of the aircraft following bird strike. Moreover, the introduction of the load case will also allow the extraction of more realistic constraints in the optimization problem, as static strength requirements for the damaged wing have not yet been taken into account in any literature involving structural optimization with respect to bird strike requirements.

The first section is going to present the wing which will be used as the subject of the optimization process, along with the hypothetical load case aforementioned. Afterwards, a discussion on which design variables should be taken into account is going to be made. The third section is going to formulate the constraint problem and translate it to an inequality constraint function. Finally, the last section is going to present in detail the Abaqus model of the wing.

4.1. Case study

A wing segment consisting of five bays delimited by six wingbox ribs and two spars will be considered as a case study. An even number of ribs has been chosen, such that the bird impact location is in the middle of the leading edge along the wing span, in the central bay, which is assumed to be a critical case. A single critical location will be considered, although more constraints can be formulated in order to represent more impact locations, which can be easily accommodated in the present methodology. Twelve stringers will be placed on the lower and upper skin panels, with six stringers on each side. Furthermore, six leading edge or nose ribs are present, placed on the same planes as the respective wingbox ribs.

It is assumed that considering only a limited portion of the whole wing is going to be sufficient in terms of methodology validation, although it is recognized that the responses of the constraint functions are going to suffer some modifications. For instance, the dynamic response of the wing subjected to bird impact is going to be affected by the smaller mass and the boundary conditions which will be considered for simplicity, but a trade-off is desperately needed between accuracy and computational efficiency, especially for the scope of this report.

On the one hand, a larger model would have greatly increased the computational time of a single analysis, while bringing no significant contribution to the overall methodology development. On the other hand, a smaller model would have needed to consist of at most four ribs. In such a case, because the root wingbox rib will be clamped for simplicity, it is expected that the solution would be greatly affected by the boundary condition. Indeed, the overly-stiff root rib would be only one bay away from the impact location, and the displacement field may be severely altered as a consequence. Nevertheless, for the five-bay wing, it will be assumed that the kinetic energy induced by the bird on the structure will transform into plastic or damage dissipation energy locally, away from the root bay, and that the boundary conditions will not affect the dynamic analysis greatly.

The implementation of novel structures to protect against bird strike damage is not the scope of this report. Therefore, the design choices will be quite traditional, in the sense that only metallic structures will be investigated. Some interesting examples of such novel structures, though, should be briefly mentioned. For instance, Liu et al., 2017 and Liu et al., 2018 investigate triangular nose ribs, while Zhao and Li, 2011 propose the “Optimized Corrugate Board Leading Edge” structure, where the leading edge skin is strengthened by riveting it to a corrugated aluminium sheet. However, it is expected that the developed methodology can accommodate such structures as well, given that suitable geometrical parameters and constraints can be formulated.

Moving back to the case study considered in the present report, the wing structure and loads are going to be adapted from the Fokker F27 Friendship and Fokker 50, which are similar turboprop aircraft. The reasoning behind this choice is based on its rather low cruise velocity, which simplifies the bird strike analyses, and on the abundance of data available on aerodynamic coefficients, supplied by van der Vaart and Muhammad, 1983, in order to derive sufficiently representative load cases.

4.1.1. Wing structural model

As mentioned earlier, only a five-bay wingbox will be considered. As the Fokker F27 does not have a constant rib spacing based on the cutaway drawing supplied by Flight Global Image Archive, 2011, a constant rib spacing of 450 mm has been chosen based on the Lockheed P-3 Orion data supplied by Niu, 1988, p. 268, which has a similar wing span.

While the Fokker has a tapered wing, a constant chord of 2400 mm is going to be assumed in the present report, which is an approximation of the value of the mean aerodynamic chord calculated by van der Vaart and Muhammad, 1983, p. 8. A constant value has been chosen for simplicity in order to have a symmetric structure with respect to the bird trajectory, while keeping the same aerodynamic coefficients, which will come into play in the next section. Page 9 of the same report specifies that NACA 64(4)-421 and NACA 64(2)-415 are the root and tip section airfoils, the root and tip chords being 3465 and 1400 mm, respectively. Moreover, in order to circumvent the study of the critical location of impact in the thickness-wide direction of the wing section location, the airfoil will be assumed to be symmetric, and the critical location is assumed to be directly in the leading edge, i.e. in the middle of the airfoil thickness. A thickness to chord ratio of 15% will be chosen to preserve the thickness to chord ratio of the tip section, therefore a NACA 0015 airfoil will be assumed to fulfil the needs aforementioned.

Akin to the Fokker F27, there will be two spars, with six stringers on each side of the skin panel between the spars, as showcased in a cutaway drawing found on Flight Global Image Archive, 2011. However, the stringer spacing is not known, therefore it will be assumed that the stringers are distributed uniformly between the spars. Moreover, a variety of cross-sections are available within the MDM package supplied by GKN Fokker. Without any knowledge regarding their nature, hat stringers have been chosen, with a height of 25 mm and lower and upper widths of 15 and 20 mm, respectively.

The spars themselves will be fixed at 20 and 80% of the wing chord, respectively. Their cross-sections will be a simple C-shape, with the flanges directed towards the leading edge. The width of the flanges will be 14 mm, representing three times a classical fastener diameter of 4 mm, with one more millimeter on each side, in order to alleviate stress concentrations. The ribs will have the same flange widths. The root rib will have the flanges directed towards the tip, while the others will face the root.

However, it should be noted that no fasteners will be included, the joints between various components being modelled using tie connections, in a similar way to a perfect bond. This reduction in complexity is mainly owed to the fasteners not having been included in the Abaqus interface during the author’s internship, with further development of this capability being outside the scope of the present thesis. It is acknowledged, however, that the introduction of fasteners would improve the structural response to bird impact, as a portion of the kinetic energy of the bird would be dissipated by means of bearing and shear deformations around the fasteners.

The final structure, as rendered by the MDM package, is shown in figure 4.1, excluding the upper skin panel. The components are rendered as shells with no thickness. It should be mentioned that while the trailing edge can be modelled in MDM, everything behind the rear spar will be excluded from the Abaqus model, as usually that portion is represented by a moveable part, and its inclusion would not have a paramount effect during bird strike.

Moreover, another aspect which is apparent in figure 4.1 is that the ribs and spars have no cutouts. Therefore, there are no stress concentrations due to such holes in the components. On the other hand, as the nose ribs are less compliant, the risk of leading edge rupture is higher, as the nose ribs accommodate less deformations. As a consequence, the net effect of this simplification cannot be quantified beforehand, but it is acknowledged that the overall accuracy of the constraint function is affected. Nevertheless, the methodology is expected to work on models of higher fidelity.

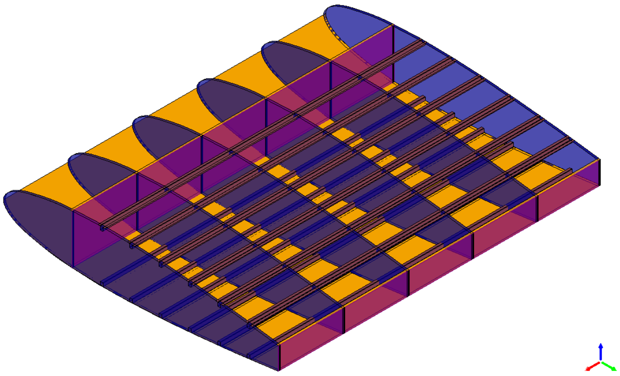


Figure 4.1: The geometry of the wing model used for methodology demonstration. Note: upper skin panel not included.

4.1.2. Design variables

The proposed methodology will be tested on a 19-dimensional design space, the variables being represented by the thicknesses of various components. While the Abaqus interface is able to generate complete architectural changes via MDM, the inclusion of variables such as spar locations or flange widths would overcomplicate the assessment of the methodology, as more failure modes would be activated. It will be assumed that only considering the thicknesses as design variables will suffice for this level of research.

Moreover, as mentioned earlier, the bird impact location will be fixed in the center of the leading edge, halfway through the wing span. It is acknowledged that considering only one impact location may be an oversimplification of the problem. However, it is envisioned that, in a similar manner to the application conducted by Ollar, Jones, and Toropov, 2017, a different surrogate model can be built for each impact location, which does come at an additional computational cost. Each new location gives rise to an extra constraint, which can be incorporated without any problem in the Bayesian optimization procedure based on the constrained expected improvement acquisition function, as seen earlier in section 3.5. However, the choice of such critical locations remains an open research question, especially if variables which change the architecture are considered, such as spar location or number of ribs. A possibility would be to use an analytical procedure for its determination, such as the one proposed by Yu et al., 2019 for metallic structures. For composites, though, the author of the present thesis has not been successful in finding such analytical procedures.

It is hypothesized that, the further away a component is from the leading edge, the smaller the effect its stiffness will have on the penetration depth after bird strike. To offer an example, the rear spar probably has a negligible contribution. Thus, this hypothesis will be tested on component thicknesses which vary in the chordwise direction. However, the contribution of wingbox components to the residual strength of the spar is expected to be significant due to changes in load distributions with varying thicknesses of the parts. The following variables will be considered for the optimization study, which are also showcased in figure 4.2:

- **Spar thicknesses**

  The front and the rear spar will have different thicknesses. However, these thicknesses will not vary in the spanwise direction. The flange and web thicknesses are the same, and will vary from 1.5 to 3.5 mm.

- **Stringer thicknesses**

  Each stringer station will be assigned with a different thickness, i.e. all six stringers on a given panel will vary. However, a stringer on the upper panel will have the same thickness as its lower correspondent, although it is acknowledged that this equality is not optimal in terms of static strength, as the lower panel is subject to tension and the upper one to compression. The thicknesses will range from 0.8 to 2.0 mm.

- **Rib thicknesses**

  All nose ribs will have the same thickness, as well as all wingbox ribs. However, the nose and wingbox thicknesses can be different. The flange and web thicknesses are the same, varying between 0.8 and 2.5 mm.

- **Skin panel thicknesses**

  The upper and lower skin panels will have the same thickness distributions. However, they will be split into 9 material zones of different thicknesses in the chordwise direction: 2 belonging to the leading edge, and 7 to the wingbox. The thicknesses will range from 1.0 to 3.5 mm.

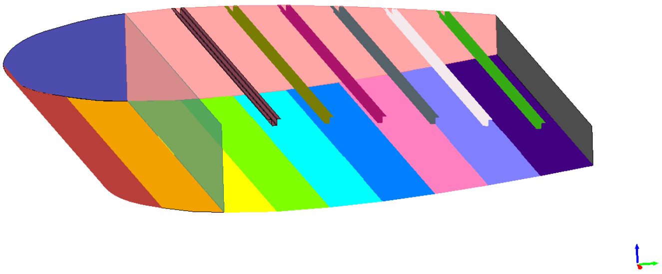


Figure 4.2: Optimization variables. Each colour represents a different thickness. Note: the upper skin panel, lower stringers, and the remaining nose and wingbox ribs have not been included for clarity.

The thickness ranges have been chosen somewhat arbitrarily, although it has been tried to preserve a resemblance with reality. However, the upper bounds are relatively large, mainly due to the exclusion of fasteners which, as discussed before, have a beneficial effect on energy absorption capabilities. Finally, table 4.1 presents the optimization variables, their ranges, as well as various geometrical properties relevant to the components or the wing itself.

Component Property	Geometrical features	Thickness range [mm]
Airfoil	NACA 0015	-
Span	2250 mm	-
Chord	2400 mm	-
Spars	Flange width: 14 mm Front spar position: 20% chord Aft spar position: 80% chord	1.5 - 3.5
Stringers (6x upper, 6x lower)	Lower flange width: 15 mm Upper flange width: 20 mm Height: 25 mm	0.8 - 2.0
Ribs (6x nose, 6x wingbox)	Flange width: 14 mm	0.8 - 2.5
Skin panels	-	1.0 - 3.5

Table 4.1: Summary of the geometrical features of the wing, as well as design variables.

4.2. Constraints definition

As mentioned in the introductory chapter, bird strike crashworthiness requirements are usually taken into account much later in the design process, and their incorporation in the preliminary design stages may yield significant weight savings. However, it was also mentioned that such requirements are not set in stone from a structural point of view: EASA CS 25.571(e) requires that "the aeroplane must be capable of successfully completing a flight during which likely structural damage occurs as a result of bird impact as specified in CS 25.631. The AMC to 25.571 (in paragraph 2.7.2) specifies the loads associated with get home conditions that have to be met for this case". Therefore, the damaged wing must be able to withstand a set of loads without experiencing further damage, and one of the constraints will target this certification aspect. The load case will be treated much more simplistically, and will be a combination of bending and torsion loads. This constraint will be tackled in section 4.2.1.

On the other hand, there may be some crucial components situated in the leading edge of the wing, such as a set of batteries in the case of an eVTOL, for instance. In this case, the depth penetration should not exceed a prespecified value, which will be formulated as a separate constraint in section 4.2.2.

These two constraints are going to be used for demonstrating in chapter 6 the proposed methodology, which will be detailed in chapter 5. As these two functions have different significant variables, as presented in chapter 6, they will also showcase how the methodology operates with different reduced design spaces.

4.2.1. The ability of the damaged structure to sustain a predefined set of loads

For simplicity, a constant set of loads is going to be used across the design space, although it is acknowledged that, realistically, it may vary significantly with some variables. Nevertheless, this simplification can still prove the effectiveness of the proposed methodology.

A combined load case consisting of bending and torsion loads will be considered starting from the Fokker F27 data supplied by van der Vaart and Muhammad, 1983. The following constraint will be imposed: the maximum Mises stress in the front spar should not exceed 80% of the yield stress of the material. In the present study, this value is approximately 295 MPa, and has been chosen as a safe figure in order to emulate the introduction of a safety factor and to allow room for various uncertainties which could be related to wind gusts, material properties, or even bird weight. Mathematically, this inequality constraint will be coined as $c_1(\boldsymbol{x}) \leq 0$, where $c_1(\boldsymbol{x})$ denotes the difference between the maximum Mises stress on the front spar and 295 MPa, in order to follow the same notation used in section 3.5.

A preliminary attempt was made, where a constraint function was defined as the maximum Mises stress encountered in the whole wingbox, regardless of the structural element. However, this definition has been abandoned as the function exhibited a rather erratic behaviour due to changes in the most affected component. Moreover, this all-encompassing constraint did not offer much information regarding the failure mode due to the same reason.

Similar constraints can be formulated with other structural parts which are deemed critical, but, for simplicity, only the front spar will be tackled. As no additional analyses are needed for the extraction of the maximum Mises stress in each component, it is advised to construct individual surrogate models instead of the all-encompassing constraint function mentioned earlier, as the accuracy of the individual metamodels is expected to be better thanks to a smoother behaviour. The front spar has been chosen mainly due to it being the wingbox component closest to the impact site.

Going back to the load case definition, for an assumed angle of attack of 2°, the lift coefficient for the NACA 0015 could be calculated as $C_L = 0.1357 + 0.1093/^\circ \cdot 2^\circ = 0.3543$, according to van der Vaart and Muhammad, 1983, p. 31. Moreover, the pitching moment coefficient is $C_m = 0.159 - 0.016/^\circ \cdot 2^\circ = 0.127$, as presented in van der Vaart and Muhammad, 1983, p. 34.

With the assumption of a cruise altitude of 7600 m, which is the maximum operating altitude of the similar Fokker 50, as given by Airlines Inform, 2012, the corresponding density is approximately $\rho = 0.55\ kg/m^3$, according to the International Standard Atmosphere. Taking the cruise speed of the Fokker 50 of $V_C = 530\ km/h \approx 147\ m/s$ according to the same source, with a wing surface of $S = 70\ m^2$, a mean chord of $c = 2.4\ m$, values supplied by van der Vaart and Muhammad, 1983, p. 8, the total lift and pitching moment become:

$$\begin{cases} L = \frac{1}{2}\rho V_C^2 C_L S \approx 147,380\ N \\ M = \frac{1}{2}\rho V_C^2 C_m c \approx 1,812\ Nm \end{cases} \tag{4.1}$$

However, in the present case study, only a five-bay wing segment is considered. For simplification purposes, by considering the whole wing as a beam of length $b = 29\ m$, it is assumed that a constant line load of $L/b \approx 5\ N$ acts perpendicular to the beam, in the z-direction, yielding an overall upwards force of $F_z(x) = L/b \cdot x \approx 5x\ N, x \in [0, b/2]$, where x is the coordinate in the spanwise direction. Therefore, a bending moment is also present, which varies according to $M_y(x) = \int_0^x F_z(\xi)d\xi = 5x^2/2\ Nm,\ x \in [0, b/2]$.

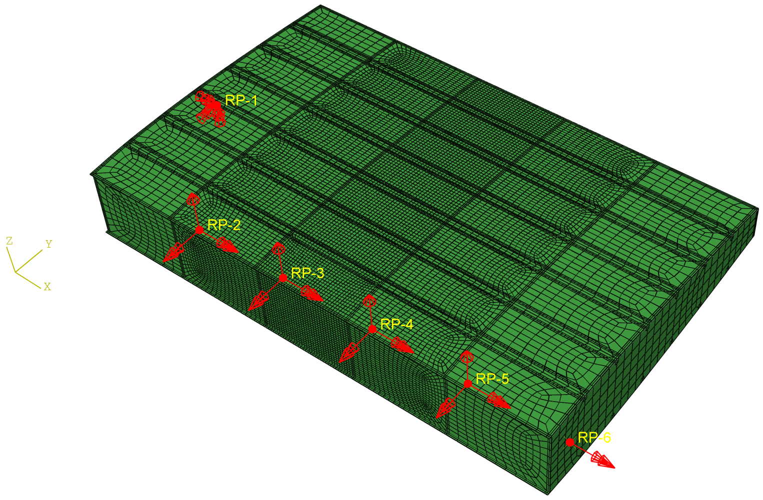
For an even further simplification, it is assumed that these distributed forces and moments act at concentrated reference points situated on each wingbox rib, at a quarter of the overall chord. These reference points spread the loads further to the nodes situated on all four flanges. As the rib pitch is 450 mm, by assuming that the wing segment is situated right at the tip of the wing, the following concentrated forces and moments are going to take place, as presented in table 4.2:

Wingbox rib number	Torsion moment [Nm]	Bending force [N]	Bending moment [Nm]
6	1,812	0	0
5	1,812	2,250	-506.250
4	1,812	4,500	-2,025.000
3	1,812	6,750	-4,556.250
2	1,812	9,000	-8,100.000

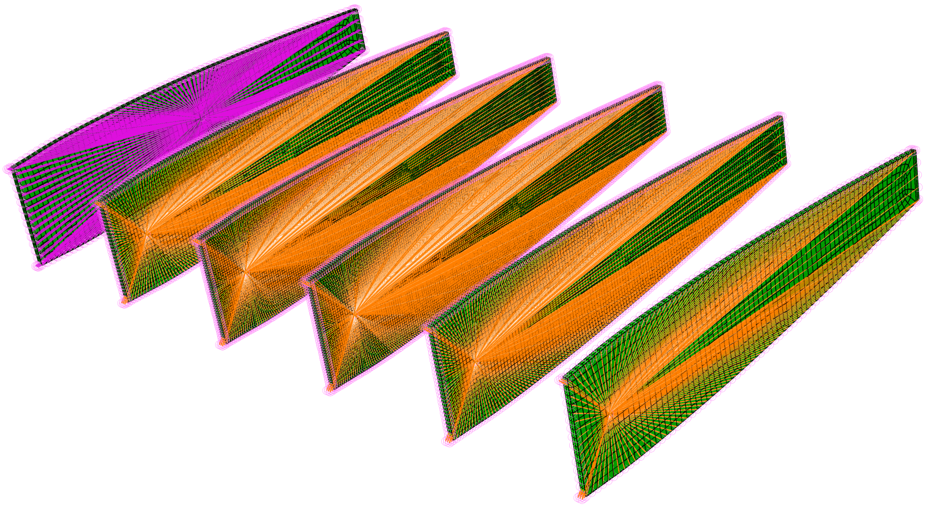
Table 4.2: Concentrated forces and moments applied to reference nodes situated on the wingbox ribs, at a quarter chord length from the leading edge. Note: the wingbox ribs are numbered from 1 to 6 in the spanwise direction, and a clamped boundary condition is applied on the first rib.

As a final assumption, only the elements of the wingbox will be considered for analysis. It is hypothesized that the main goal of the damaged structure is for the wingbox to have enough residual strength such that no further damage is ensued after bird strike, regardless of the severity of the damage caused to the leading edge or nose ribs. It is acknowledged that, on the one hand, disregarding the leading edge can yield an increase in the overall stress, as the leading edge may carry some load as well. On the other hand, if there is significant damage after bird impact, the penetrated leading edge can also lead to stress concentrations, resulting in an increase as well. However, it will be assumed that the two surges in stress are of comparable magnitude, and the leading edge will be excluded. Another argument to support this decision stems from pure time resources: the damaged leading edge may significantly increase the computational time, as numerous plastic strains and damage quantities would need to be solved at each increment for the numerous elements in the leading edge.

To foreshadow the Abaqus implementation and to illustrate the load scenario taken for this constraint, figure 4.3 presents the concentrated forces and moments acting on the reference points, alongside the couplings between the latter and the corresponding wingbox rib nodes. As a final note, it is acknowledged that the proposed load case is not severe due to the small magnitude of the forces and moments, but it has been chosen to emulate the get-home conditions which should not be critical. Indeed, for the intact wingbox of a minimum weight design, where all thicknesses are kept at their minima, the load case considered in the present section does not cause any plastic deformation to the front spar, the maximum Mises stress being 283 MPa, which corresponds to 76.7% of the yield stress of the material. Nevertheless, when applying the loads on the damaged structure, the narrative changes, as, in some cases, the maximum Mises stress of the front spar exceeds the yield stress of the material. Such cases are going to become apparent in chapter 6, when discussing the results of the constraint functions after acquiring an initial dataset.



(a) Concentrated forces and moments acting on reference points.



(b) The coupling between reference and wingbox rib nodes.

Figure 4.3: Abaqus wingbox model with concentrated forces and moments in reference points.

Although the Abaqus implementation of the static analysis responsible for extracting the value of this constraint function will be presented in the next section, it is important to highlight that its evaluation is highly demanding in terms of computation. As the analysis is conducted on the damaged structure, the static simulation should be performed after the bird strike analysis. The material state of the elements and the nodal displacements thus have to be imported from the dynamic to the static analysis, the latter starting with finding the static equilibrium of the structure, and then applying the loads in a uniformly increasing manner until their full magnitude is achieved. However, generally, the bird strike simulation solves the following basic equation of dynamic equilibrium:

$$[\boldsymbol{M}]\{\ddot{\boldsymbol{u}}\} = \{\boldsymbol{P}\} - \{\boldsymbol{I}\} \tag{4.2}$$

In the expression above, $[\boldsymbol{M}]$ denotes the mass matrix, $\{\ddot{\boldsymbol{u}}\}$ is the nodal acceleration vector, while $\{\boldsymbol{P}\}$ and $\{\boldsymbol{I}\}$ are the vectors of externally applied and internal element forces at the start of each increment, respectively. Therefore, the dynamic equilibrium simply states that the inertial forces acting on the body must be equal to the total nodal forces. On the other hand, the subsequent static analysis tries to find a solution to the static equilibrium equation:

$$\{\boldsymbol{P}\} - \boldsymbol{I} = \boldsymbol{0} \tag{4.3}$$

Therefore, it is clear that the static analysis can find an initial solution of the static equilibrium equation only if the inertial forces, $M\ddot{u}$, are approaching 0. However, due to the high kinetic energy of the bird, which is approximately $19.5\ kJ$, the vector of nodal accelerations requires a considerable amount of time to decrease. Two scalar quantities may be used to monitor whether this reduction is sufficient: the kinetic and total elastic strain energies of the wing structure. Indeed, while the kinetic energy of the bird is mainly dissipated through plastic deformation and damage accumulation, a portion of the kinetic energy is also transformed into sound energy and vibrations, which are closely related to the two energies aforementioned. As a consequence, the kinetic and elastic strain energies of the structure should become sufficiently small in order to continue with the static analysis, which might require a considerable simulation duration.

4.2.2. Imposing a maximum penetration depth

The second constraint regarding the maximum penetration depth is much easier to implement. No costly static analysis needs to be done on the damaged structure, which, as mentioned in the last paragraph, requires also the stabilization of the elastic strains and a sufficient decrease in the kinetic energy of the wing before the termination of the bird strike dynamic analysis. Because the elastic deformations are negligible with respect to the penetration depth at impact location, a simple termination criterion can be imposed, which mainly concerns the stabilization of the total plastic and damage dissipation energies, which is achieved in a much smaller duration. Indeed, because the bird strike is a local event, the plastic deformation is going to be concentrated near the impact location, i.e. in the two central nose ribs, the middle bay of the leading edge, the front spar being seldom affected. Therefore, there is no need to wait for the stress wave to propagate through the entire structure in order to extract the maximum penetration depth after impact, as all other components will undergo only elastic deformation.

Moreover, as the leading edge is penetrated by the bird, on one hand, severe plastic deformation occurs, but thanks to the strain hardening of the metal, the leading edge also starts resisting further deformation. Assuming a perfectly rigid target, the collision duration should be $L_{bird}/V_{initial} \approx 1.5\ ms$, therefore the real duration would be slightly higher. Figure 4.4 illustrates the relative difference between consecutive outputs of the elastic strain, kinetic and plastic and damage dissipation energies versus simulation time for the optimal design which will be further elaborated in chapter 6, when the results of the present case study are going to be discussed. It is apparent that the kinetic energy of the wing stabilizes rather late, while the plastic and damage dissipation energies do so approximately ten times faster. More details on the termination criterion for the Abaqus analyses are going to be given in the next section, but the main point of the present discussion is that the depth constraint is cheaper to acquire than the residual strength.

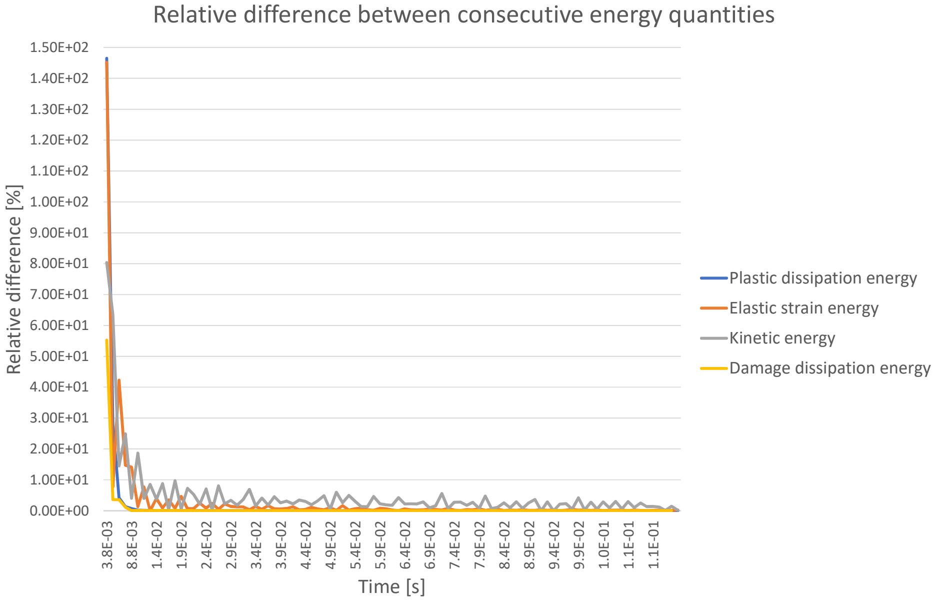


Figure 4.4: Relative differences between consecutive outputs of various energy quantities.

As the spar is fixed at 20% of the chord, i.e. at 480 mm from the leading edge in the chordwise direction, it is proposed to enforce a maximum penetration depth of half this value. It is hypothesized that, in the case of an eVTOL, for instance, battery packs may be placed in the leading edge, near the front spar, and that this distance should be enforced to ensure that the batteries do not suffer any damage. Any value would have been sufficient to demonstrate the efficacy of the methodology given that both feasible and unfeasible designs are discovered, and in the end, 68% of the initial data points were feasible with respect to the proposed constraint. Mathematically, this constraint will be denoted by $c_2(\boldsymbol{x}) \leq 0$, where $c_2(\boldsymbol{x})$ represents the difference between the maximum penetration and 240 mm.

As a side note, a preliminary exercise has been carried out in the early stages of the study, involving the minimization of the weight of the wing constrained by a no-rupture requirement. The latter was interpreted as no elements being deleted throughout the bird strike analysis. Only two variables were included, namely the leading edge skin and nose rib thicknesses. To model this constraint, a Variational Gaussian process (VGP) surrogate has been employed, which acts as a binary classifier, predicting whether rupture occurs. The exercise is briefly presented in appendix C, together with a review on VGPs. Although such surrogate models raise numerous problems concerning the variable screening procedure as the latter is no longer feasible under the formulation presented in section 3.4, the exercise shed light on an important aspect: that compliant nose ribs benefit the energy absorption capability of the leading edge, as their distortion accommodates the leading edge deformation without it rupturing. Therefore, the contradicting nature of the two constraints formulated in the present section is highlighted once more, further supporting their simultaneous inclusion in an optimization framework.

4.3. Constraint results extraction

Now that the constraint functions are settled, it is time to discuss how the results are going to be extracted. Abaqus has been chosen to carry out the bird impact analyses, and both the model and the workflows are going to be discussed, followed by a mesh convergence study for a design of minimum thickness.

4.3.1. Abaqus model and workflow

In the preliminary phases of the present work, numerous bird strike dynamic and static analyses have been conducted in order to find a suitable workflow for the extraction of both constraints. It was discovered that, without any damping in the explicit dynamic simulation, the kinetic energy of the wing decreased very slowly, or sometimes not at all. While the first case can be solved simply by increasing the simulation time, the second one is not realistic at all, as the metallic structure itself should provide partial damping. Indeed, as mentioned briefly in the last section, the dynamic analysis should be terminated when a quasi-static state is approached such that the static analysis which follows can achieve static equilibrium, therefore the kinetic energy of the model should decrease. Not only can this process be sped up by introducing material damping, but it is also more realistic to include it in the model. To give an order or magnitude, it was discovered that the analyses could be terminated 15% faster compared to the ones without damping, according to a termination criterion which will be discussed shortly.

As a consequence, in order to extract the results of the constraint functions detailed in the last section, the proposed Abaqus workflow is as follows:

1. Firstly, the first natural frequency of the model is extracted, which will be used to apply a Rayleigh damping to the model during the bird impact analysis.
2. Afterwards, the bird strike is analyzed using the explicit procedure. The simulation duration varies depending on the kinetic and elastic strain energy of the wing, as previously discussed. Moreover, automatic checks are made to ensure that the results are not affected by hourglass resistances.
3. Finally, a static analysis is carried out, which is used to assess whether the damaged structure can withstand the design loads detailed earlier in the present chapter.

For illustration purposes, a flowchart of the steps aforementioned is presented in figure 4.5. Before detailing each of them, there are three aspects to be elaborated, which are fixed for all analyses: the wing material, bird and bird material files.

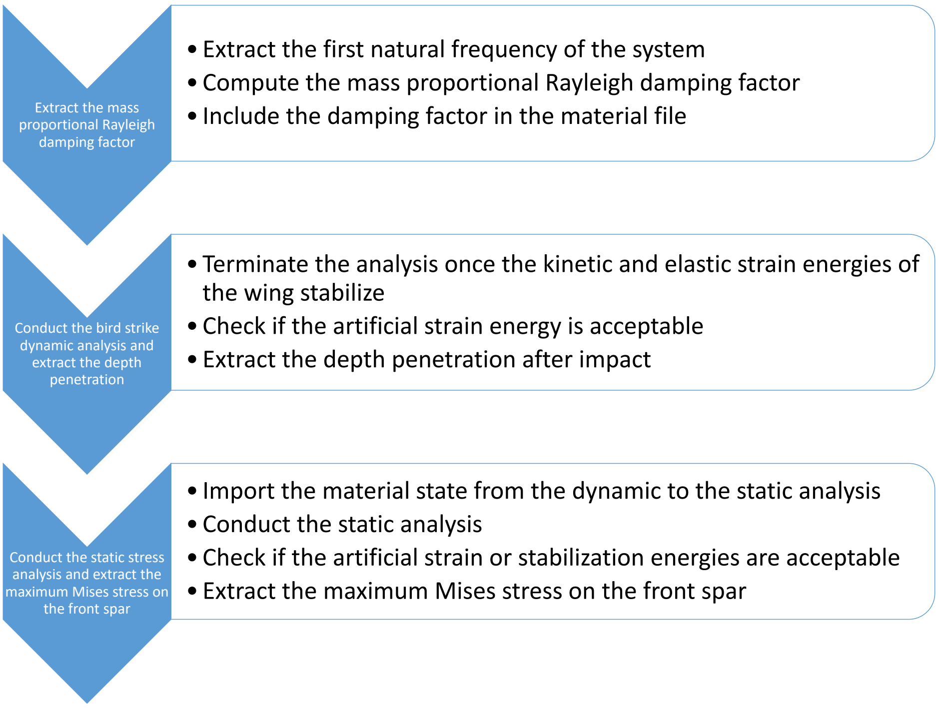


Figure 4.5: Flowchart of the tree analysis steps used for extracting the values of the two constraint functions.

Wing structural material input file

The aluminium 2024-T3 has been selected as the material of the whole structure, and, for simplicity, it will not be an optimization variable, although a mixed-integer Kriging surrogate model would be able to handle it. This choice has been made thanks to its popularity within the aerospace domain, and to the sheer abundance of data found in literature. It is recognized that different components can have different materials. However, aluminium 2024-T3 seems to be a popular choice in bird strike literature as well. Starting with the leading edge skin, Liu et al., 2018 used 2024-T3. While Guida et al., 2022 and McCarthy et al., 2004 both proposed Fiber Metal Laminates as the leading edge materials, the former used Al 7475 T-761 for the metal layers, while the latter employed 2024-T3. Nevertheless, Guida et al., 2022 used an Al 2024-T3 spar. Di Caprio et al., 2019 proposed a metallic sandwich leading edge, where the outer and inner layers were also Al 2024-T3. As for ribs, both Di Caprio et al., 2019 and Liu et al., 2018 employed the proposed 2024-T3. As a consequence, for simplicity, it will be assumed that the whole structure is made from the same material, although it is recognized that it is not realistic. Nevertheless, as hypothesized and as exposed in appendix C, more compliant structures have better energy absorption capabilities as they are able to dissipate a larger portion of the kinetic energy of the bird by deformation, therefore the 2024-T3 aluminium is expected to perform well for the leading edge skin, at least.

Most properties of this material are given by Vershinin, 2015, who conducted a validation study of plasticity and fracture models applied to a high velocity problem. The elastic properties are taken from Vershinin, 2015, the elastic modulus being 74,660 MPa, with a 0.3 Poisson ratio. The density is 2,770 kg/m^3.

The bird impact phenomenon can cause significant damage, and can induce large strain rates due to it being inherently a high-velocity phenomenon. Therefore, a rate dependent plasticity model is needed, along with a suitable damage criterion. As discussed earlier in section 3.1, a widely used plasticity model is the Johnson Cook hardening, which will be reiterated here for convenience:

$$\sigma(\varepsilon, \dot{\varepsilon}) = [A + B(\varepsilon_p)^n]\,[1 + C \ln \dot{\varepsilon}] \tag{4.4}$$

Fortunately, the model can be easily employed in Abaqus by defining the material constants. The study conducted by Lesuer, 2009, p. 8 gives these values, along with the rate dependent factor. A summary of these parameters is given in table 4.3.

Density [kg/m^3]	Poisson ratio [-]	A [MPa]	B [MPa]	n [-]	C [-]
2,770	0.3	369	684	0.73	0.0083

Table 4.3: Aluminium 2024-T3 material constants used in the analysis.

Moreover, Lesuer, 2009 also provides the parameters for the Johnson Cook damage initiation criterion. Unfortunately, though, this damage model can only be used with Abaqus/Explicit (Dassault Systèmes, 2022f, Johnson-Cook criterion section). Because the constraint function is formulated such that a static analysis needs to be run in Abaqus/Standard based on the results from the explicit simulation, this criterion cannot be used, as the material, which is imported from one analysis to the other, is not recognized. Therefore, another ductile damage criterion is going to be used. The damage initiation is given in tabular form by Vershinin, 2015, table 4, although the original data has been obtained by Bao, 2003. Vershinin, 2015 has been used as a source, though, only for convenience, as Bao, 2003 does not give the data in tabular format.

In order to foreshadow the implementation of both the dynamic and static analyses, the elements which are completely damaged are deleted. Therefore, the damage evolution needs to be defined. In Abaqus, there are two types of damage evolution criteria for elasto-plastic materials: either in terms of the plastic displacement after the initiation of damage or of the energy required for failure. The two are equivalent, as between them the following relationship takes place (Mol et al., 2022), with G_f being the fracture energy, σ_{y0} the ultimate stress, and $\overline{u}_f^{pl}$ the plastic displacement at failure:

$$G_f = \frac{\sigma_{y0} \times \overline{u}_f^{pl}}{2} \tag{4.5}$$

Therefore, the plastic displacement formulation has been chosen, as its value of $\overline{u}_f^{pl} = 0.001\ mm$ was readily found in Klosak et al., 2021, table 1.

Bird formulation and its material

Moving on to the bird, which is going to remain unchanged throughout all impact analyses, the SPH formulation has been chosen, as detailed earlier in section 3.1. The bird file, which includes the node locations and section controls, has been generated internally within GKN Fokker, not by the author, following the algorithm proposed by Siemann and Ritt, 2019. The resulting bird has 10,000 particles. As mentioned in section 3.1, a cylinder with hemispherical caps is the most suitable shape for bird impact simulations according to recent literature, with a length-to-diameter ratio of 2. The overall length of the bird is approximately 227 mm.

The material was chosen to have a density of 950 kg/m^3, as outlined in the data collected by Budgey, 2000. Its equation of state, relating the pressure to the volumetric strain, was given in tabular form using the data provided by Marulo and Guida, 2014, which is also given in table 4.4.

Volumetric strain [-]	0	-0.105	-0.118	-0.128	-0.137	-0.154	-0.169	-0.183	-0.195	-0.217
Pressure [MPa]	0	237	425	586	727	972	1180	1370	1540	1840

Table 4.4: Data for the bird equation of state, as given by Marulo and Guida, 2014.

Now that the material and bird input files are done, the actual Abaqus workflow is going to be described. As each analysis has its own input file with different procedures and solvers, they deserve a detailed discussion regarding how they are carried out. The only common aspects among them is that the same mesh is used, which will be covered in the next subsection, along with a convergence study. Moreover, the same boundary condition is defined, the wing being clamped at the root. The boundary condition is enforced via a kinematic coupling between a separate reference node situated in the middle of the root wingbox rib and all the nodes on the root wingbox rib flanges.

Moreover, the components are tied between each other using a surface to surface formulation, which “minimizes numerical noise for tied interfaces involving mismatched meshes”, according to Dassault

Systèmes, 2022j, although the only mismatched meshes present in the model are between the stringers and skin panels, as will become apparent later in the mesh convergence study.

Last but not least, all analyses are conducted with domain parallelization for a much smaller wall time. However, as all analyses are carried out on the university cluster, the number of available processors strongly depends on the time of the day, the number of available license tokens, and which queues are free, as they have different architectures.

Frequency analysis

While damping in structures can be caused by friction at various connections, or even by plastic or damage dissipation, it can also be caused by the material itself, according to Orban, 2011. A simplifying assumption mentioned in the introductory chapter was that the contact will be assumed frictionless, and, because the structure is essentially continuous thanks to the tie connections, no damping is provided by these mechanisms. However, including damping in the model can be very useful during high-velocity analyses such as projectile impact, as it can speed up the process of achieving a quasi-static state by helping to dissipate the kinetic energy of the structure, as exposed in the last section when defining the constraints.

Let $[\boldsymbol{M}]$, $[\boldsymbol{C}]$, $[\boldsymbol{K}]$ denote the mass, damping and stiffness matrices, respectively, $\{\ddot{\boldsymbol{u}}\}$, $\{\dot{\boldsymbol{u}}\}$, $\{\ddot{\boldsymbol{u}}\}$ the nodal acceleration, velocity and displacement vectors, and $\{\boldsymbol{F}(t)\}$ the nodal force vector at time t. Recall that the equation of motion states that the external load applied to a system is equal to the sum of the inertia, damping and stiffness load of the system, as shown in equation 4.6:

$$[\boldsymbol{M}]\{\ddot{\boldsymbol{u}}\} + [\boldsymbol{C}]\{\dot{\boldsymbol{u}}\} + [\boldsymbol{K}]\{\boldsymbol{u}\} = \{\boldsymbol{F}(t)\} \tag{4.6}$$

The damping force is then proportional to the nodal velocity, and Abaqus offers various ways to describe this proportionality coefficient, which is the damping matrix. However, the latter is mesh-dependent, and it is desirable to have an immutable parameter to characterize this damping. Such a parameter is the damping ratio, which is the ratio between the actual and the critical damping of a system, the latter being, simply put, that damping which is right at the crossroads between the system having an oscillatory or a non-oscillatory response. Orban, 2011 states that the damping ratio of a continuous metallic structure is between 0.02 and 0.04. In the present workflow, a ratio of 0.02 is assumed, in order to avoid overdamping.

Unfortunately, it seems not to be possible to define a constant damping ratio in Abaqus, as presented in Dassault Systèmes, 2022i. What is possible, though, is to use Rayleigh Damping, which assumes that the damping matrix is a weighted average between the mass and the stiffness matrices, as exhibited in equation:

$$[\boldsymbol{C}] = \alpha[\boldsymbol{M}] + \beta[\boldsymbol{K}] \tag{4.7}$$

The only coefficients which need to be specified are the weight factors. According to Dassault Systèmes, 2022i, though, for a given mode with a natural frequency ω, the corresponding damping ratio ξ is related to the Rayleigh factors in the following manner:

$$\xi = \frac{\alpha}{2\omega} + \frac{\beta\omega}{2} \tag{4.8}$$

It is clear from equation 4.8 that low frequencies render the second multiplying coefficient β negligible, whereas high frequencies yield the first multiplying coefficient α insignificant. Thus, the equation implies that the mass and stiffness proportional factors, α and β, are responsible with damping out the low and high frequencies, respectively.

One could, for instance, extract the first two natural frequencies, assume that the associated damping factors are equal to 0.02, and extract the Rayleigh parameters. However, Dassault Systèmes, 2022i demonstrates that the stiffness proportional damping factor, β, can dramatically reduce the stable time increment, and recommends that its value should be less than the initial time increment without damping. As the latter is no easy task, a workaround is proposed using solely mass proportional damping, where the α coefficient can be extracted using the first natural frequency:

$$\alpha = 2\xi\omega_1 = 0.04\omega_1 \tag{4.9}$$

This workaround is substantiated by the fact that, for higher natural frequencies, the damping factor is reduced, thus preventing overdamping. Moreover, the high frequencies are automatically dampened by Abaqus/Explicit with the introduction of bulk viscosity, which is also referred to as “truncation frequency damping”, as exposed in Dassault Systèmes, 2022h, section Bulk viscosity. As this mechanism is associated with volumetric straining and is activated by default, no further elaboration on the numerical stability is made.

However, it should be noted that the natural frequency of the structure may exhibit large variations due to the element deletion mechanism during bird strike analyses, as the mass matrix can alter significantly. Furthermore, because of the progressive damage behaviour of the material, the stiffness matrix can also undergo variations. As a consequence, overall, the damping matrix in the Rayleigh formulation can change drastically throughout the bird impact analysis. For a single spring mass system, a decrease in mass results in an increase of damping ratio, which is also the case for stiffness. Unfortunately, though, the author did not succeed to find a way to plot the actual damping ratio throughout the bird strike analysis, therefore it will be assumed that these changes will not result in a damping ratio higher than 0.04, which is twice as large as the value adopted at the beginning of the simulation, still being in the range recommended by Orban, 2011.

Finally, moving on to the actual implementation of this damping, the Abaqus analysis is straightforward. The default Lanczos eigensolver is used, and only the first natural frequency is requested. After job completion, an Abaqus Python script is launched in Abaqus/CAE which extracts its value, computes the mass proportional damping factor, and generates a new material input file to be used by the bird strike analysis, with the damping factor included.

Bird impact analysis

For all analyses, reduced-integration quadrilateral, S4R, elements are used. According to Dassault Systèmes, 2022d, reduced integration significantly reduces computational time, which is paramount to an optimization study, as many analyses are needed. A downside, though, is that hourglassing may occur, and this phenomenon will be presented in greater detail in the next subsection, during the mesh convergence study, along with a concrete example. Therefore, it is paramount that, after a bird strike analysis, the ratio between artificial and total strain energies is computed. A general rule-of-thumb obtained after discussing with various stress engineers within GKN Fokker would be to keep the ratio aforementioned under 5%. Dassault Systèmes, 2022d also recommends to define hourglass controls. Therefore, a section control is included in the input file, which states that an enhanced hourglass control formulation is to be used, along with the specification that elements are to be deleted once they fail, which is simply the default when modelling progressive damage.

All bird nodes are assigned an initial velocity of 150 m/s, roughly 540 km/h, a little bit higher than the Fokker 50 cruise speed, on the axis perpendicular to the leading edge. For simplicity, frictionless contact is assumed among all the components of the wing structure and the bird, with self-contact included. All the contact definitions aforementioned are needed, as the bird can collide with the leading edge, nose ribs and front spar, and the leading edge can come into contact with the nose ribs, front spar and even itself.

As the bird has an overall length of 227 mm, a rough calculation would yield an impact duration of $L_{bird}/V_{initial} \approx 1.5\ ms$. Nevertheless, this equation assumes a perfectly rigid target, which is not the case. Therefore, the impact duration is actually slightly larger. However, one of the constraint functions to be considered concerns the ability of the damaged structure to sustain a set of loads, which entails a static analysis after the explicit one. Therefore, the actual simulation time should be even larger, in order to wait for the stabilization of the elastic strains and the kinetic energy of the wing structure, as discussed in the last section.

Moreover, this stabilization may occur earlier or later depending on each specific case, due to different damping ratios and failure modes of the structure. A workaround has been to implement an Abaqus Python script which monitors real-time the elastic strain and kinetic energies of the structure, and terminates the analysis once a restart checkpoint has been reached, with the energies having been constant for a user-defined number of iterations. The restart parameter is needed to periodically write restart files either to recover failed analyses due to power outages or other random failures, or to serve as a starting point for the static analysis.

However, the term “constant” needs to be discussed, as the energies do exhibit some form of oscillation. Firstly, it was found that setting an analysis time of 0.2 seconds, i.e. more than a hundred times the rough duration of impact calculated beforehand, is a good maximum value for this case study, although most analyses were terminated at approximately 0.1 seconds. Nevertheless, if, in isolated cases, the analysis has not settled after 0.2 seconds, the energy outputs are verified, and a step continuation can easily be requested, beginning from the last available restart interval.

Secondly, the number of history outputs, which include the artificial, elastic, and total strain, as well as plastic and damage dissipation energies, was set such that an output is generated every 1 millisecond. Finally, it was found that checking if the relative difference between the maximum and minimum energy values for the last 0.01 seconds, i.e. for 10 history outputs, is below 3% could constitute a good termination criterion, as most subsequent static analyses were found to converge with this threshold.

Now that the termination criterion was discussed, the restart parameter deserves more details. As stated before, 0.2 seconds will be simulated, and 8 restart files are requested, resulting in one file every 0.025 s. Although a higher frequency is desired in order to promote a faster analysis termination, restart files can be large, and their writing can take significant time. For instance, an 8-interval restart file amounts to more than 7GB for this specific case study. Therefore, disk space is also of concern, and 8 restart intervals seem to offer a reasonable balance between memory usage and overall computational time.

Last but not least, field outputs are requested every 0.002 seconds, totalling 100 time intervals. Displacements, elastic and plastic strains, as well as the status of all elements are written. The constraint function concerning the maximum penetration depth of the leading edge will be obtained from the field outputs. It is acknowledged that, with this writing frequency, the actual maximum value can be overlooked, but in the end this problem is inherent to all discrete calculations. A workaround could have been to request nodal displacement output in the middle leading edge skin section with a larger frequency, but it is assumed that this is not needed. Indeed, plastic deformations dominate the displacements, any spring-back is expected to be captured by the dynamic analysis based on the stabilization of the kinetic energy, and elastic deformations are simply negligible, as their order of magnitude is much lower than the plastic ones.

Static analysis

The final analysis to be discussed is the static simulation needed for the constraint function involving the ability of the damaged structure to withstand the load case detailed in section 4.2.1. Abaqus/Standard will be used for the static analysis. The part instances are imported from the last restart file written by Abaqus/Explicit, along with the current material state of all elements, which includes stresses, equivalent plastic strains and information related to the damage model, according to Dassault Systèmes, 2022c. However, there are some limitations, as specified by Dassault Systèmes, 2022m: contact, tie connections, boundary conditions and multi-point constraint definitions need to be specified once again.

While the material of the bird strike analysis includes a damping model, the latter will not be used in the static analysis. What is of more interest, now, is the non-linear static step definition. It was found that good results in terms of convergence were obtained by applying initially a load of 5% of the full load, and then allowing small increments down to 0.001% of the full load during the solution scheme.

However, a small change is made to the model: the leading edge skin and the nose ribs are not considered in the analysis in order to speed up the computational time. Indeed, significant damage to the leading edge can exponentially increase the analysis time due to the large number of plasticity calculations. It will be assumed that the wingbox, whose components may exhibit rather large plastic strains after bird impact, especially the front spar, is the only one sustaining the load scenario. Even if, overall, the stresses in the wingbox are going to rise, the absence of the damaged leading edge can get rid of various stress concentrations which have an opposite effect. Thus, the net effect of this simplification cannot be easily established, and a comparative study would be interesting to pursue. However, as the accuracy of the test constraint function is not the primary scope of the report, a decrease in computational time is welcome.

Dassault Systèmes, 2022k states that, in large problems, the quasi-Newton solution technique can be less computationally intensive than using the full Newton method. Indeed, it has been found to be true, helping the minimum thickness design converge, as presented later during the mesh convergence study in section 4.3.2, when the full Newton method did not terminate the analysis in 100 iterations. Moreover, the line search algorithm is activated by default for the quasi-Newton technique, according to Dassault Systèmes, 2022l. However, to force convergence, the maximum number of line search iterations is increased from the default of 5 to 10, as it was found that some analyses needed more.

However, in the unfortunate cases of the proposed solution scheme not yielding a solution due to more iterations being needed than the default of 100, convergence can be forced by activating an adaptive automatic damping algorithm, as presented in Dassault Systèmes, 2022b, which dissipates a fraction of the total energy of the system. The “ALLSDTOL” parameter can be specified in order to set the maximum allowable ratio of the stabilization energy to the total strain energy, such that the solution is not greatly affected. A threshold of 5% is imposed, although it was found that the automatic stabilization algorithm overestimates the maximum ratio, and that, usually, a ratio of approximately 1% is obtained in the end. Nevertheless, most analyses converge without the introduction of any artificial stabilization, and the undampened solution scheme will be initially employed.

Finally, the loads defined earlier during the constraint function formulation are specified as concentrated loads on reference nodes which do not belong to the mesh of the wing structure, being solely created to distribute the load to the nodes on the wingbox rib flanges. This distribution is created using a distributing coupling constraint, which ensures that the loads and moments applied on the reference point are equivalent to the resultant of all the forces and moments on the coupling nodes.

4.3.2. Mesh convergence study

Before delving into the mesh convergence study, it is important to mention a few key aspects and assumptions that govern this investigation. Firstly, the MDM meshing interface developed during the author’s internship obeys the following rules:

1. All the edges in the geometry are seeded with an even number of segments, unless the edge length is smaller than a user-specified minimum value. In the latter case, the edge is seeded with a single element.

   This rule has been implemented to ensure that most faces would be meshed only with quadrilateral elements. As the meshing toolbox supplied by ParaPy, which is the core package used in the development of MDM, is based on the Salome platform, the generation of exclusively quadrilateral elements on a given geometrical face is enforced in the meshing interface via the “Quadrangle preference” 2D meshing algorithm. The latter has the restriction that the total number of segments on all sides of the face needs to be even (CEA/DEN et al., 2017).
2. Common faces between primitives have the same discretization.

   The FEM segmentation module present within MDM, which has not been developed by the author, splits the shell representations of various primitives in order to facilitate interface and section definitions. For instance, a skin panel may have more material zones, and it is therefore necessary to have separate faces with unique material assignments. However, the faces of the spar flange, which are in contact with the skin panels, would no longer have a one-to-one correspondence to the faces of the panel. Therefore, the FEM segmentation module also splits the spar with the same planes.

   Generalizing this action to all of the primitives, parts that are in contact have faces which are geometrically analogous. The meshing module thus enforces that congruent faces are discretized in the same way by projecting the mesh constructed on one of the faces on the other face. It is acknowledged that in Abaqus it is not necessary to have a bijective correspondence between the meshes on surfaces which should be tied together, although it is advised to choose the main surface of the constraint such that its mesh is coarser than the secondary one, according to Dassault Systèmes, 2022j. However, the same chapter does mention the fact that in the “node-to-surface” approach, a node on a secondary surface is tied together with nodes on the main one via an interpolating function which takes into account the distances among them. As a consequence, it is expected that conform meshes would produce more accurate results with the “node-to-surface” formulation, although in this thesis the “surface-to-surface” approach is used, which does not benefit from it.

Nevertheless, the meshing module is not based on these software recommendations, and produces conform meshes among primitives in contact. This choice was also made to preserve the rules provided by an automatic meshing tool used within the CoC Design department. The only exception to this rule are the stringers, which are meshed independently due to splitting problems in the FEM segmentation module. These issues concern excessively small faces created on the skin panels when upper and lower stringers have nearly coincident end of flange planes, which are used for splitting.

3. Common edges between primitives are seeded with the same number of elements.

   While this is a natural consequence of the rule aforementioned, some parts may not have entire faces in common, but only edges, such as the interface between the upper and the lower skin panel, for instance. Therefore, common edges between primitives also have a one-to-one node correspondence.

With these rules in mind, the mesh will have the following characteristics:

1. Each material zone of the skin panel may have different element sizes. The material zones are obtained by splitting the surfaces representing the wing components with various cutting planes, which include the planes used to obtain the conform meshes aforementioned, or the planes such that different thicknesses are assigned to each face, as illustrated back in figure 4.2.
2. Spar flange faces are meshed the same as the corresponding faces on the skin panels. This also results into the longitudinal edges of the web of the spar being already seeded. Only the widths of the spar are seeded according to the mesh configuration parameters. A detail of the meshes of the front spar and lower skin panel can be found in figure 4.6, which also showcases different element sizes for different material zones of the same component.

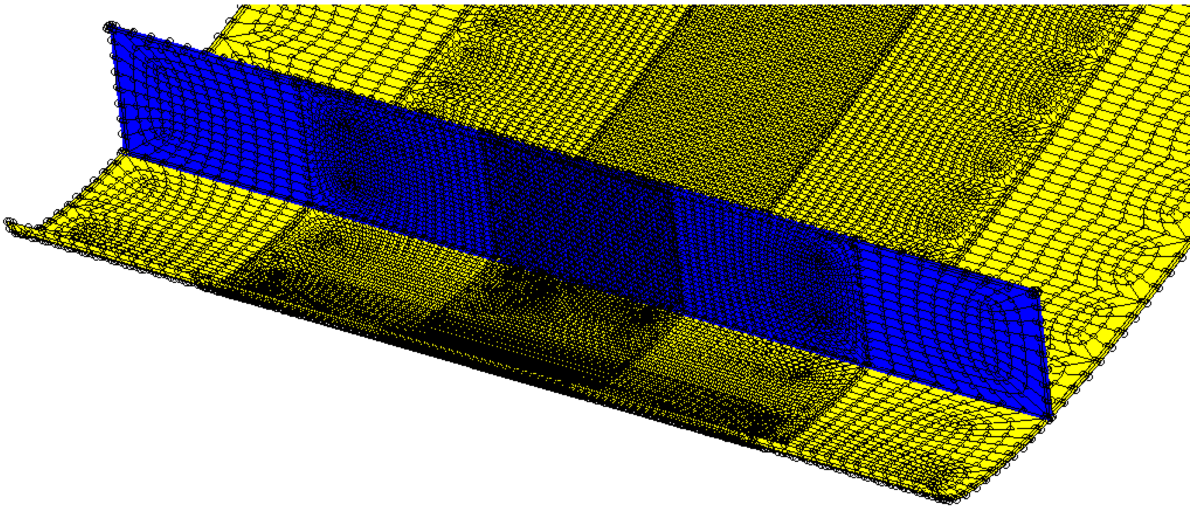


Figure 4.6: Mesh detail of the front spar and lower skin panel.

3. The number of nodes on the edges of a rib are entirely determined by the meshes of the skin panels and spars. However, in the case of a nose rib without any additional splits, the “Quadrangle preference” algorithm fails to generate an exclusively quadrilateral mesh. Therefore, some additional splits are performed as shown in figure 4.7, which transform the trilateral face to four quadrilateral faces. These interior edges of the nose rib are seeded according to the element size specified by the user.

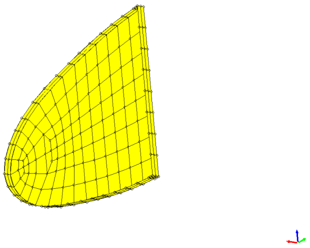


Figure 4.7: Mesh detail of a nose rib. Note: a large element size has been chosen for illustration purposes.

Even with the characteristics aforementioned, there are still a multitude of parameters which can be varied for the mesh convergence study. Moreover, as the bird impact location is always the same, it is expected that only the parts in the vicinity of the bird strike are going to undergo significant deformation. However, it is expected that a multitude of failure modes are possible due to the enormous number of thickness combinations. Nevertheless, it is assumed that performing the mesh convergence study on the design with all thicknesses set to the minimum of their respective ranges will suffice, where the most severe deformations are expected to occur.

As will become apparent later in this section, the minimum thickness design experiences significant nose rib crushing, leading edge skin penetration, and high stresses in the wingbox components. However, due to the expected large total number of elements, it is desired to vary the target element size with the components. Therefore, the following steps are proposed for the mesh convergence study:

1. Perform a mesh convergence study on the leading edge structure, comprising of only the leading edge skin, front spar, and nose ribs. The element size is assumed to be constant for all components, but as the complexity of the assembly is diminished, the overall computational time should not be prohibitive. Only the extraction of the first natural frequency and the bird strike analysis are conducted, as the structure is not representative for the static test.
2. The element size found in the first step is fixed for the components in the vicinity of the bird impact location, namely for the two central nose ribs and the leading edge skin between them. Afterwards, a total of 9 meshes are analyzed, where the element sizes vary in the chordwise and spanwise directions.

Leading edge mesh convergence study

As outlined earlier, the first step of the mesh convergence study is to analyze a smaller portion of the whole wing, consisting solely of the leading edge skin, nose ribs and front spar. The boundary conditions will be changed from clamping the wing at the root to constraining all the nodes on the front spar web in all degrees of freedom. The model geometry is presented in figure 4.8, which is a screenshot of the ParaPy model.

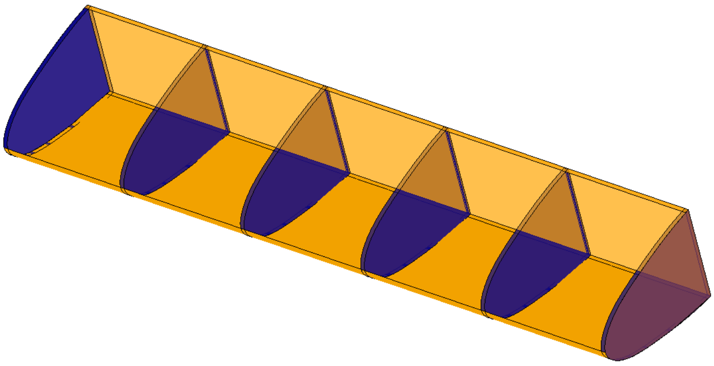


Figure 4.8: Leading edge model geometry for the mesh convergence study. Note: upper skin panel not included.

Four different meshes with increasingly small element sizes are studied, ranging from 20 to 5 mm, as presented in table 4.5. The analysis time was 0.2 seconds, assumed to be enough to judge which element size performs best. A longer simulation time would have only captured the stabilization of the solution without having a significant contribution to the plastic deformation of the components. Moreover, the stabilization of the solution and the diminishing of the various energetic quantities are also greatly influenced by the boundary conditions and the overall mass of the model, aspects which are not representative for the whole model.

Case number	1	2	3	4
Element size [mm]	20	15	10	5

Table 4.5: Element sizes used for the leading edge mesh convergence study.

Although these simulations do follow the framework of first extracting the natural frequency, as the boundary conditions and architecture are changed, no comparison is made among the cases considered. Therefore, only the bird strike dynamic analyses are used to establish the most suitable element size.

Starting with a 20 mm element size, the artificial strain energy comprised more than 12% of the total strain energy of the whole model, which comes as no surprise, given the coarseness of the mesh. Although enhanced hourglass formulation section controls have been defined, as outlined earlier in the present chapter, the employed elements have reduced integration, which are prone to hourglassing. The latter is a phenomenon whose resistive energy is included in the artificial strain energy, and occurs when zero-energy deformation modes are excited.

More specifically, as illustrated in figure 4.9, there may be some scenarios when the nodes of an element undergo significant displacements, thus resulting in large deformations of its edges. However, the integration point in the center of the element does not observe any deformation, as the two dotted lines are still perpendicular and of the same length. As a result, no strain is computed, which results in a null strain energy, in spite of the significant deformation, as explained in Abaqus, Inc., 2009. Nevertheless, as the artificial strain energy is too high, this case is dismissed.

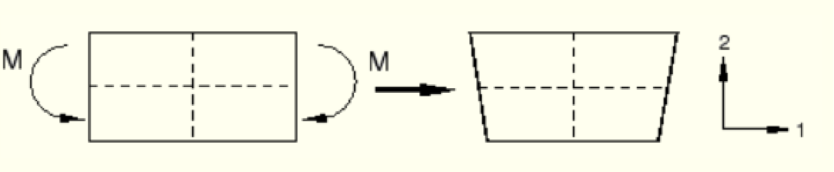
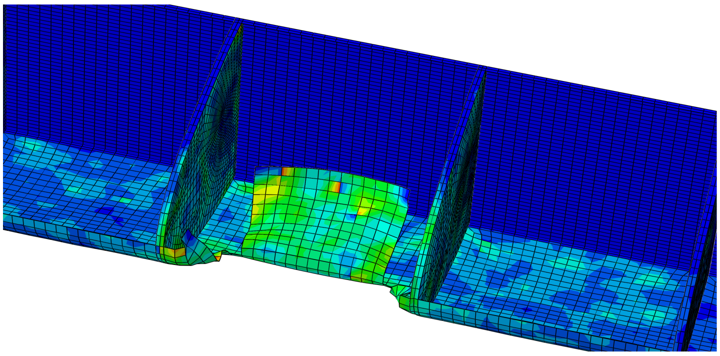


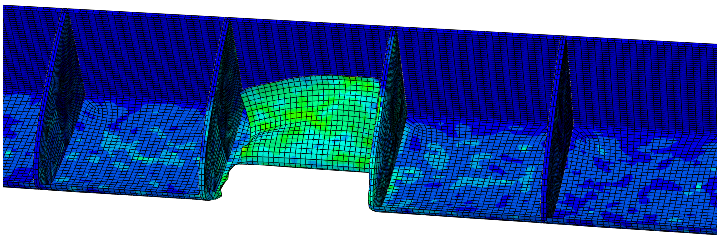
Figure 4.9: The hourglass phenomenon in reduced-integration elements. Image source: Abaqus, Inc., 2009.

Moving on to the second case corresponding to an element size of 15 mm, at the time mark of 0.072 seconds, a surge in total energy has been observed, increasing by more than four times. The analysis is automatically stopped once there is a surge of the kinetic and total energy of the entire model, by means of a Python script that automatically reads the status file written by Abaqus. After examining the output database file, it has been concluded that the increase was attributed to an upsurge in artificial strain energy, which was suddenly equal to roughly three times the initial total energy of the whole model.

It is quite interesting that this phenomenon has not been observed in the coarser mesh. However, as the strain distribution is vastly different, it is possible to induce hourglassing modes in more elements than before. As will be seen later, the true failure mode for this analysis is composed by both rib crushing and leading edge skin penetration. Although it is quite clear that both meshes are completely unusable, it is interesting to discuss the difference in artificial strain energy. As illustrated in figure 4.10, which shows the central section of the leading edge without the top skin, the second case attempts to capture the rib crushing phenomenon, as some elements are severely distorted on the top flange of the middle-right rib. This distortion could account for the drastic artificial strain energy increase, a distortion which cannot be seen in the first case.



(a) Case 1 deformation after bird impact, with an element size of 20 mm.



(b) Case 2 deformation after bird impact, with an element size of 15 mm.

Figure 4.10: Difference between the solutions obtained in the first two cases.

The third case of a 10 mm element size is no different, the artificial strain energy reaching a fraction of nearly 9% of the total strain energy of the whole model. The same comments can be made for this case, as illustrated in figure 4.11. No rib crushing has been captured, the rib flanges being seemingly stiffer, resulting in the leading edge skin tearing down in their vicinity.

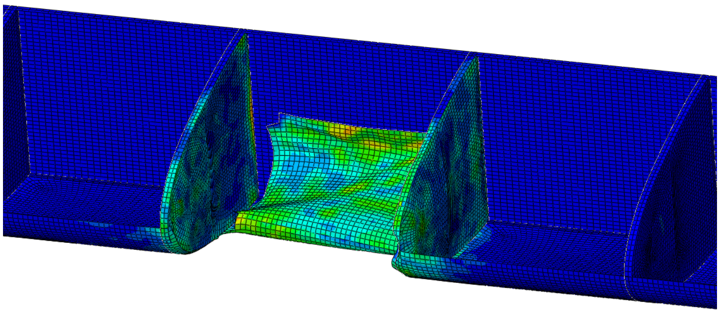
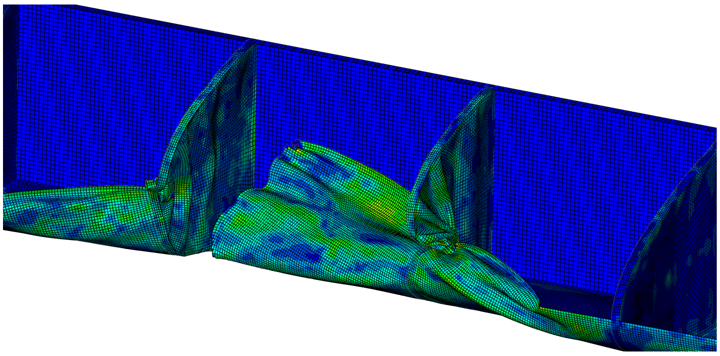


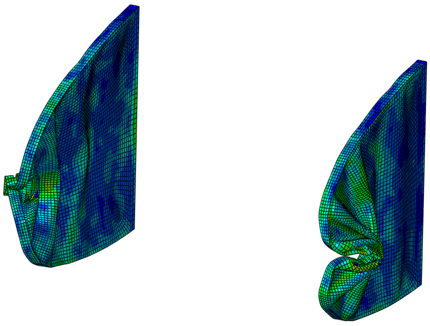
Figure 4.11: Case 3 deformation after bird impact, with an element size of 10 mm.

The final case of a 5 mm element size behaves much better in terms of the ratio between artificial and total strain energy, with a value of under 3%. As the minimum thickness leading edge design is expected to exhibit the largest deformations due to a lower energy absorption capability compared to the whole wing, it is assumed that, when analyzing the latter, the ratio will decrease even further.

Figure 4.12 illustrates the actual severity of the bird impact. As seen in the lower screenshot, the ribs are brutally crushed, with the leading edge skin tearing down on the left hand side, the fine mesh capturing also the wrinkling behaviour of the skin connected to the central rib on the right.



(a) Case 4 deformation after bird impact, with an elements size of 5 mm.



(b) Case 4 central ribs deformation after bird impact, with an element size of 5 mm.

Figure 4.12: Case 4 deformation after bird impact, with an element size of 5mm. The lower screenshot shows only the central ribs, which were severely crushed.

It is recognized that it would be recommended to try even finer meshes, but, as presented in table 4.6, the CPU time increases exponentially with a decrease in element size. The analyses have been conducted with 20 processors, therefore the wall time for the last case reached more than 6 hours, i.e. the CPU time divided by the number of processors. Moreover, the second case has not been included due to its failure. As the ratio between the artificial and the total strain energy is less than 5% for the minimum thickness design, it is assumed that there would not be a dramatic decrease in its value for an even finer mesh.

Moreover, the scope of the present thesis is the development of a methodology, and, as long as the failure modes are captured and the analyses do not experience failure and over-stabilization, the responses of the constraint functions are expected not to differ greatly from a qualitative point of view. Therefore, it is assumed that the observations fed into the surrogate model will still exhibit the same non-linearity and that the variable ranking will yield the same important variables, although there may be slight inaccuracies in the contribution quantification. As a consequence, the mesh convergence study will continue with a fixed size of 5 mm of the central leading edge bay.

Case number	1	3	4
CPU time [h:min]	16:54	27:18	121:11
Artificial-total strain energy ratio [%]	12.1	8.67	2.72

Table 4.6: Summary of the leading edge convergence study.

Whole wing mesh convergence study

As foreshadowed earlier, it is assumed that, as the bird strike is a local event, the further away an element is from the impact, the less affected it is. Therefore, it is proposed to increase the element size in both the chordwise and spanwise directions according to pre-established growth ratios. As the geometry of the whole wing consists only of shells with quadrilateral faces, these factors represent the ratio between the element sizes on two opposing edges. All other settings are kept constant besides the mesh, including the total time step of 0.2 s, for a better comparison between the performances of the cases.

These growth ratios are depicted in figure 4.13, and are exaggerated for illustration purposes. More specifically, starting from the central leading edge skin panel, the element sizes of the neighbouring panels in the spanwise direction are increased by the first spanwise ratio. The element sizes of the root and tip leading edge panels are then obtained by increasing the size aforementioned by the second spanwise ratio. Afterwards, for a given leading edge skin panel, the element sizes in the chordwise direction are increased in a similar manner.

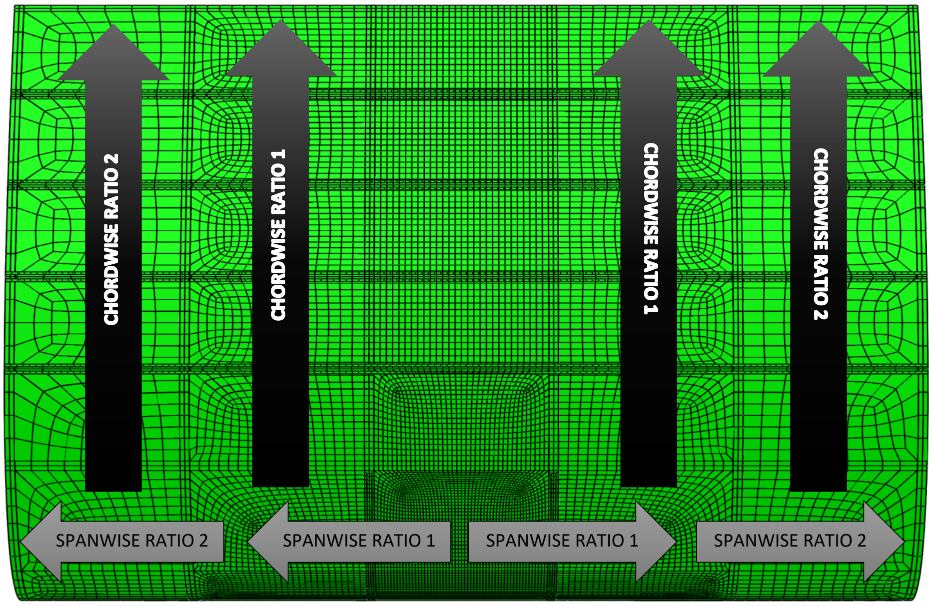


Figure 4.13: Growth ratios for the element size increase in both chordwise and spanwise directions.

It was found that the maximum growth ratio to ensure a non-distorted mesh is 2; otherwise, Abaqus issues a distortion warning. Therefore, two ratios are going to be tested: 1.5 and 2. It is assumed that one can have a more aggressive growth the further the component is from the impact location. For this reason, the second chordwise and spanwise ratios will be equal or larger than the first ones, respectively. In total, 9 cases will be considered for this phase of the convergence studies, with the ratio values presented in table 4.7. It should be noted that the cases are not enumerated in ascending order of the total number of elements.

Case no.	Chordwise 1	Chordwise 2	Spanwise 1	Spanwise 2	No. elements
1	1.5	1.5	1.5	1.5	219,526
2	1.5	1.5	1.5	2.0	212,530
3	1.5	1.5	2.0	2.0	189,560
4	1.5	2.0	1.5	1.5	219,529
5	1.5	2.0	1.5	2.0	212,536
6	1.5	2.0	2.0	2.0	189,551
7	2.0	2.0	1.5	1.5	159,752
8	2.0	2.0	1.5	2.0	153,570
9	2.0	2.0	2.0	2.0	137,344

Table 4.7: Growth ratio values and total number of elements for each of the cases considered.

Starting with the extraction of the first natural frequency, all analyses were successful, with the results presented in table 4.8. No case can be dismissed solely on erroneous frequency extraction, as the difference between the maximum and minimum values is only 0.07 rad/s. However, there is a large relative difference in CPU time among the cases, which is interesting to note. For instance, the first and fourth cases have nearly the same number of elements, but a large difference in solution time. It is not clear why this happened, though; it could be unrelated to the finite element modelling altogether, as the first and fourth cases were run on 32 and 20 processors, respectively. For short wall times, the overhead related to the domain parallelization may have taken a significant portion of time compared to the actual frequency and eigenvalue extraction.

Case no.	First natural frequency [rad/s]	CPU time [s]
1	236.09	1759.6
2	236.67	1602.9
3	236.71	1245.8
4	236.67	1215.8
5	236.09	1151.2
6	236.71	1268.6
7	236.95	861.44
8	236.95	762.39
9	237.02	630.2

Table 4.8: First natural frequencies of the cases considered in the mesh convergence study.

Moving on to the bird strike analyses, the eighth case can be easily dismissed based on the ratio of artificial strain energy to the total strain energy, which is much larger than the rest, as seen in table 4.9. However, as all other cases have an acceptable ratio, a more in-depth analysis needs to be made, based on all available energetic quantities.

Case no.	1	2	3	4	5	6	7	8	9
Ratio [%]	2.83	2.90	2.94	2.92	2.89	3.01	2.85	7.91	2.91
CPU time [h:min]	224:52	223:19	200:19	219:11	221:59	210:51	232:35	236:58	217:18

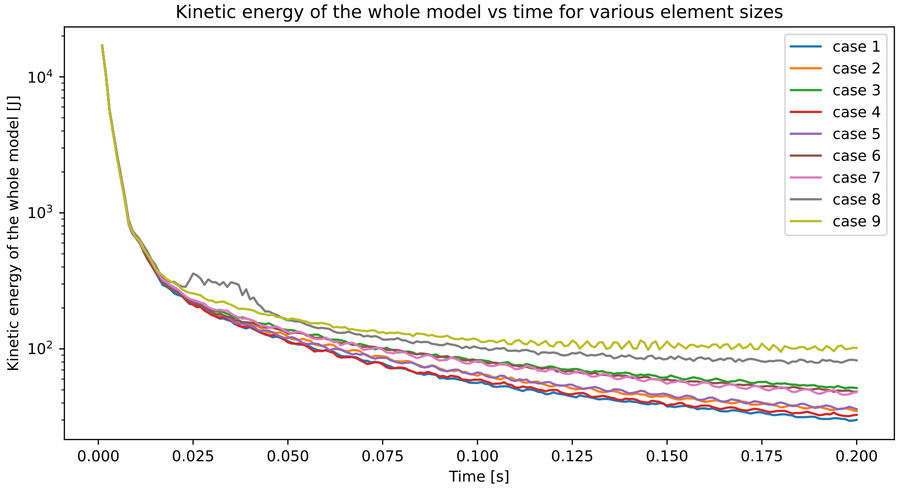
Table 4.9: Ratio between artificial and total strain energies, along with the total CPU time for the bird strike analyses.

Starting off with the total energy of the whole model, all cases are relatively the same throughout the whole simulation time, with a maximum relative difference of just under 0.025% among them. However, the kinetic energy of the whole model, which is included in the total energy, exhibits large variations, as seen in figure 4.14. Firstly, the fact that cases 8 and 9 do not exhibit a decreasing trend is worrisome, as the kinetic energy of the whole model should present a steady decline caused by damping. Therefore, case 9 will also be dismissed. Secondly, the lower figure shows the relative difference in kinetic energy, the reference being case 1, as it shows the most abrupt decrease. Unfortunately, cases 3, 6 and 7 show large discrepancies, which are apparent throughout the whole duration. As the differences surpass 50% at the end of the analysis, these cases will also be dismissed.

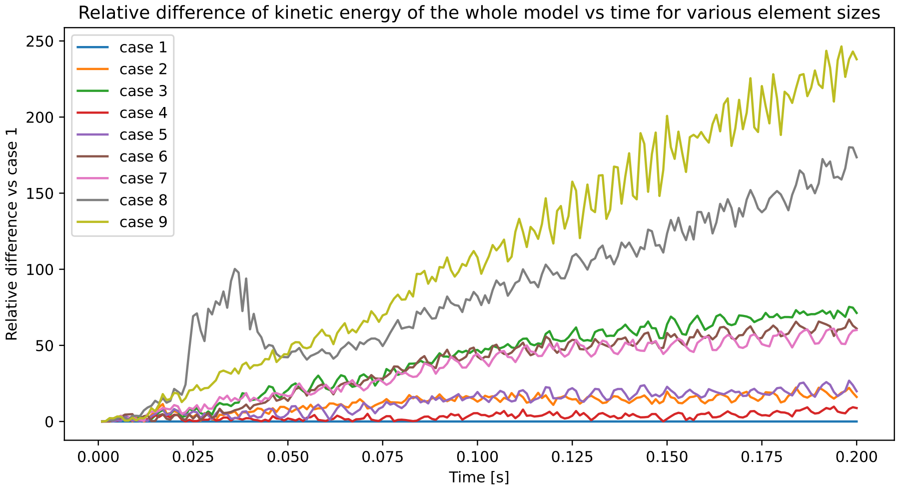
The remaining energetic quantities, namely total, plastic and elastic strain energies, as well as damage dissipation energy, are relatively the same across the remaining cases, with less than 0.5, 0.6, 5 and 2.6%, respectively. Moreover, the maximum penetration depths after bird strike are 426.296, 422.223, 421.712 and 420.262 mm, respectively. As a consequence, all the predicted depths have close values, the maximum absolute difference of approximately 6 mm being less than 1.5% of all reported depths. Therefore, no further cases can be dismissed based on the bird strike analysis. The remaining meshes will be compared against each other based on the static analysis.

The first try to obtain the solution to the static problem after a 0.2 second dynamic simulation is done using the default full Newton method. As discussed earlier, the loads are ramped up linearly over the simulation period, with an initial application of 5% of the total magnitude, and a minimum increment of 10^{-5}, as a smaller one would result in a prohibitive computational time. Moreover, the default number of 100 maximum iterations is left for the same reason. However, none of the cases managed to converge, having completed only 0.290, 0.642, 0.733 and 0.261 of the total time, respectively.

In the second attempt, the quasi-Newton was employed, as suggested earlier. The only analysis which converged was case 4, with all others needing a minimum increment lower than 10^{-5}. Therefore, the fourth case will be used as the base input file for all analyses. As the design variables consist solely of thicknesses, these input files are easily changed without resorting to the MDM package. However, as mentioned before, the platform is needed for architectural changes, such as spar positioning, for instance.



(a) Absolute values.



(b) Relative difference compared to case 1.

Figure 4.14: Kinetic energy of the whole model.

It is acknowledged that thickness changes, or, even more drastically, architectural changes, can result in different meshes performing better. However, a convergence study for every single design is unfeasible. Therefore, a constant mesh will be employed for all designs in the present study. The following quantities are checked in order to confirm a valid result: the ratio between the artificial and total strain energies, which should be kept under 5%, and the convergence of the static analysis. For the latter, no additional quantities are verified, if no artificial stabilization is employed. However, if stabilization is needed, the ratio between the artificial stabilization and total strain energy is verified to be under 5% as well. If these checks are not passed, the results are deemed inaccurate and will not be used for training. Results will be imputed to deter the optimization algorithm from requesting failed analyses, in a manner which will be discussed in further detail in chapter 5.

5

Proposed methodology

As foreshadowed back in the introductory chapter, the proposed methodology consists of the following general steps for a given constraint, as illustrated in figure 5.1, which was also presented in chapter 1. Firstly, an initial dataset will be acquired, based on an SLHD, thanks to its space-filling nature. As the proposed variance-based variable ranking procedure necessitates an initial surrogate model, the initial dataset will be divided in training and verification points. Two Kriging surrogate models will be built with the exponential and squared exponential kernels, which are suitable for underlying functions of different smoothness, and the most suitable one in terms of the root mean square error is chosen. The variable ranking procedure will then be conducted, and the model will be retrained on a reduced design space. These steps are repeated for each constraint, which will then be incorporated into the Bayesian optimization procedure.

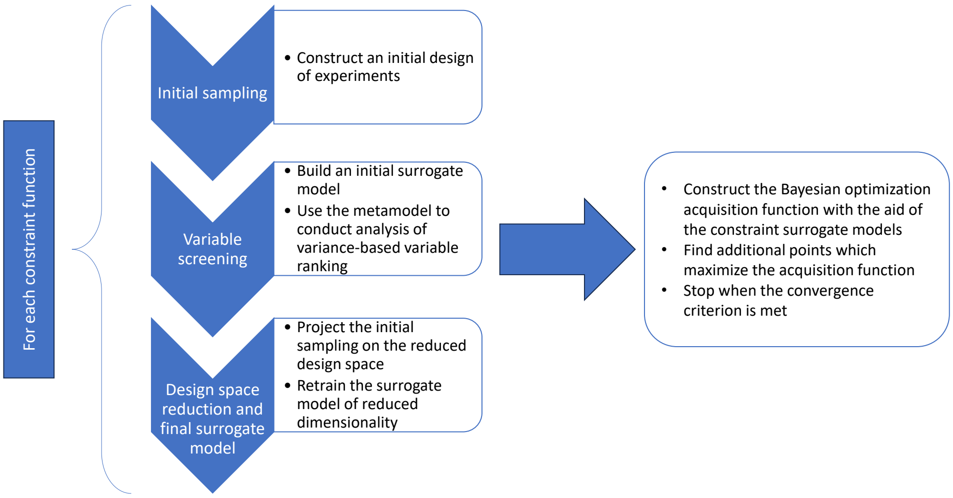


Figure 5.1: Flowchart of the proposed methodology.

In the present chapter, it is assumed that the constraint functions have already been defined, and that mesh convergence studies, if needed, have been carried out. Moreover, all inputs are assumed to have already been normalized to the unit hypercube. As mentioned earlier in chapter 4, it is acknowledged that there may not exist a single appropriate mesh for the whole design space. Therefore, it is proposed to conduct the convergence study only for a critical case, such as the minimum thickness design in the problem presented earlier, in section 4.3.2. If, by any chance, some analyses fail, values will be imputed for those sample points in a manner which will discussed shortly.

5.1. Initial data generation

An initial surrogate model will be built and serve as the foundation for the variable ranking procedure. This step is rather delicate, as a disadvantage of the proposed methodology is the lack of a recovery mechanism against an erroneous identification of significant variables, therefore a sufficient number of training points is needed. Indeed, as the screening procedure is dependent on the surrogate model predictions to compute the variance induced by each set of variables, the metamodel should be sufficiently accurate in order to prevent an erroneous computation of the contributions. On the other hand, as the data acquisition is computationally intensive, the exact opposite is wanted: to gather the least number of samples. Therefore, a careful balance needs to be achieved.

A prevalent figure for the number of initial samples seems to be 10 times the number of variables, as advised by Khatouri et al., 2020, who performed multi-fidelity constrained Bayesian optimization using surrogate models. However, the problem was only bi-dimensional, leading to a 20-point training sample, which is far from prohibitive in terms of computational time. In the case of the proposed 19-variable problem, though, 190 samples would be needed, each having a computational wall time of approximately 15 hours, such that a significant amount of time would be allocated to the initial data acquisition.

This advice has probably stemmed from Schonlau et al., 1998, who are also the authors of the famous EGO. Nevertheless, this value is cited only as a typical figure, as 120 training points did suffice for the 36-variable non-linear problem considered in Jones et al., 1998. Moreover, a sufficient accuracy of the ordinary Kriging metamodel has been achieved for the 20-variable function proposed by Welch et al., 1992 with 66 samples. Indeed, 3-4 times the number of variables seems more reasonable for the 19-variable case study which will be tackled in the next chapter.

Moreover, the global accuracy of the surrogate model is not of interest in the context of optimization; it is rather its ability of finding a minimum which is important, as put forward also by Viana et al., 2009. The goal of the initial surrogate model, in the end, is only to identify some trends in a qualitative manner in order to identify the correct significant variables via the analysis of variance procedure, while the quantification of the contribution of each set can be seen as an educational bonus.

Viana et al., 2021 also recommend splitting the sampling points in a validation and a training dataset, the former being half the size of the latter. As a consequence, it is proposed to generate an initial data set four times the number of variables, and to keep roughly a third for validation purposes. For the proposed case study, 76 samples will be generated, according to the algorithm proposed by Shan Ba and Brenneman, 2015, which produces Sliced Latin Hypercube designs.

5.1.1. Analysis workflow for data acquisition in the proposed case study

The responses of the constraint functions are going to be acquired by conducting the analyses proposed in chapter 4, which will be briefly reiterated here. The first natural frequency is extracted to include a damping ratio in the material definition, achieving roughly a damping ratio of 0.02, which is characteristic to continuous metal structures.

Afterwards, the bird strike dynamic analysis is carried out. An Abaqus Python script is surveying the latter in real-time, and does the following checks: if there is a surge in the total energy of the whole model or if the ratio between the artificial to total strain energy exceeds 5%, the analysis is terminated and deemed as failed. On the other hand, if the kinetic energy and the elastic strain energy of the model stay relatively the same for a number of increments, the analysis is terminated and the maximum depth penetration is extracted. A value which alleviated the convergence of the static analysis was found to be 3% for the kinetic energy and 1% for the elastic strain energy for 10 increments or 0.02 simulation time, as 3% showed a deceleration in the decrease of kinetic energy, while 1% exhibited a stabilization of the elastic strain energy. Moreover, the term “relatively the same” is judged as following: the average value of the energy quantity is computed, and the maximum relative difference from the average to an incremental value is compared to the values aforementioned.

As a side note, the Abaqus Python script is called via Abaqus/CAE, which takes one so-called CAE license seat, or license token. This aspect is totally undesirable, and the script was only called during night-time, when sufficient free seats were available. During daytime, on the other hand, the analyses were let to run without surveillance, for the whole duration of 0.2 seconds. A much better workaround would have been the implementation of a user subroutine embedded in the solver, which would not have used any CAE seat. A VUAMP subroutine has been tried (Dassault Systèmes, 2022n) thanks to its ability to parse sensor output, forcing step termination if the quantities aforementioned have stagnated. However, sensors specified as energy outputs are apparently not supported as of Abaqus 2021, and the author of the present report has not been successful in finding another solution, mainly also due to limited experience with Abaqus user subroutines in general.

Going back to the Abaqus workflow, the static analysis is carried out on the damaged structure. If it converges and the ratio between artificial stabilization, if employed, and total strain energy is less than 5%, the maximum stress of the front spar is extracted; otherwise, the analysis is deemed as failed and a value is imputed during the optimization procedure. Note that failure in the present section should be seen as the inability to extract the constraint results, not as infeasible designs.

5.1.2. Proposed sequential addition in case of failed analyses

It has been repeatedly stated that some analyses may fail as a result of an incorrect choice of mesh, and that such cases are inherent to the present methodology. In order to preserve the initial number of data points, the following procedure is proposed: after all samples are attempted, a surrogate model is trained, in a manner detailed in the next section. The remaining number of data points are then added according to those locations of the design space which exhibit the largest variance, which is, as discussed in section 3.5, a measure of uncertainty of the surrogate model.

This procedure can be done either sequentially, by adding the points one-by-one, retraining the surrogate model at each step, or in parallel, as detailed back in section 3.5, under the term of batch acquisition. Even though the latter has been presented in the context of the maximization of the constrained expected improvement, it is easily adapted to the variance, as it also needs to be maximized. Indeed, changing the acquisition function in the Trieste package is done virtually effortlessly, which is one of the reasons why it has been chosen as a tool for the present thesis.

It should be noted that, in the proposed case study, as the static analysis depends on the dynamic one, it may be the case that more data points are procured for the depth constraint. However, to foreshadow the results, the addition of more points did not yield much change in the root mean square error mainly due to the enormous design space. Therefore, although this sequential addition may result in an increased global accuracy, it is advised against its use with a smaller number of initial samples in the hopes of obtaining an initial surrogate model faster.

5.2. Initial model fitting

The initial model fitting treated in this section will not cover any aspects regarding the optimization of the hyperparameters, as the computational algorithms are not the scope of the present thesis. In both Trieste and SMT packages, the default hyperparameter optimizers will be used. The SMT package, version 2.3.0, employs the Cobyla optimizer, according to its documentation on the Kriging metamodel (Hwang, 2020). Trieste 1.2.0 wraps the “GaussianProcessRegression” model (The Trieste Contributors, 2020c) around GPFlow’s “GPR” class, whose optimizer is based SciPy’s minimize function (The GPflow Contributors, 2023) with the L-BFGS-B algorithm. However, what will be discussed is the generation of the validation and training datasets, and the kernel choice.

5.2.1. Splitting the dataset for validation and training

Before choosing the appropriate covariance function for the ordinary Kriging model, the validation points need to be picked from the initial dataset. These samples are needed to compute error measures such as the root mean square error (RMSE), although other metrics which do not employ other data points are available, such as the cross-validation error. The employment of the latter as a quality measure would be beneficial because an even smaller number of initial points would theoretically be needed, e.g. three times the number of variables. Indeed, the so-called k-fold cross-validation error is computed by splitting the whole dataset in k groups of roughly the same size. The error is then equal to the average prediction error of the model trained on the whole dataset except for one such group, keeping the latter as a validation set. Unfortunately, though, the results of chapter 6, which will also provide a comparison between the two metrics, will advise against relying solely on the cross-validation as a quality measure due to a large discrepancy to the RMSE. Therefore, the validation dataset is strongly necessary.

As mentioned before, Viana et al., 2021 recommends setting aside a third of the total samples for verification purposes. However, no aspects regarding how this split should be made have been discussed. As this initial split will be used to decide the most appropriate kernel, the following is proposed.

The discrepancy of a set is a measure of its degree of space-filling in a unit hypercube. The lower the discrepancy, the better. This measure is readily available within the SciPy package (The SciPy community, 2022), under the three criteria which are discussed in Y.-D. Zhou et al., 2013. However, the mathematical details are not of interest in the present section.

Given the initial dataset, a user-defined number of random subsets of the whole dataset are generated, and the one with the lowest discrepancy is chosen as a validation dataset. Therefore, the most space-filling subset will be used for verification purposes. As the root mean square error is also computed using a limited number of data points, it is also prone to error itself. It can be affected, for instance, by validating the model against points which are crowded in a small volume of the whole design space. Therefore, in order to assess the global accuracy of the surrogate model, which is solely needed for choosing the kernel, the most space-filling subset will be employed for validation.

5.2.2. Choosing the most appropriate kernel

Recall that the two kernels which have been discussed in chapter 3 differ solely based on the power of the distance function, one being 1, yielding the so-called Ornstein–Uhlenbeck process, the other one being 2, resulting in the Squared Exponential. Due to the power aforementioned, the latter is more appropriate for smooth functions than the former. However, with no prior knowledge on the nature of the underlying constraint function, a kernel comparison is needed.

Now that the initial dataset for a given constraint, consisting of the analyses which did converge, has been split in validation and training subsets, two models are going to be built, one for each kernel. After hyperparameter optimization, error measures can be computed to assess the quality of each surrogate. The most popular ones are the root mean square error (RMSE), and the k-fold cross-validation error. The latter is assumed by Viana et al., 2009 to be directly proportional to the former, and recommends choosing the number of folds such that each fold contains roughly 5 points. Moreover, as previously discussed, the RMSE computation requires a validation dataset, while the cross-validation does not. Their computations are done in the following manner.

k-fold cross-validation error

The n-point dataset is divided into k subsets, or folds, each fold containing $m_i, i \in \overline{1,k}$ data points; the number of points should be approximately the same for all the folds. One of the subsets is retained, and the surrogate model is trained on the remaining $k-1$ subsets. The mean square error is calculated based on the test or left-out subset, and the process is repeated for all the folds. If the approximation obtained by leaving out the i-th fold is denoted by $\hat{f}^{(-i)}$, the k-fold cross-validation error is then:

$$PRESS_k = \frac{1}{k}\sum_{i=1}^{k}\frac{1}{m_i}\sum_{j=1}^{m_i}\left(y^{(j)} - \hat{f}^{(-i)}(\boldsymbol{x}^{(j)})\right)^2 \tag{5.1}$$

Root mean square error

In the context of surrogate modelling, the root mean square error is approximated by choosing n_v validation or test points in the design space, which are different from the training points. The RMSE is then computed by:

$$RMSE = \frac{1}{n_v}\sqrt{\sum_{i=1}^{n_v}\left(y(\boldsymbol{x}_v^{(i)}) - \hat{f}(\boldsymbol{x}_v^{(i)})\right)^2} \tag{5.2}$$

In the expression above, $\boldsymbol{x}_v^{(i)}$ represents the i$^{\text{th}}$ validation point, and $y(\boldsymbol{x}_v^{(i)})$, $\hat{f}(\boldsymbol{x}_v^{(i)})$ denote the true and predicted values, respectively.

In the following chapter, though, it will be seen that the cross-validation error is an underestimator for both constraints. Moreover, more often than not, it does not even follow the trend of the RMSE, i.e. the two may not both decrease or increase at the same time. Therefore, it is advised to use the RMSE for choosing the most appropriate kernel out of the two. After the right model is obtained, additional data points are created according to the highest Kriging variance predicted by the model, in order to supplement the missing data.

5.3. Variable ranking

Now that the optimal kernel for each surrogate model has been chosen, it is time to perform variable screening, as outlined back in section 3.4. It is proposed to retrain the surrogate model on the whole dataset, i.e. to include the validation points as well in the training process. Indeed, as no further points will be added to the dataset to perform ANOVA, the overall accuracy of the model will benefit by the addition of the validation dataset. The ANOVA procedure has been implemented in Python using the SMT toolbox. The reason behind not choosing Trieste for this step is mainly due to the author having already implemented the procedure in the SMT package during the initial literature study of the present thesis, before having switched to Trieste for the Bayesian optimization procedure. Supposing that all main and interaction effects have been computed, the following methods are proposed to judge which variables should be included in the model.

5.3.1. Variance coverage as a function of included variables

The first proposed method is to compute the total variance coverage as a function of included variables according to the following steps:

1. The variable with the most significant main effect is first included in the model. The variance coverage will be given solely by its main effect.
2. The next variable is added such that the sum of the variance coverage with its main effects and the interaction effects with all variables already included in the model is maximized.
3. Step 2 is repeated until all variables are included in the model.

These variance coverages can then be presented easily using barplots, for instance. Not only do such graphs show the total variance coverage of main and interaction effects, which could also hint that there are other significant higher order interactions between the variables, but they also present quantitatively the gain with each sequential addition of a variable.

5.3.2. Computing error measures as a function of included variables

With the sets of significant variables acquired earlier, it is also possible to build surrogate models in each reduced dimension. It is proposed to use the split dataset once again in order to assess the root mean square error of each model using the validation points. This step can prove especially useful if there are no clear-cut jumps in variance coverage with the addition of more variables.

5.3.3. Plotting the marginal effects

Forrester et al., 2008, p. 10 propose to also conduct a visual screening procedure by plotting the interaction effects of any two variables. After fitting a surrogate model in the whole design space, all variables are kept at a chosen baseline value except for the two whose interaction is to be studied. However, it may not be always obvious which baseline value to choose, and some important trends may be misidentified due to a wrong choice.

It is then proposed to plot instead the marginal effect used for the variable screening procedure, which was derived back in equation 3.35, reiterated in the following:

$$\hat{f}_{\boldsymbol{e}}(\boldsymbol{x}_{\boldsymbol{e}}) = E[\hat{f}|X_i, i \in \boldsymbol{e}] = \int_{\bigotimes_{j\notin \boldsymbol{e}} \chi_j} \hat{f}(\boldsymbol{x}_{\boldsymbol{e}}, \boldsymbol{x}_{-\boldsymbol{e}}) \prod_{j\notin \boldsymbol{e}} w_j(x_j) dx_j \tag{5.3}$$

Indeed, as the marginal effect averages out the contribution of all other variables, it should paint a better picture of the interaction between the two variables. These plots, which could also be augmented in three dimensions in order to showcase an overall variance of the Kriging predictor as well by integrating out the variance instead of the predictor in the equation above, could provide additional help when deciding which variables are significant, especially since this step probably requires the most user input throughout the whole methodology.

5.4. Optimization procedure

The constrained expected improvement approach presented in section 3.5 is going to be used as an acquisition function during the Bayesian optimization procedure, which will be reiterated here for convenience:

$$E[I(\boldsymbol{x})]_{constrained} = E[I(\boldsymbol{x})] \prod_{i=1}^{N_c} P[c_i(\boldsymbol{x}) \leq 0] \tag{5.4}$$

However, it may be the case that the objective and constraint functions are defined on different reduced design spaces. Let $\mathcal{D}$ be the reunion of all design spaces:

$$\mathcal{D} = \mathcal{D}_0 \cup \mathcal{D}_1 \cup \cdots \cup \mathcal{D}_{N_c} \tag{5.5}$$

In the equation above, $\mathcal{D}_0$ denotes the design space of the objective function, while $\mathcal{D}_i, \forall i \in \overline{1, N_c}$ represents the domain of the i$^{\text{th}}$ constraint. It should be noted that $\mathcal{D}$ can differ from $\mathbb{R}^k$, if the reunion of the reduced spaces is different from the original design space. It is proposed that all variables in $\mathbb{R}^k - \mathcal{D}$ should be kept at their minimum values in order to minimize the objective function, which is the weight of the wing in the present study, as they have been deemed as insignificant by the variable ranking procedures corresponding to each constraint function.

Let $Pr_i(\boldsymbol{x}) : \mathbb{R}^k \rightarrow \mathcal{D}_i$ denote the projection of a point in the original design space to the reduced design space, which is simply a vector containing all the significant coordinates of the reduced domain. As all surrogate models are now defined on different domains, the acquisition function to be maximized becomes:

$$\begin{aligned} \max_{\boldsymbol{x}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{y_{min} - \hat{f}(Pr_0(\boldsymbol{x}))}{s(Pr_0(\boldsymbol{x}))}} \left(y_{min} - \hat{f}(Pr_0(\boldsymbol{x})) - su\right) \exp\left(-\frac{1}{2}u^2\right) du \cdot \\ \cdot \prod_{i=1}^{N_c} \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi} s_i(Pr_i(\boldsymbol{x}))} \exp\left[-\frac{1}{2}\left(\frac{t - \hat{f}_i(Pr_i(\boldsymbol{x}))}{s_i(Pr_i(\boldsymbol{x}))}\right)^2\right] dt \end{aligned} \tag{5.6}$$

Recall that the weight function can be represented analytically thanks to the choice of design variables, which are the thicknesses of various components of the structure. As a consequence, the wing does not experience any architectural changes, and all components have a constant area. Going back to the specific case treated in the present report, the expected improvement can thus be replaced by a known improvement:

$$E[I(\boldsymbol{x})] \rightarrow w_{min} - w(\boldsymbol{x}) \tag{5.7}$$

It will be assumed that all projections of the reduced design spaces are carried out as recommended, as all projection notations will be dropped for simplicity. In equation 5.7, w_{min} denotes the already observed minimum weight feasible design, while $w(\cdot)$ represents the known weight function. Of course, if the weight function is known beforehand due to the inclusion of variables such as spar position or number of ribs, for instance, the constrained expected improvement will still be used in its original form, dependent on the surrogate model of the objective function. In the case presented, though, the acquisition function to be maximized becomes:

$$[w_{min} - w(\boldsymbol{x})] \cdot P[c_1(\boldsymbol{x}) \leq 0] \cdot P[c_2(\boldsymbol{x}) \leq 0] \tag{5.8}$$

However, troubles may arise if the algorithm points towards a failed design, failed in the sense of a non-converged analysis. This aspect is also covered by Forrester et al., 2008, p. 133, who proposes a method to impute data in order to deter the optimizer towards other points. Adding the failed design to the training set with any value bigger than 0, such that the design would not satisfy the constraint, will increase the predictor and decrease its variance in the surrounding region, thus diverting the optimizer regardless of the imputed value.

However, there is the risk of the model achieving a poor accuracy due to the imputed values affecting the hyperparameter optimization, or becoming less smooth in regions of converged designs. The solution proposed by Forrester et al., 2008 aims to counteract this deficiency in the following manner:

1. Train the surrogate model on all data points whose analyses converged.
2. For each failed design $\boldsymbol{x}$, impute the observation given by $\hat{f}(\boldsymbol{x}) + s^2(\boldsymbol{x})$.
3. Retrain the surrogate model on the whole dataset, including imputed values.
4. Maximize the constrained expected improvement function.
5. Repeat steps 1-4 until a suitable convergence criterion is achieved for the Bayesian optimization procedure.

The use of $\hat{f}(\boldsymbol{x})+s^2(\boldsymbol{x})$, where $\hat{f}$ and s^2 are the Kriging predictor and variance of one of the constraints, ensures that the imputed data does not change the smoothness of the surrogate model, but only deters the optimizer from the failed designs. Indeed, if the failed data point $\boldsymbol{x}^f$ is very close to a training point $\boldsymbol{x}^{(i)}$, i.e. $||\boldsymbol{x}^f - \boldsymbol{x}^{(i)}|| \to 0$, then the variance becomes extremely low, $s^2(\boldsymbol{x}^f) \to 0$, and, as the predictor interpolates the data, $\hat{f}(\boldsymbol{x}^f)+s^2(\boldsymbol{x}^f) \to y_i$, where y_i is the observation at the training point. Conversely, if the variance is high, the imputed value would also be high, and the optimizer would not attempt sampling in that region due to an increased probability of the constraint function not being satisfied.

The last aspect which should be discussed is the convergence criterion of the Bayesian optimization procedure. Of course, in a realistic setting, the most decisive factor is time itself. However, computational resources and time aside, a simple convergence criterion can be formulated, where the optimization procedure stops when the acquisition function is relatively constant over a number of iterations.

During the preliminary phases of the study, though, as the objective function is the weight in the present study, a threshold of the improvement was attempted to be defined. For instance, if a threshold of 100 grams is defined, the optimization can be stopped once the improvement $w_{min} - w(\boldsymbol{x})$ is lower than this threshold value. However, it may be the case that the optimization algorithm exploits a region for a number of iterations, then jumps to another area to explore it, thus experiencing a surge in improvement, although the probabilities of feasibility are consistently low.

Therefore, it is proposed to stop the procedure once the constrained expected improvement is lower than the threshold for a number of consecutive iterations. It is advised to also export the improvement and probabilities of feasibility at each iteration, in order to gain insight on whether the algorithm is explorative or exploitative.

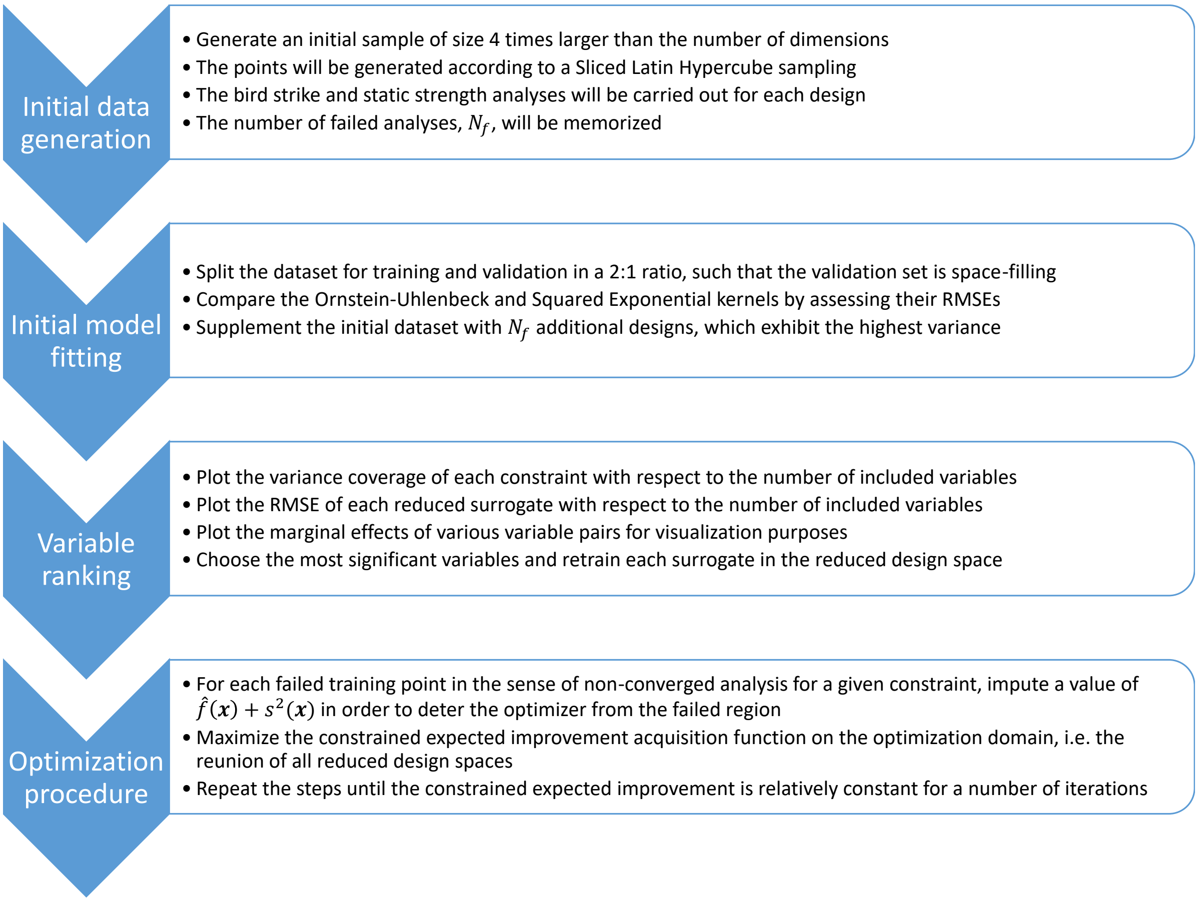
5.5. Summary of the proposed steps

To summarize the proposed methodology for the given case study, figure 5.2 illustrates a flowchart to visualize the sequential steps to be undertaken, their results being presented in the next chapter. Firstly, an initial Sliced Latin Hypercube design consisting of $4k$ samples is generated, where k is the number of dimensions. The bird strike and static strength analyses will be carried out for each design, and additional checks are going to be made to ensure the accuracy of the simulations. If the artificial strain energy is larger than 5% of the total strain energy of the structure, the bird strike analysis is deemed as failed. Similarly, if the static simulation does not converge or is overstabilized, the static analysis is also deemed as failed. The total number of failed analyses will be memorized for each constraint function, and will be denoted by N_{f1} and N_{f2}.

For each constraint function, the initial dataset is split for training and validation in a 2:1 ratio, such that the validation dataset is space-filling. In order to ensure this aspect, it is proposed to generate a number of splits and to assess the discrepancy of each validation dataset, and then choose the most space-filling one. Afterwards, two kernels are going to be compared against each other, based on the root mean square error. The initial model fitting is finalized by supplementing the initial dataset of each constraint function with its corresponding N_f additional designs.

The variable ranking procedure follows, which is probably the step which requires the most user input. Various methods are proposed to choose the most significant variables. Variance coverage and error metrics can be plotted against an increasing number of variables, being quantitative indicators of the quality of the reduced design space, although plotting the marginal effects among various variables can also help to judge whether the contributions were estimated correctly. After choosing the most significant variables, the surrogate model corresponding to each constraint function is retrained on its reduced design space.

The Bayesian optimization procedure is the final step, and is carried out over a number of iterations until the value of the constrained expected improvement function stabilizes, i.e. no surges are present over a number of optimization steps. At each iteration, values are imputed at each design where the simulations failed, such that the optimizer is deterred from its vicinity. The constrained expected improvement is then maximized in order to find the next infill points. Finally, when convergence is reached, the optimal design is chosen such that the objective function is minimized, and feasibility is ensured against both constraints.



Initial data generation

- Generate an initial sample of size 4 times larger than the number of dimensions
- The points will be generated according to a Sliced Latin Hypercube sampling
- The bird strike and static strength analyses will be carried out for each design
- The number of failed analyses, N_f, will be memorized

Initial model fitting

- Split the dataset for training and validation in a 2:1 ratio, such that the validation set is space-filling
- Compare the Ornstein-Uhlenbeck and Squared Exponential kernels by assessing their RMSEs
- Supplement the initial dataset with N_f additional designs, which exhibit the highest variance

Variable ranking

- Plot the variance coverage of each constraint with respect to the number of included variables
- Plot the RMSE of each reduced surrogate with respect to the number of included variables
- Plot the marginal effects of various variable pairs for visualization purposes
- Choose the most significant variables and retrain each surrogate in the reduced design space

Optimization procedure

- For each failed training point in the sense of non-converged analysis for a given constraint, impute a value of $\hat{f}(\boldsymbol{x}) + s^2(\boldsymbol{x})$ in order to deter the optimizer from the failed region
- Maximize the constrained expected improvement acquisition function on the optimization domain, i.e. the reunion of all reduced design spaces
- Repeat the steps until the constrained expected improvement is relatively constant for a number of iterations

Figure 5.2: Flowchart of the proposed methodology, applied to the case study problem.

6

Case study results

In the present chapter, the proposed methodology will be applied to the case study problem suggested back in chapter 4. The initial data generation will be discussed, along with any issues which have arisen, mainly due to failed analyses. This data will be then split into two sets, for training and validation purposes. Afterwards, a comparison between two kernels will be made, and an initial surrogate model will be fitted. The variable ranking will be carried out, followed finally by the optimization procedure in the reduced design space.

As mentioned in section 4.1, the thicknesses of various parts of the wing will act as design variables. However, realistically, metal sheets are available in a limited set of thicknesses. As a consequence, the design space will be treated as discrete, with the thickness values having a step of $0.05\ mm$. Nevertheless, as outlined in chapter 3, integer Kriging models rely on the continuous formulation as well, although a different covariance is proposed by Garrido-Merchán and Hernández-Lobato, 2020.

For simplification, though, the continuous Kriging model will be employed in the present study, in order to conduct the variable screening procedure as exposed in section 3.4, on the continuous Kriging kernel. Employing discrete variables in this manner is assumed to influence only the Bayesian optimization process, as the infill points suggested at each iteration will be rounded to the nearest discrete design point. However, due to the rather small step of $0.05\ mm$, each component of a design variable is going to be at most $0.05/2 = 0.025\ mm$ away of the nearest discrete value. As will become apparent later during the variable screening procedure, the optimization space will consist of 8 dimensions, therefore an infill point suggested by the continuous formulation will be at most $\sqrt{0.025^2 \cdot 8} \approx 0.07\ mm$ away from the nearest discrete value, which is assumed not to be a sufficiently large distance to impact negatively the outcome of the solution.

6.1. Initial model generation

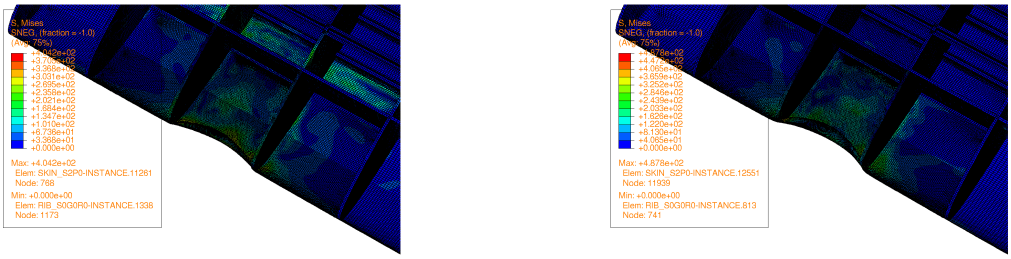
The initial model serves as the foundation for the variable ranking procedure presented back in section 3.4. As discussed in chapter 5, the initial dataset will be split for verification purposes, in order to choose the most appropriate kernel.

6.1.1. Initial data generation

A 76-sample Latin Hypercube was generated with the R package developed by Ba, 2015. The resulting data points were rounded up to the nearest discrete value, and the analysis files were written. All simulations were carried out on the university cluster, and the amount of employed CPUs varied with the available number of nodes and Abaqus licenses. To give an idea about the computational intensity of these analyses, a single workflow took, on average, 10 hours on 32 CPUs. Of course, more analyses could be run in parallel, depending once again on the availability of resources.

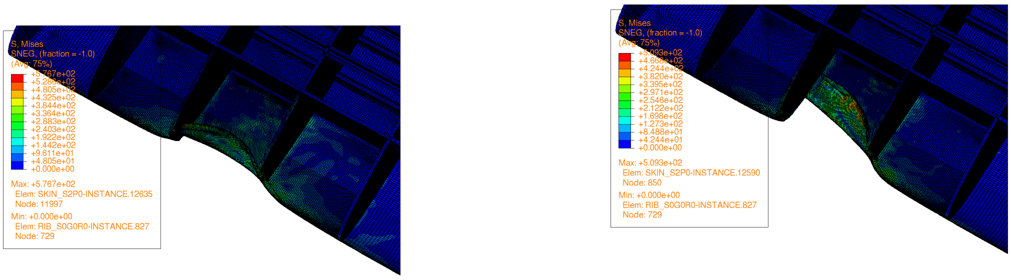
Due to the rather large design space, a variety of failure modes can be observed, some of them being included in figure 6.1. A small amount of penetration can be observed in figure 6.1a, where the nose ribs are virtually intact. In figure 6.1b, though, the nose rib flanges have started deforming a bit, whereas in figure 6.1c, severe nose rib crushing occurs. In both figure 6.1d and 6.1e the leading edge skin has completely ruptured in the vicinity of the left-central nose rib, the latter case also showing a significant plastic deformation of the front spar. Both of the last cases occur with rather thick nose ribs. Indeed, more compliant ribs can accommodate a larger penetration depth without causing leading edge skin rupture.

All in all, this variety of failure modes supports the exploration of the design space with surrogate models. What would also be interesting, for a future study, would be to build a classification surrogate to predict different failure modes.



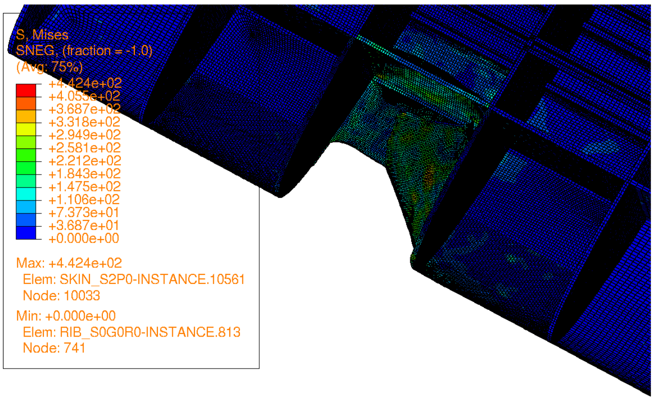
(a) Small penetration and negligible nose rib distortion.

(b) Small penetration and small nose rib distortion.



(c) Medium penetration and severe nose rib crushing.

(d) Large penetration, leading edge skin rupture and severe nose rib crushing.



(e) Critical penetration, leading edge skin rupture and negligible nose rib distortion. Front spar affected, strong indentation.

Figure 6.1: Various failure modes throughout the design space.

6.1.2. Initial model fitting and additional data points

Out of the 76 initial data points, though, 2 bird strike analyses failed due to high artificial strain energies, and, on top of those, 9 static analyses did not converge. Therefore, 11 more points are needed. However, instead of generating them without any input from a surrogate model, the initial 74-point set for depth and 67-point set for static strength will be split to create the 25-point validation datasets. This partition is carried out as described in chapter 5, where 10^6 subsets are randomly produced, and the one with the smallest discrepancy is chosen for validation. The following results are obtained for the two models:

Constraint	Static Strength	Maximum penetration depth
Ornstein–Uhlenbeck MSE	$2694.787\ MPa^2$	$15623.898\ mm^2$
Squared Exponential MSE	$2667.602\ MPa^2$	$15623.556\ mm^2$
Ornstein–Uhlenbeck PRESS	$1529.833\ MPa^2$	$4629.445\ mm^2$
Squared Exponential PRESS	$1517.588\ MPa^2$	$4629.335\ mm^2$

Table 6.1: Error measures for the two kernels, for each constraint function.

The Squared Exponential kernel seems to give the best results for both constraints, although the results are quite close, especially for the maximum penetration depth constraint. This may be an indicator that the power of the distance function, i.e. p_l in equation 3.9, should not be fixed. Indeed, this functionality is available in the SMT toolbox, but not in Trieste. Moreover, the ANOVA implementation developed by the author only supports these two kernels. Therefore, the squared exponential correlation function will be fixed for both constraints. The 15-fold cross-validation error, in this case, predicts the same result.

It is now time to add the remaining training points. Although there were rather many failed static analyses, failed in the sense of unconverged solutions, the MSE of the model is not high at all. A root mean square error of less than 52 MPa represents less than 15% of the yield strength of the spar material, which can be seen as acceptable, given that the constraint enforces an 80% ratio between maximum stress and yield strength. On the other hand, the RMSE of the depth model is nearly 125 mm, which is more than half of the maximum acceptable value. Therefore, as the dynamic and static analyses are done sequentially, it has been decided to employ the depth surrogate model to search for additional points with the highest variance.

Figure 6.2 shows the evolution of the mean square and 15-fold cross-validation for the depth model built on the initial training set. The PRESS error severely underestimates the RMSE, while also predicting a much more aggressive increase in global accuracy. In reality, though, the RMSE did not alleviate much. Therefore, the idea that validation points are still needed is reiterated.

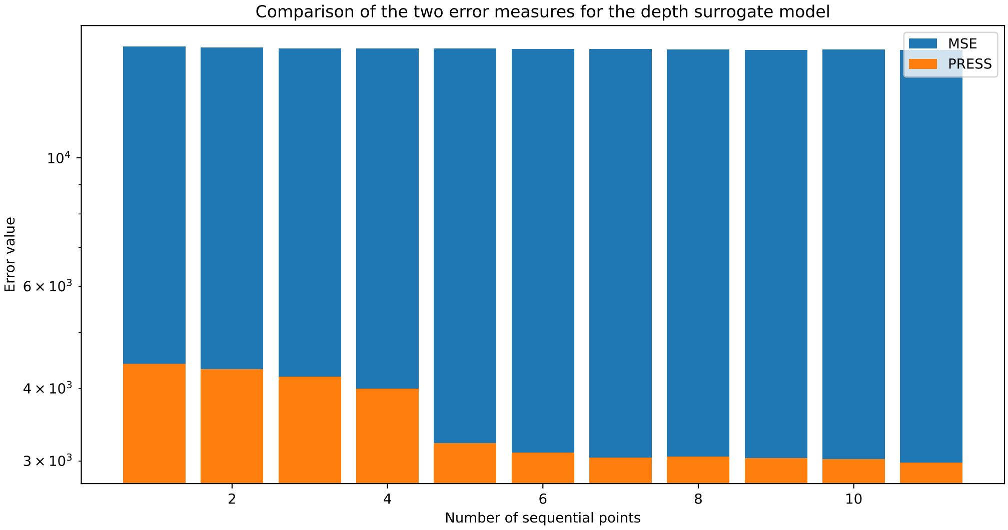


Figure 6.2: The mean square and cross validation errors of the depth surrogate model versus the number of additional points.

6.2. Variable ranking

The next step is to conduct the variable screening procedure. As outlined in the last chapter, the training dataset is now augmented with the validation points to provide a greater accuracy of the initial model.

6.2.1. Spar residual strength constraint

Although the initial model for this constraint showed the least RMSE, indicating a good fit, the number of significant variables is quite high, as will be presented in the following. First and foremost, figure

6.8 presents all main and interaction effects as a percentage of the total variance of the model. The main effects are situated on the diagonal of the graph, while all interaction effects are in the lower-right part of the plot. All effects with a value of less than 0.1% have been masked for simplicity. Even if the leading edge skin and nose ribs have not been included in the Abaqus model of the static analysis, the variable screening method identified them as being the most important variables, their interaction also being deemed important.

The wingbox rib is also judged as important, although it does not show any interaction with another variable. Along with the other components of the wingbox, such as the stringers and skin panels, they probably were identified as significant due to their contribution to the torsional stiffness of the wingbox, as the torsion moment has a rather large magnitude. This explanation will be confirmed when plotting the marginal effect of the wingbox rib shortly.

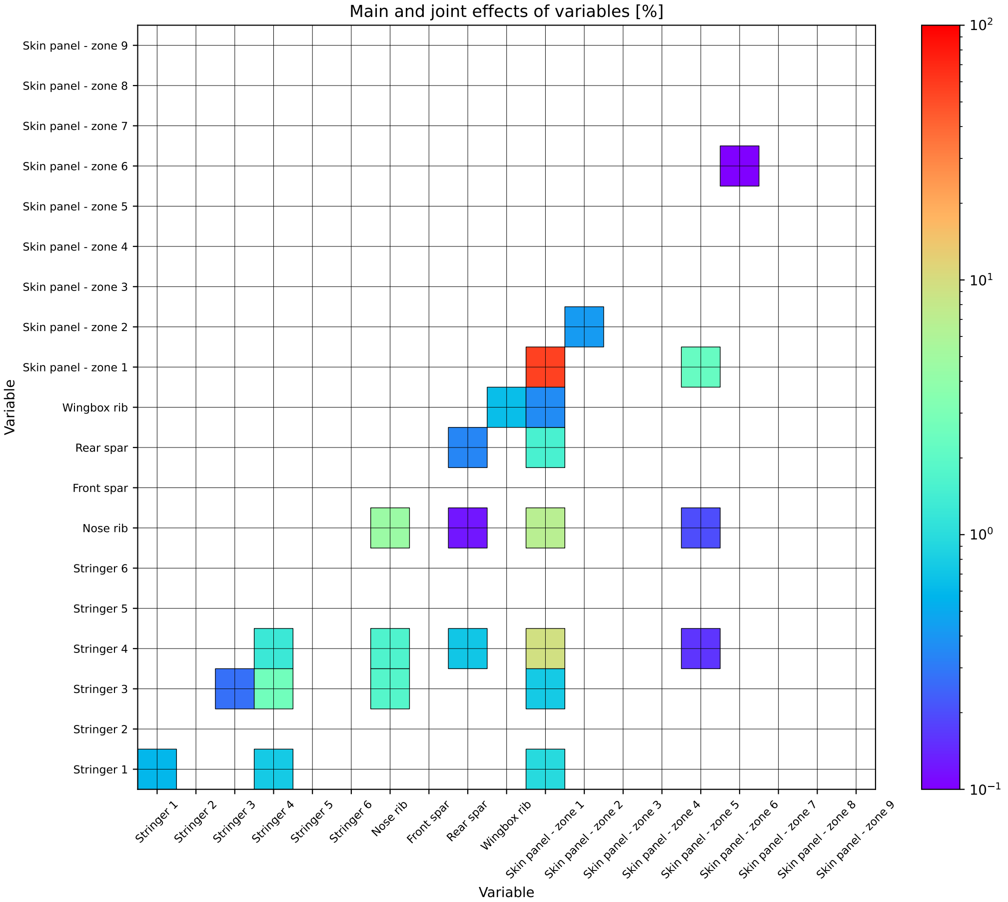


Figure 6.3: Main and interaction effects for the maximum maximum Mises stress constraint on the front spar. Note: all effects with a value of less than 0.1% have been masked.

Moreover, it is curious that the front spar does not contribute significantly towards the response, although the maximum Mises stress is taken from it. However, this phenomenon could also be explained by the choice of bending loads, which are probably not high enough to cause a significant increase in the stress of the damaged spar. Indeed, the maximum stresses encountered before and after load introduction seem to be heavily correlated, as presented in figure 6.4. The Pearson correlation coefficient between the two results has been computed as well, yielding a value of 0.90, further confirming the relationship between the two quantities. The computed R-squared, which is a statistical measure indicating how much the variation of a model is explained by a variable, yields a value of 0.82, proving further that there is a linear relationship between the stresses before and after load introduction in the front spar. Of course, as both correlation coefficients are not very close to 100%, it is acknowledged the stress after load introduction is not explained solely by the one before, which explains why the other design variables have been found as significant.

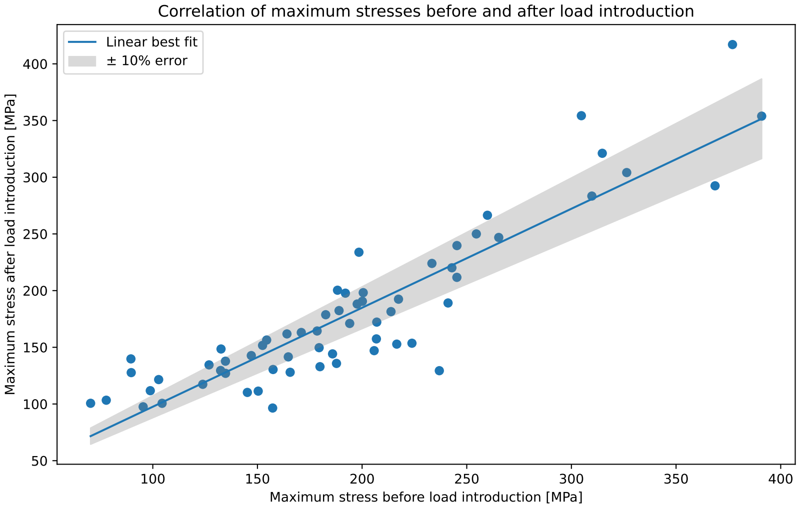


Figure 6.4: Scatter plot between the maximum stress before and after load introduction, along with the best linear predictor and a ± 10% error area.

Identifying the most significant variables

Following the recommendation in the last chapter, a bar plot showcasing the evolution of the variance coverage with respect to the number of included variables is presented in figure 6.5. Firstly, it is apparent that higher-order interactions take place, given the less than 95% coverage of all included variables. Secondly, from 8 variables onwards, the benefit of introducing more dimensions in terms of variance seems to be less than 1%.

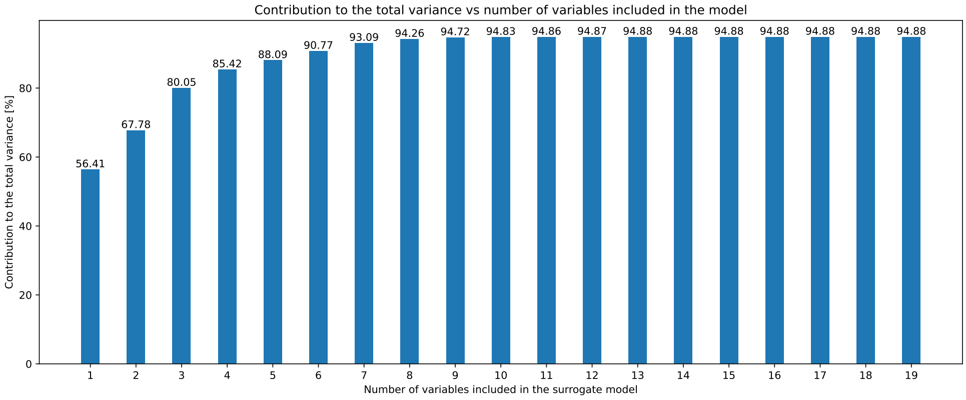


Figure 6.5: Variance coverage versus number of included variables for the maximum stress constraint.

Another aspect which may aid in decision making would be refitting surrogate models in the reduced design space on the initial training dataset, and evaluating their MSE and PRESS. Firstly, it should be noted that the decrease in MSE is owed to the addition of training points in lieu of the failed analyses. Moreover, the cross-validation error underestimates the MSE by a rather high margin, and it also does not follow the same increase and decrease trends.

It is curious, though, that both MSE and PRESS show a sudden decrease at the 5-variable mark. However, as only 88% of the total variance is covered with this number of dimensions, it will be dismissed, as the variance coverage is not sufficient. After 8 dimensions, though, both errors show a constant trend, with a slight exception of the MSE with 14 variables. It should be reiterated, though, that the MSE itself is an estimation of the global accuracy, and it should be taken with a grain of salt. Such sudden decreases may be solely artifacts caused by the measure itself and a lucky hyperparameter optimization in that specific reduced design space. However, it is still encouraging that the error measures of the 8-dimensional surrogate are not much higher compared to the 19-variable model, therefore, even if higher-order interactions are present, the reduced metamodel is not expected to be severely inaccurate.

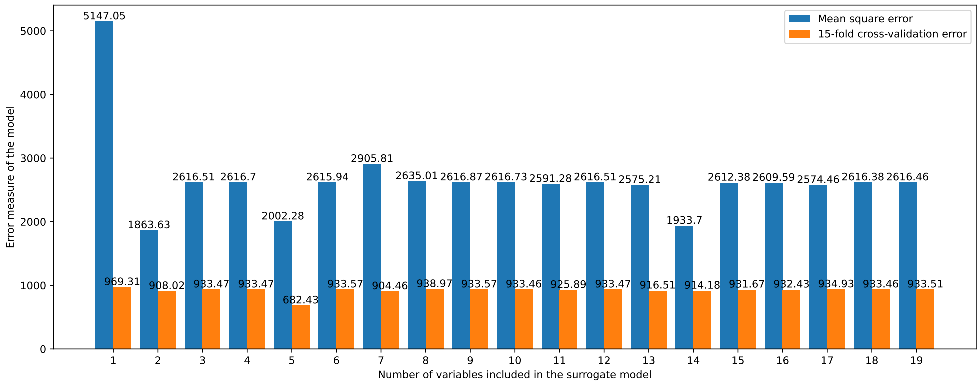


Figure 6.6: The root mean and 15-fold cross validation errors for reduced-space surrogates.

Based on the observations aforementioned, 8 significant variables have been identified, namely the leading edge skin, nose and wingbox ribs, the first, third and fourth stringer stations, the fifth material zone of the skin panel (i.e. between the second and third stringers), and the rear spar. It is acknowledged that the exclusion of the second stringer or other skin panel material zones may seem odd, and that the inclusion of higher-order effects may have revealed a complex interaction with other variables thus rendering it significant as well, exposing a weakness of the present methodology.

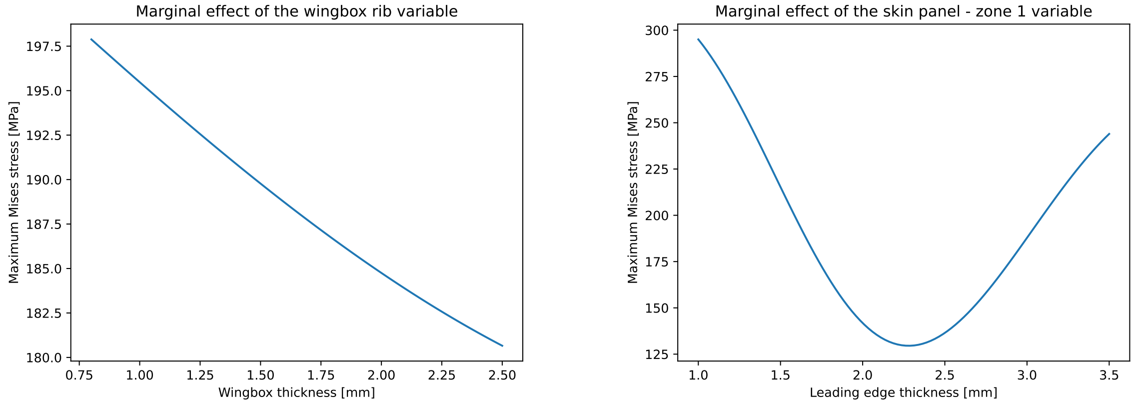
The main effects of the significant variables are given in table 6.2. It is clear that only the leading edge skin, followed from a rather big distance by the nose rib, have sizeable main effects. The interaction among all components is therefore quite notable, therefore it is advised that such static tests use a higher-order ANOVA formulation, which has not been implemented in the present work. The rear spar has a more significant effect than the front spar probably due to the bending force point of application, which is further away from the rear spar, thus creating an additional torsion moment.

Variable	Main effect [%]	Variable	Main effect [%]
Skin panel - zone 1	56.41	Skin panel - zone 5	0.06
Nose rib	4.5	Rear spar	0.34
Stringer 4	1.26	Stringer 1	0.58
Stringer 3	0.28	Wingbox rib	0.63

Table 6.2: Main effects of significant variables.

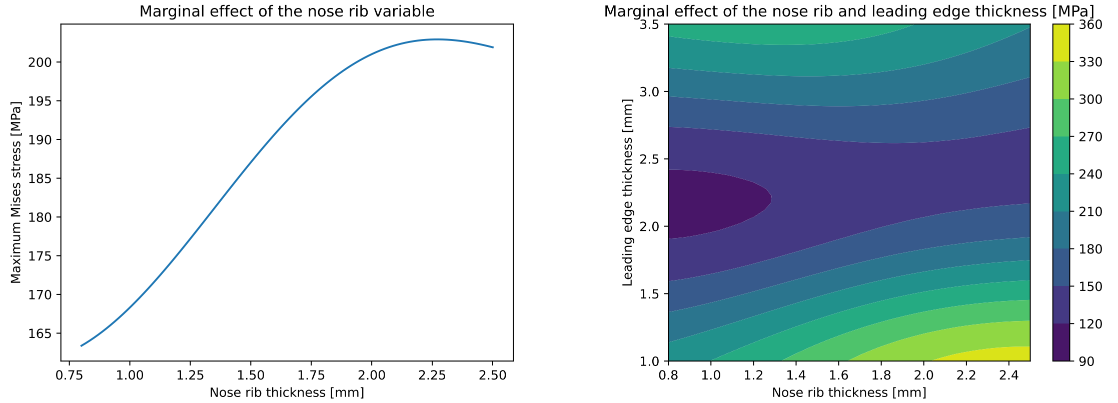
Illustrating the marginal effects

Because a rather large number of significant variables have been identified, only some of the marginal effects will be illustrated. It was mentioned in the last section that the wingbox rib may have been estimated as being significant thanks to its contribution to the torsional stiffness, which is now confirmed by the linear, decreasing relationship depicted in figure 6.7a. The marginal effect of the most significant variable, the leading edge skin, is depicted in figure 6.7b. It seems to have a parabolic shape, with its minimum lying approximately in the middle of the range. Indeed, a very small thickness may lead to a significant amount of penetration, where the skin may get ruptured, allowing for the spar itself being impacted. On the other hand, an excessively thick skin may be too rigid, therefore having reduced energy absorption capabilities, and the kinetic energy of the bird being transferred to the rest of the structure, spar included.



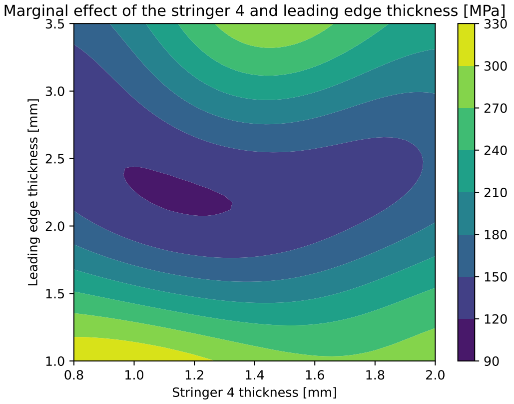
(a) Marginal effect of the wingbox rib.

(b) Marginal effect of the leading edge skin.



(c) Marginal effect of the nose rib.

(d) Marginal joint effect of the leading edge skin and nose rib.



(e) Marginal joint effect of the leading edge skin and stringer 4.

Figure 6.7: Various marginal effects for the maximum Mises constraint.

Figure 6.7c presents the marginal effect of the nose rib. A first observation is that the range of the effect is much smaller than the leading edge skin, with 35 and nearly 175 MPa, respectively, thus confirming the large difference of variance coverage. Moreover, the main effect of the nose rib seems to be linear. As previously stated, a more compliant nose rib is desired, as it absorbs more energy through plastic deformation, while also allowing a larger penetration of the leading edge without rupture. Rigid nose ribs would also transfer energy directly to the spar behind them, due to reduced energy absorption capabilities, while leading edge skin rupture would lead to bird particles, or even pieces of the skin itself, impacting the spar.

Figure 6.7d shows the interaction between the leading edge and the nose rib, whose effect was estimated at 6.87%. As predicted from both main marginal effects, the highest stress is encountered with a thin leading edge and thick nose rib, as it promotes skin rupture. Conversely, the lowest stress is predicted with the thinnest nose rib, and a leading edge skin of roughly 2.0 - 2.5 mm.

Although there are many effects which could be illustrated, the interaction between the stringer at the fourth station and the leading edge skin will be the last depiction thanks to its rather high joint effect of 9.37%, which is shown in figure 6.7e. Once again, the lowest stress is encountered with a leading edge skin between 2.0 and 2.5 mm, as predicted by its main effect. The stringer probably has a significant contribution due to the bending rigidity it offers against the considered load case.

6.2.2. Maximum penetration depth constraint

Although this constraint has a large RMSE, a very small number of significant variables can be identified, as showcased in figure 6.8. The leading edge skin is by far the most notable contributor, having a main effect of 88.60%. It does interact with the nose rib with an estimated effect of 1.35%, although the nose rib itself has a small main effect of only 0.78%. The wingbox rib has a larger main effect of 1.31%, although it does not interact with any other variable. Interestingly, stringer 2 has a main effect of 0.44% and is estimated to interact with the leading edge skin in a proportion of 1.89%, but this contribution could be also misidentified due to a rather poor global accuracy.

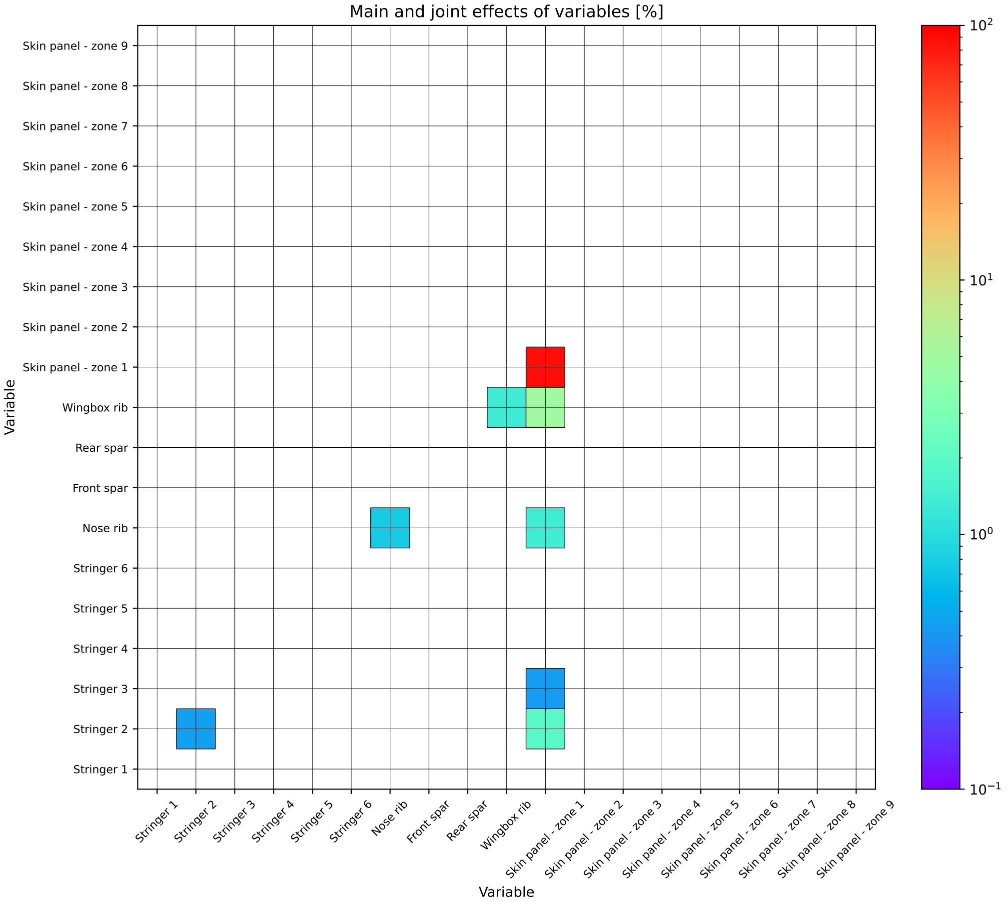
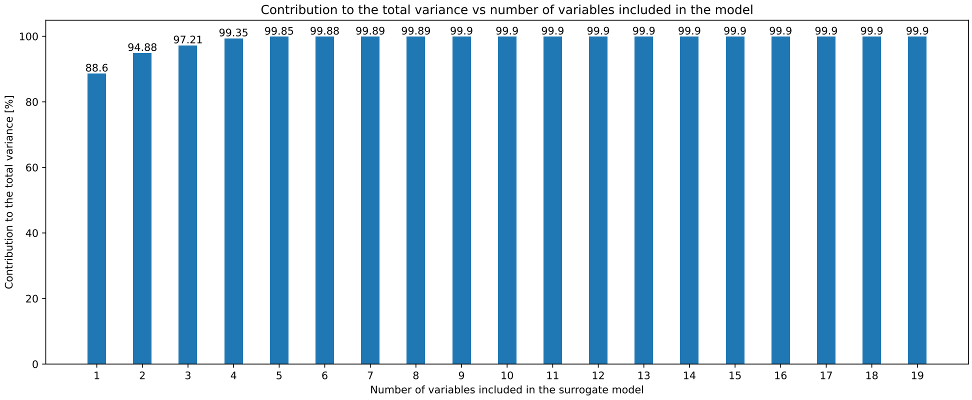


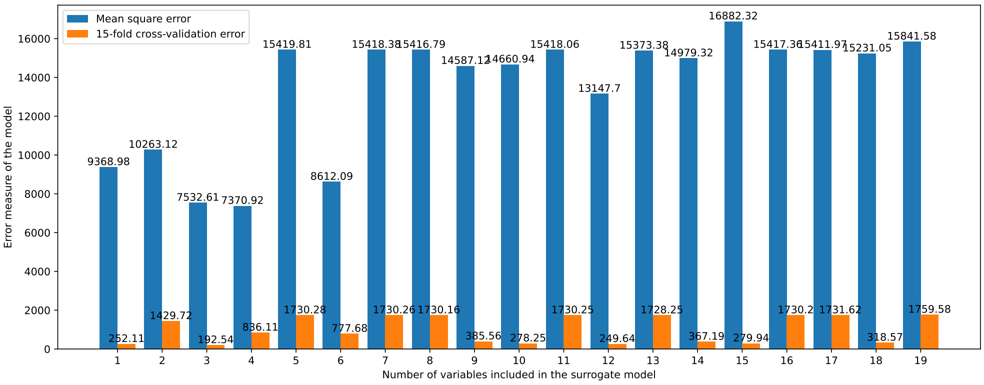
Figure 6.8: Main and interaction effects for the maximum penetration depth constraint.

Identifying the most significant variables

Figures 6.9a and 6.9b present the variance coverage and the mean square and cross-validation errors, respectively. It can be observed that, from 4 variables onwards, there is not much improvement in terms of coverage. However, the three-dimensional design space presents good variance coverage and mean square error reduction, with a remarkable decrease of 80% compared to the full space model. The cross-validation errors severely overestimate the global accuracy of the model, though, and do not follow the same trend as the MSE in higher dimensions.



(a) Variance coverage versus number of included variables for the maximum penetration depth constraint.



(b) Error measures versus number of included variables for the maximum penetration depth constraint.

Figure 6.9: Barplots of variance coverage and error measures with an increasing number of variables for the maximum penetration depth constraint.

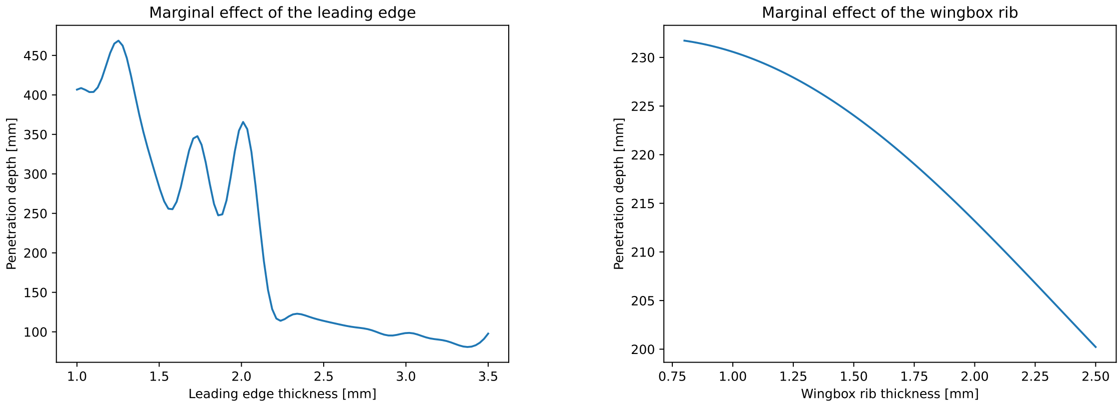
Table 6.3 presents the main and joint effects of the four most significant variables, which are, in the order of the algorithm presented in chapter 5: leading edge skin, wingbox rib, second station stringer, and nose rib.

Variable(s)	Effect [%]
Skin panel - zone 1	88.60
Stringer 2	0.44
Nose rib	0.78
Wingbox rib	1.31
Stringer 2 - skin panel zone 1	1.89
Nose rib - skin panel zone 1	1.35
Wingbox rib - skin panel zone 1	4.97

Table 6.3: Main and joint effects of the first four most significant variables.

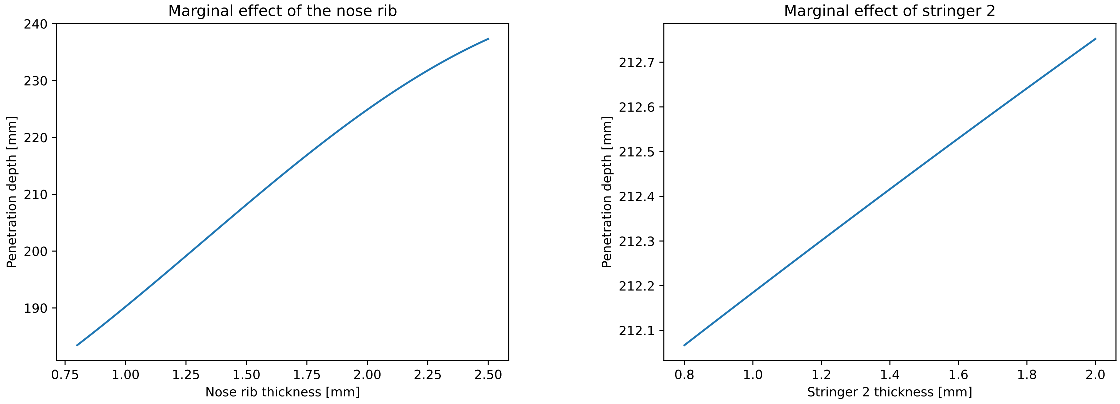
Figure 6.10 showcases some of the marginal effects of the first four most significant variables. While it is clear from figure 6.10a that the leading edge has a significant contribution, a highly multimodal landscape is predicted. On the other hand, the nose and wingbox ribs depicted in figures 6.10c and 6.10b, respectively, exhibit nearly linear trends. The second station stringer does not have a significant main contribution, though, as the depth range is under 1 mm, as shown in 6.10d. However, a strong interaction was estimated with the leading edge, which is illustrated in 6.10g. The bands of constant depths, though, indicate a poor accuracy of the metamodel. Figures 6.10f and 6.10e, which show the interaction of the leading edge with the nose and wingbox ribs, respectively, also suffer from the same shortcome, although a smaller multimodality can be observed.

Based on engineering judgement, the leading edge and nose and wingbox ribs will be chosen as significant variables, the second station stringer being dismissed. After a re-evaluation of the error measures of the design space, the MSE has decreased even further, from 7532.61 to 6269.87 mm^2. However, the cross-validation error shows a different story, as it has actually increased from 192.54 to 643.8 mm^2, thus reiterating the necessity of a validation set.



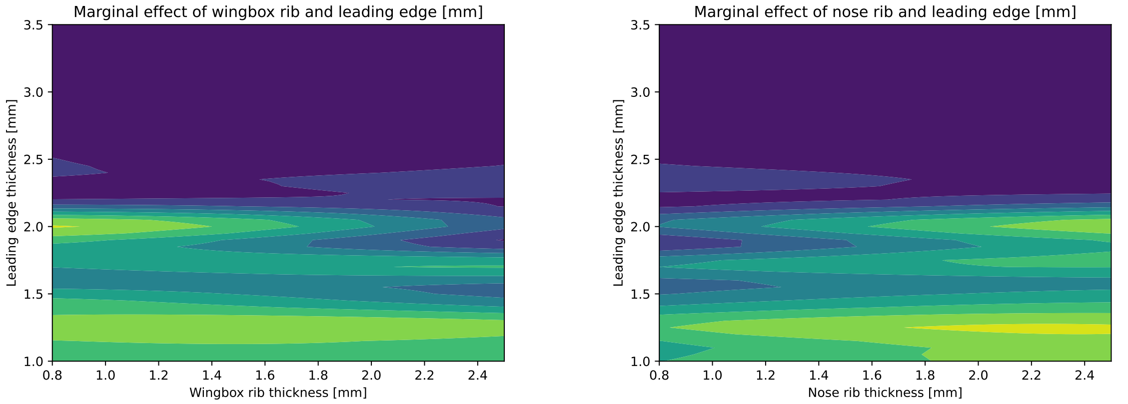
(a) Marginal effect of the leading edge skin.

(b) Marginal effect of the wingbox rib.



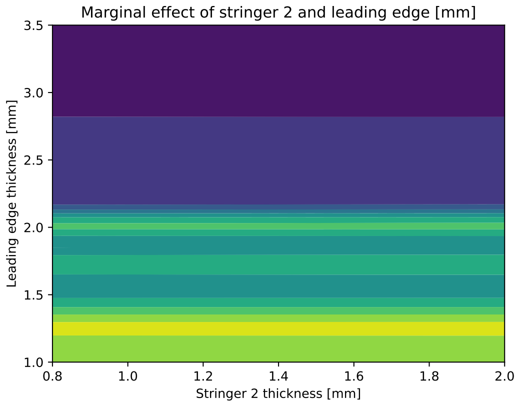
(c) Marginal effect of the nose rib.

(d) Marginal effect of the second station stringer.



(e) Marginal joint effect of the leading edge skin and wingbox rib.

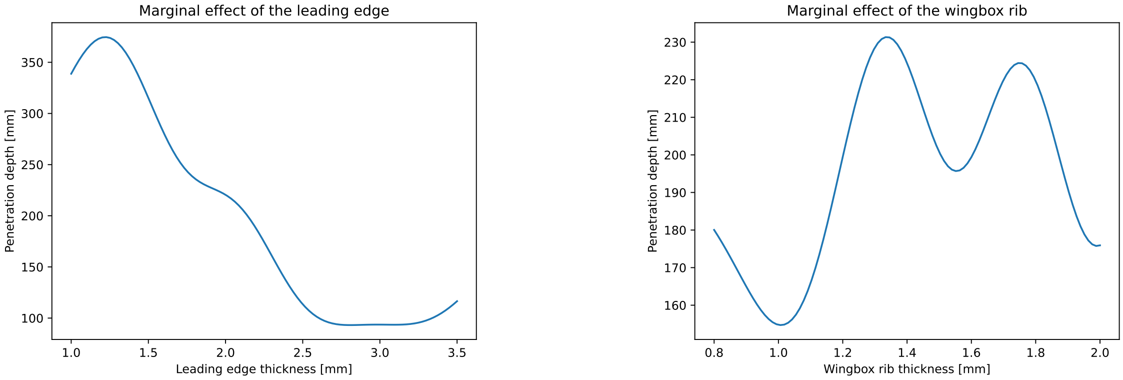
(f) Marginal joint effect of the leading edge skin and nose rib.



(g) Marginal effect of the second station stringer.

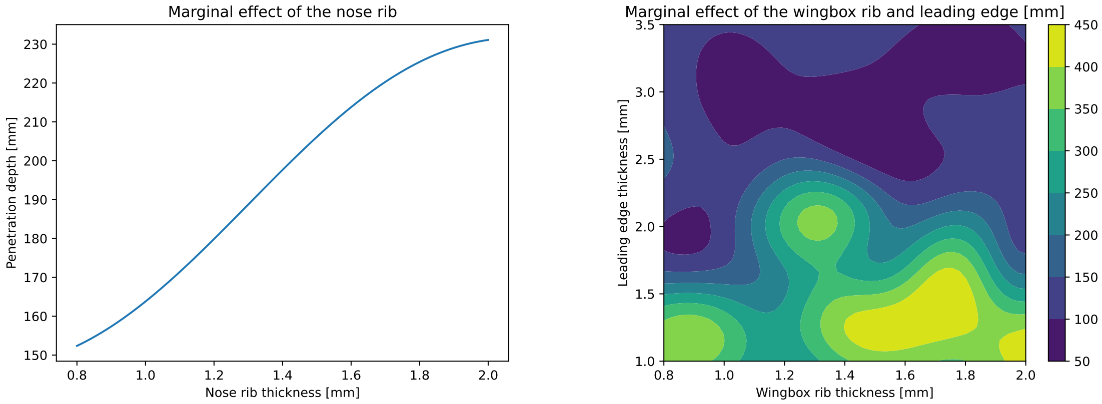
Figure 6.10: Various marginal effects for the maximum penetration depth constraint.

The fact that the variables have been well chosen is also supported by their marginal effects in the reduced design space, which are depicted in figure 6.11. The leading edge skin no longer presents the high modality aforementioned, although the wingbox rib took its place. It seems that, although a rather large variation of penetration depth can be observed in figure 6.11b, the trend is no longer linear. The nose rib, though, has more or less the same marginal effect. The interactions among the variables, though, no longer exhibit any bands of constant values or regions of sharp transitions, showing a better fit of the Kriging θ hyperparameter, which was mentioned back in section 3.3 that it is inversely proportional to the correlation distance between two points.



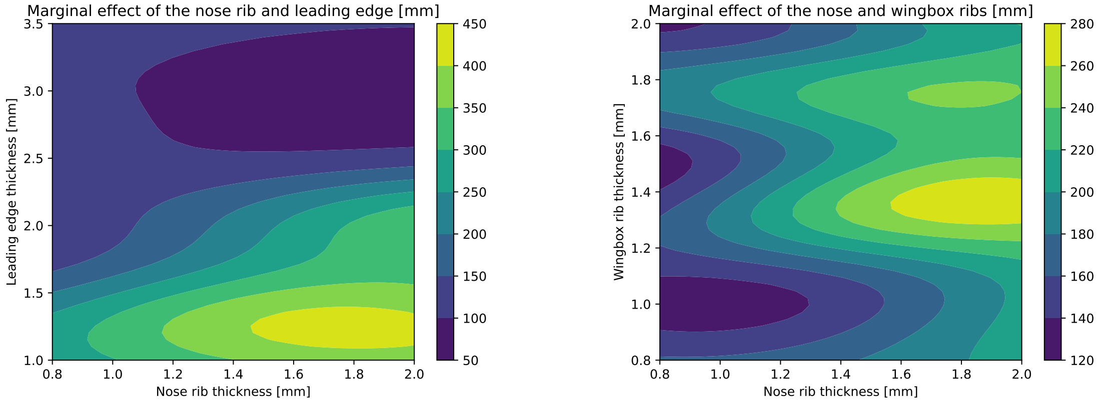
(a) Marginal effect of the leading edge skin.

(b) Marginal effect of the wingbox rib.



(c) Marginal effect of the nose rib.

(d) Marginal joint effect of the leading edge skin and wingbox rib.



(e) Marginal joint effect of the leading edge skin and nose rib.

(f) Marginal effect of the nose and wingbox ribs.

Figure 6.11: Various marginal effects for the maximum penetration depth constraint after design space reduction.

6.2.3. Additional comments

The variable ranking procedure identifies the leading edge skin and the nose rib among the significant variables for both constraints, which may raise the question whether the two constraints are correlated. Figure 6.12 presents a scatterplot of the two observations. However, no correlation seems to take place, and Pearson's correlation coefficient confirms it, having a value of 0.40. As a consequence, both constraints should be taken into account during the optimization process.

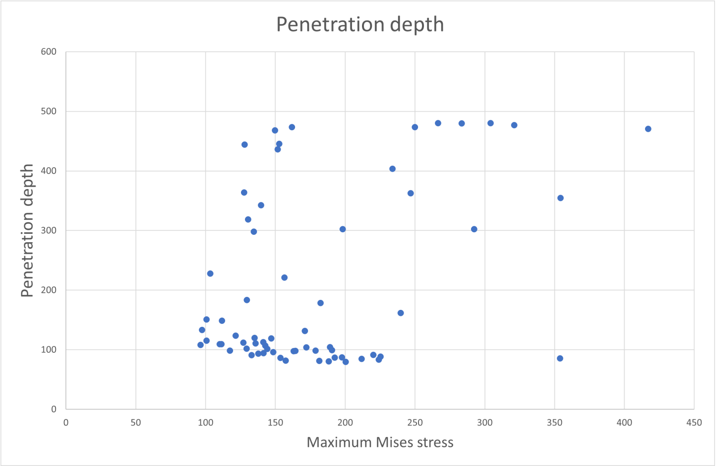


Figure 6.12: Scatterplot of the penetration depth and the maximum Mises stress of the front spar.

The largest design space compaction was shown by the penetration depth constraint which, interestingly, also showed the largest mean square error. The marginal effect plots indicate a highly multimodal and nonlinear landscape, as predicted, and the variable screening procedure helped alleviate the mean square error significantly.

As a side note, Welch et al., 1992 introduced the idea that the $\boldsymbol{\theta}$ hyperparameters can be used to assess the importance of variables for computer experiments, who illustrated the idea on a 20-variable analytical benchmark function. Indeed, as described earlier in section 3.3, the bigger θ is in a direction, the faster the correlation between a point and the training data plummets. Therefore, high values of θ show that the corresponding variable is very active. Although this procedure is very easy to achieve, it is very dependent on the number of samples used for training.

For instance, take the maximum penetration depth constraint. The evolution of the θ hyperparameters for a 19-, 38-, 57- and 76-point training set is depicted in figure 6.13. The values clearly exhibit a large variation with respect to the number of training points. Moreover, it is not obvious at all when to stop, as no quantification of information loss is available, contrary to the variance-based method. As presented in this chapter, there are many visualization methods to aid in the identification of significant variables. Coupled with engineering judgement, as done for the depth constraint, the proposed procedure can also yield significant improvements in the accuracy of a metamodel, on top of a dimensionality reduction of the design space, which greatly speeds up the optimization process.

While the identification of the nose rib and leading edge skin as significant variables was expected, the necessity of including the wingbox rib as well for the depth constraint was not immediately apparent, as the second material zone of the skin panel, which is situated between the front spar and the leading edge, as depicted back in figure 4.2, could have been a larger contributor due to it being closer to the impact location. Moreover, for the maximum Mises stress, the exclusion of the front spar as a significant variable would not have been possible without the screening procedure, as its exclusion is rather counter-intuitive.

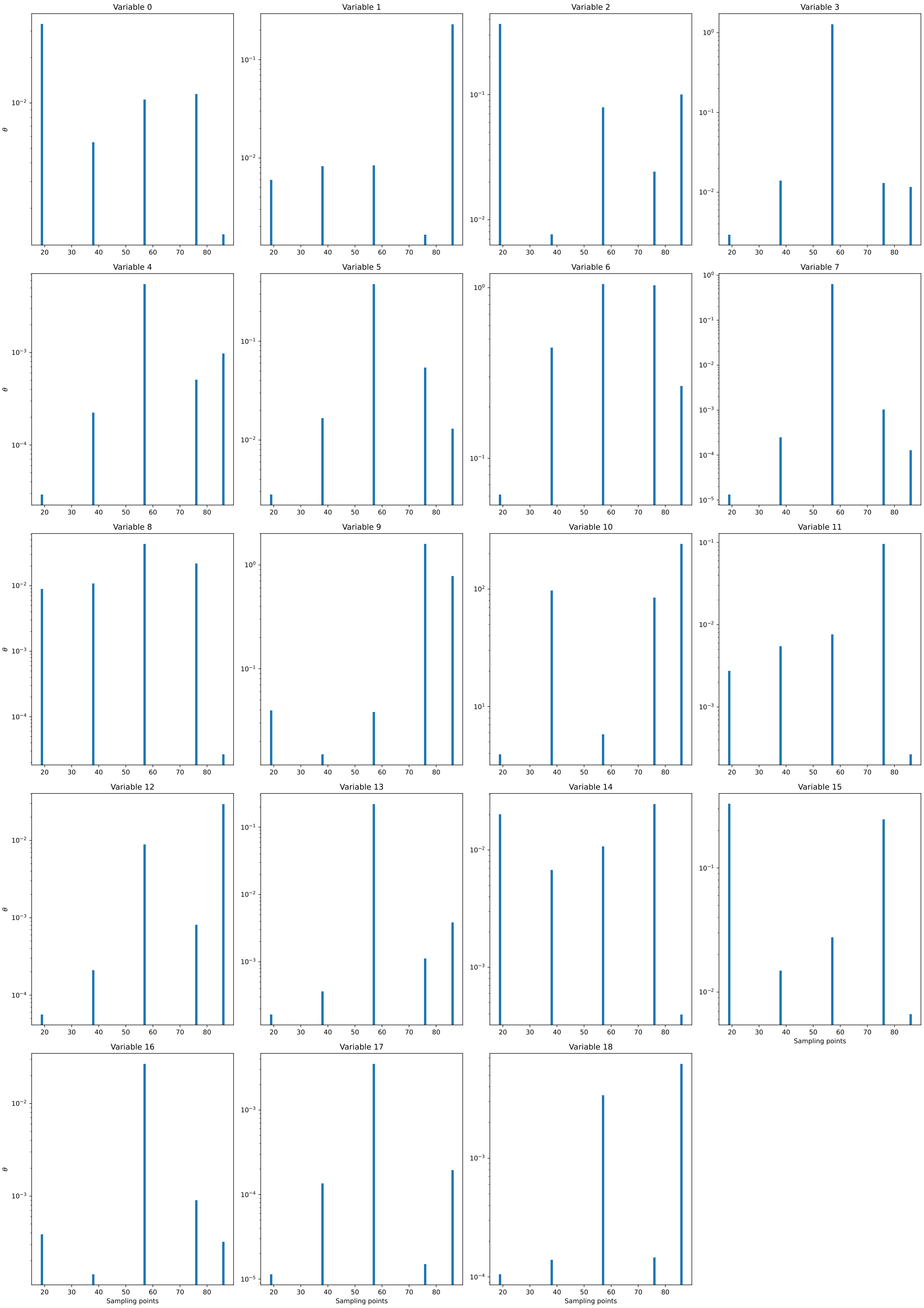


Figure 6.13: The evolution of the θ hyperparameters with respect to the number of training samples.

6.3. The Bayesian optimization procedure

The final and most important step of the proposed methodology is the optimization procedure itself. Recall that the goal of the proposed methodology is to find the thicknesses of wing components, $\boldsymbol{x}$, such that the weight $w(\boldsymbol{x})$ is minimized, while also ensuring that the front spar of the damaged wing can sustain a set of loads such that the maximum Mises stress does not exceed 80% of the yield stress, and that maximum penetration depth is smaller than 240 mm:

$$\begin{aligned} &\text{maximize} && w(\boldsymbol{x}) \\ &\text{with respect to} && \boldsymbol{x} \\ &\text{subject to} && c_1(\boldsymbol{x}) \leq 0, c_2(\boldsymbol{x}) \leq 0 \end{aligned} \tag{6.1}$$

In the equation above, $c_1(\boldsymbol{x}) \leq 0$ and $c_2(\boldsymbol{x}) \leq 0$ denote the maximum Mises and penetration depth constraints, respectively. As exposed in section 3.5, each optimization step entails the maximization of the constrained improvement acquisition function:

$$\begin{aligned} &\text{maximize} && \text{max}\,[(w_{min} - w(\boldsymbol{x}), 0] \cdot P[c_1(\boldsymbol{x}) \leq 0] \cdot P[c_2(\boldsymbol{x}) \leq 0] \\ &\text{with respect to} && \boldsymbol{x} \end{aligned} \tag{6.2}$$

As mentioned in the methodology, if the weight can be expressed analytically, there is no need to employ a surrogate model. Equation 6.2 is the resulting acquisition function. The first term is set to 0 if the improvement is negative such that the optimizer does not dwell in regions of no improvement. Indeed, the optimizer used by Trieste is SciPy's L-BFGS-B optimizer (The Trieste Contributors, 2020b), which is gradient-based. As a consequence, if the optimization search space is highly multimodal, the computational time may become quite large, especially in high dimensions. Even by setting a constant, null improvement, some optimization steps took nearly 3 hours to complete, although this is partially caused by the author's lack of experience with parallel computing in Python, as the acquisition function has been defined by the user. Therefore, this high computational time is also due to the acquisition function being computed on a single CPU. Moreover, because the design space is treated as discrete, the acquisition function value is computed by rounding each point to the nearest discrete value, thus achieving regions of constant acquisition responses, which may speed up the optimization process.

After the variable screening procedure, it was mentioned in the last section that all the training points will be projected on the reduced, 8-dimensional design space, while all insignificant variables will be kept at their minimum values. It was assumed that the constraint observations would not change with this projection. However, because of its analytical formulation, the weight decreased dramatically. The weight of the optimum design among the initial training points, before projection, was 78.83 kg. After projection, though, the weight has decreased dramatically to 52.58 kg, yielding a remarkable 33.31% weight decrease only thanks to the variable screening procedure. Of course, if the weight was not expressed analytically, many more significant variables would have been found for the objective function, which would not have resulted in such a dramatic decrease anymore. The extent of the damage after bird strike of the initial optimal design is illustrated in figure 6.14. The penetration and rib crushing are moderate, the depth being just 161.64 mm. Because the maximum Mises stress on the front spar after load introduction is 239.72 MPa, it is clear that there is room for improvement.

It was also mentioned that batch points can be requested by using the Kriging believer method, where the unknown or still ongoing observation of an optimization point can be fantasized, by imputing the value of the Kriging predictor itself. However, the following practical problems occurred: firstly, the number of Abaqus license tokens during working hours is limited. The number of computational nodes is also limited, even outside working hours. Moreover, there are two types of nodes: one with a total of 32 CPUs, and another one with a total of 20 CPUs. Therefore, the analysis computational wall time can greatly vary.

As a consequence, if an optimization batch is requested, the analyses often do not finish at the same time. Moreover, in the case of some static analyses not converging, their input files may be adjusted, such as requesting more iterations, reducing the minimum increment, or switching to using automatic stabilization coupled with the full Newton technique. Due to time constraints, it was decided to maximize the number of ongoing parallel optimization analyses, even if there were one or two ongoing simulations from the previous batch. All optimization data was kept track of in a comma-separated values (CSV) file.

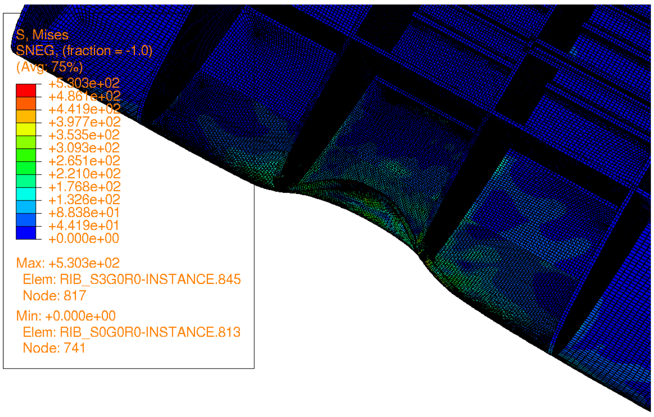


Figure 6.14: The damaged structure of the initial optimal design. Note: the upper skin has been hidden.

Practically, with the risk of a rather hectic optimization procedure, the following steps were taken: a batch of points was requested, and the coordinates of the next infill point, along with the absolute weight improvement, probabilities of feasibility, and acquisition function value were written to the CSV file. When an analysis completed, its results were written to the file. If there was room on the cluster for more simulations, another single point or a batch was requested, depending on the situation, and the Kriging believer method was used to also fantasize about the data of the other ongoing analyses.

It should also be noted that the penetration depth can be extracted earlier, in order to provide preliminary information about this constraint. If a bird strike analysis is ongoing, but another optimization point is required, the database can be consulted to extract the values of the total internal and plastic dissipation energies of the structure, along with the penetration depth. If the two energies are stable, it means that the only quantities which needs to stabilize are the elastic strain and kinetic energy, which are not expected to influence the depth result too much, as the plastic deformation had already occurred. As a consequence, this result can be used as a preliminary observation for the depth constraint, and was used to provide a more accurate information than the Kriging predictor itself.

Another strategy to reduce the optimization search space is to make use of the minimum observed feasible weight. Indeed, because the initial optimal weight was low, a large portion of the search space would have had a null constrained improvement. In some cases, the weight of a design with the maximum thickness of the wingbox rib, for instance, with all other variables kept at a minimum, was larger than the initial optimal weight. As a consequence, the maximum thicknesses of those variables were adjusted in order to reduce the search space, clipping out the regions of guaranteed null improvement. Once again, though, it is reiterated that this was possible due to the weight being formulated analytically, as the null improvement was guaranteed.

Finally, 14 batches were requested during the optimization procedure, some intertwining, as mentioned earlier. After these batches, 3 individual, sequential points were requested to assess the convergence of the procedure. The total number of optimization points was thus 47, simulated over the course of 13 days. Usually, one batch of two points was requested in the morning, and another batch of 3 in the evening, although, sometimes, this differed on whether the previous analyses converged, if modifications were needed, or if the cluster was crowded. Therefore, there is no clear-cut delimitation between different batches.

The final optimal weight was 47.40 kg, which was obtained at the 43rd optimization point, in the last batch. This result is rather remarkable, as the absolute minimum possible weight is only 4 kg lighter, at just 43.87 kg. As a reminder, the absolute minimum weight design, which is obtained by setting all design variables at their minima, is not feasible, as presented back in chapter 4. Interestingly, the last batch contained two consecutive optimal points, although the weight difference between them was just

80 grams. This aspect is illustrated in figure 6.15 as well, which presents the optimal weight achieved after each optimization iteration, the weight changing only if a lighter feasible design is discovered, which occurred at iterations 12, 27, 31, 42 and 43. It is apparent, though, that the weight improvement is increasingly smaller, hinting that the optimization process converges.

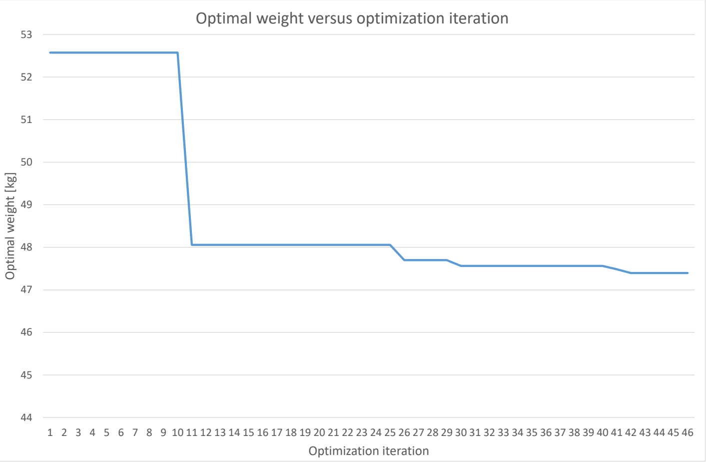


Figure 6.15: Optimal feasible weight versus optimization iteration.

The normalized constraint function values are illustrated in figure 6.16, where the $y = 0$ line denotes the boundary between feasible and unfeasible regions. All the points above the boundary violate the constraint, while those below obey it. The normalization has been done by subtracting the maximum constraint values from each observation, and scaling the results to the $[-1, 1]$ range. Although this scaling does not have any practical meaning, figure 6.16 showcases which optimization points were feasible or not; note that feasible points have been highlighted with a green circle on the x-axis. As expected, the feasible points are close to this boundary, although a rather erratic behaviour can be observed. This is partly due to the reduced accuracy of the depth surrogate model, as many optimization points seem to have large penetrations, but also due to the batch acquisition of infill points.

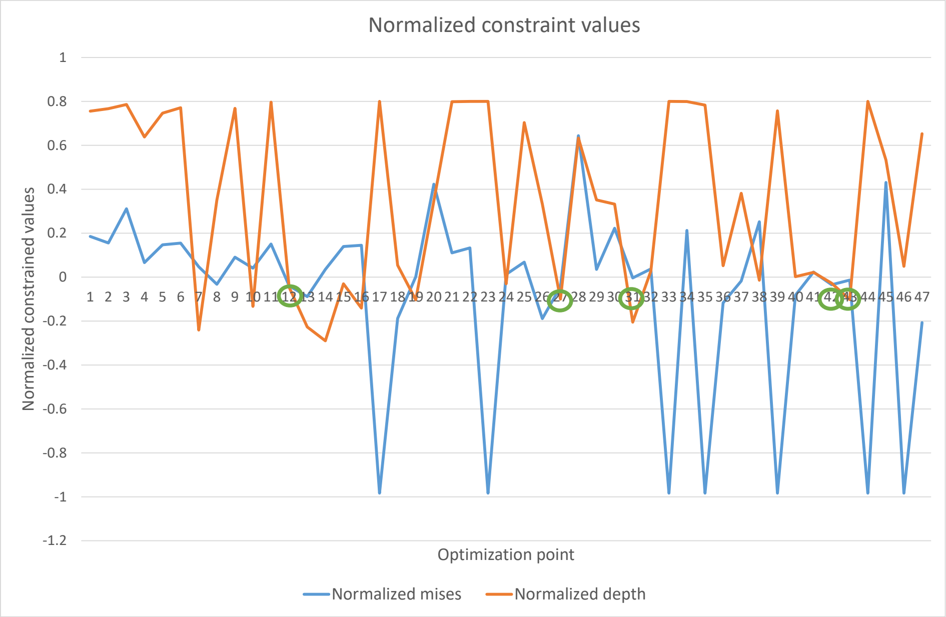


Figure 6.16: The normalized values of the constraints versus the number of optimization points. Note: the $y = 0$ line represents the boundary between the feasible and unfeasible regions, which are situated below and above it, respectively.

Figure 6.17 illustrates the probabilities of feasibility for both constraints. It can be seen that, for the first

12 points, nearly all infill points were chosen with a high probability of meeting the residual strength constraint, although some show a very low probability for the depth constraint. However, the first six points were all acquired in batch, as the optimization work started in a weekend, and many analyses could be run at the same time. The next six points also constitute a batch. Afterwards, though, the batches became intertwined due to the limited computational facilities during working days. As a consequence, among the first 2 batches, the probability of satisfying the depth constraint is low partly due to the Kriging believer method. After the discovery of a new optimum, though, the optimizer started promoting regions with a high probability of the depth constraint.

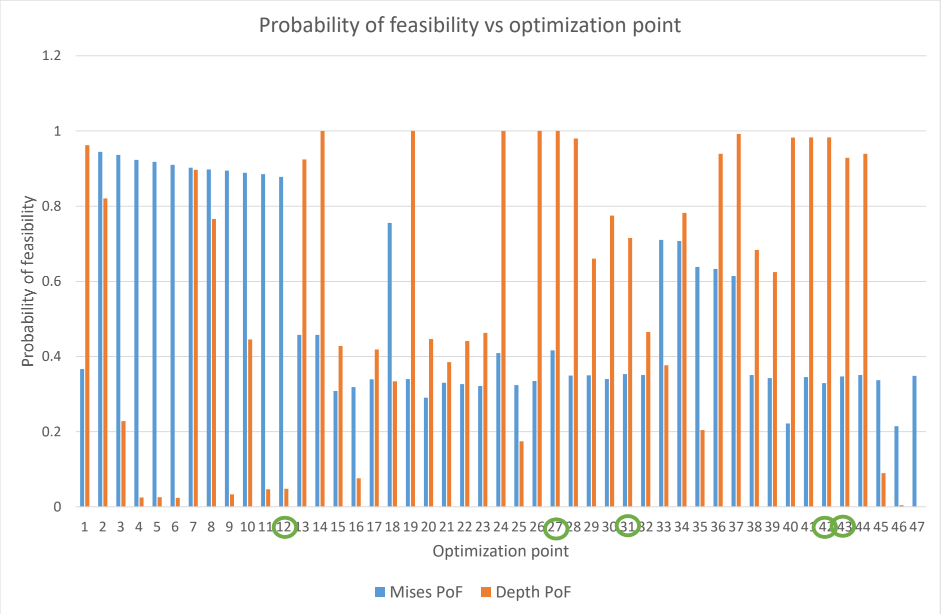


Figure 6.17: The probability of feasibility of the constraints versus the number of optimization points.

After the 44th run, though, the optimizer started suggesting points with a null constrained improvement. The number of initial samples and optimization runs has kept at the recommended values by the Trieste documentation (The Trieste Contributors, 2020b) of 8000, which is 1000 times the number of dimensions, but, apparently, the number had not been enough. This mishap does raise a question on whether the previous optimization runs had actually maximized the acquisition function, but, due to time constraints, the optimization process of the present study was not redone, as another couple of weeks would have been needed. Therefore, the remaining three analyses were carried in sequential order, with no solution being fantasized, in order to attempt finding a maximum of the acquisition function. However, the number of initial samples for the optimizer could only be doubled from 8,000 to 16,000 due to the high optimization time aforementioned, which increased the computational wall time to 3 hours.

Figure 6.18 illustrates the values of the acquisition function versus the optimization points. The last three points clearly have a low value, the 47th value being in the order of 10^{-19}. Had it not been for the doubts regarding the maximizer, it could have been concluded that the optimization process had converged. Unfortunately, though, it can only be concluded that the number of initial samples for the optimizer is too low. Figure 6.19 presents the values of the acquisition function against the leading edge and nose rib thicknesses, while keeping all other variables the same as the optimum. The scattered data points have all been projected on this two-dimensional space, and the optimum is depicted in yellow, in the lower-left corner. It is clear that there is a point in the lower-left region of the plot which maximizes the acquisition function in the lower-left, although the maximizer failed to see it.

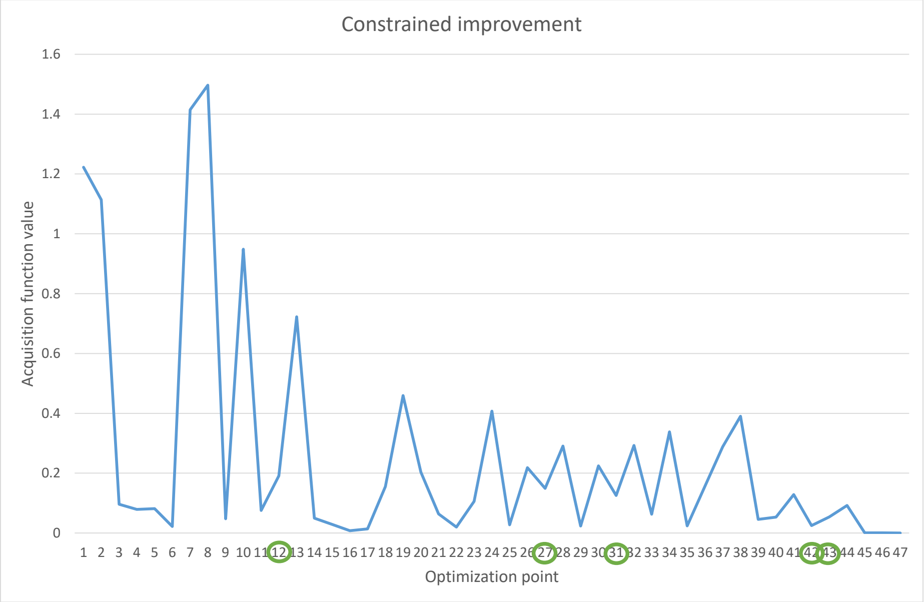


Figure 6.18: The constraint improvement acquisition function versus the number of optimization points.

Nevertheless, these problems are inherent to optimizers in high-dimensional spaces, and careful consideration must be given to their choice. The evaluation of the suitability of the optimizers with respect to the maximization of a high-dimensional acquisition function was outside the scope of the present study, and the default recommendations given by the Trieste documentation were followed. It is recommended, however, to scale up greatly the number of initial samples and runs for the optimizer, given that the acquisition function computation can make use of parallel computing.

As illustrated back in figure 6.15, the consecutive weight reductions after the optimal point obtained in the 12th iteration, which had 48.06 kg, were neither significant nor negligible. For instance, the 27th optimization dropped a further 358 grams, the next 3 successful runs decreasing the weight by 137, 78, and 80 grams, respectively. Nevertheless, when comparing the weight reduction of the final optimal solution to the initial minimum, before dimensionality reduction, the decrease is remarkable, the optimal design being 39.88% lighter than the initial one.

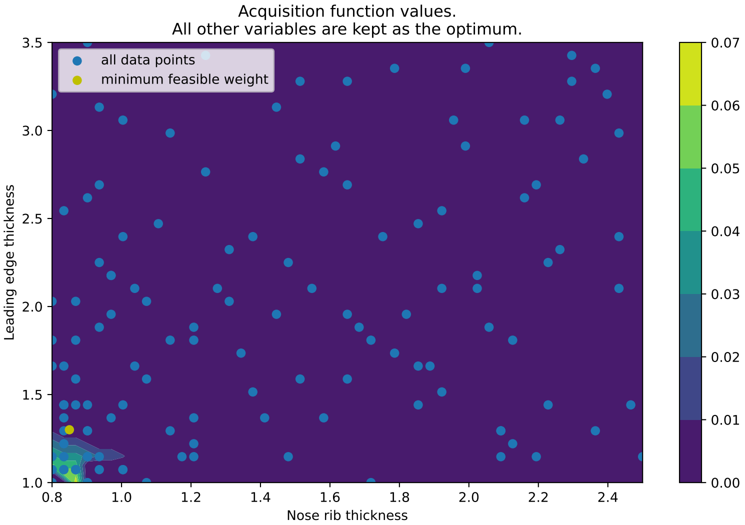


Figure 6.19: The acquisition function value against the leading edge and nose rib thicknesses. Note: all other variables are fixed at the values of the optimal design.

The damaged structure of the final optimal solution, which is reiterated that it may not be the true optimum, is depicted in figure 6.20. It is apparent that, compared to the initial design, the penetration depth is larger, and the rib crushing is more severe. The penetration depth is 209.342 mm, leaving room for improvement, and the maximum Mises stress on the front spar is 291.060 MPa.

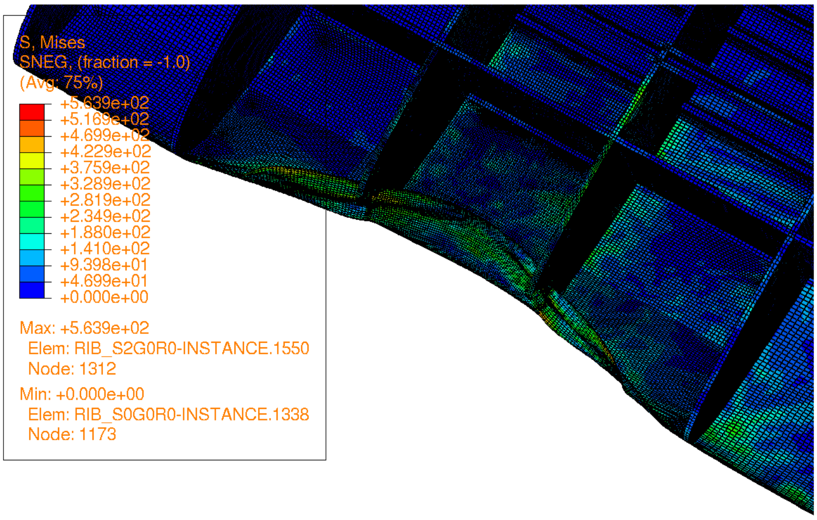


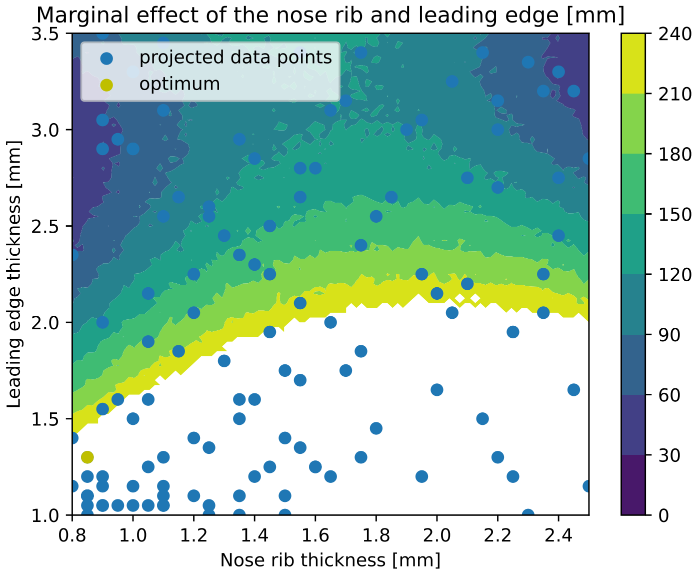
Figure 6.20: The damaged structure of the final optimal design. Note: the upper skin has been hidden.

The thicknesses of the components of the optimal design are given in table 6.4. It is interesting that a combination of thin leading edge and nose rib has been found, as the marginal effect of the leading edge with respect to the depth penetration constraint did not estimate a feasible design, as depicted in figure 6.10a, which only promotes the use of Bayesian optimization. Moreover, the variable screening procedure applied on the same constraint uncovered a high non-linearity of the marginal effect of the wingbox rib, as illustrated in figure 6.10b, but, in the end, a low thickness has been found to suffice. A component whose thickness could have been improved is probably the third station stringer, which is quite thick.

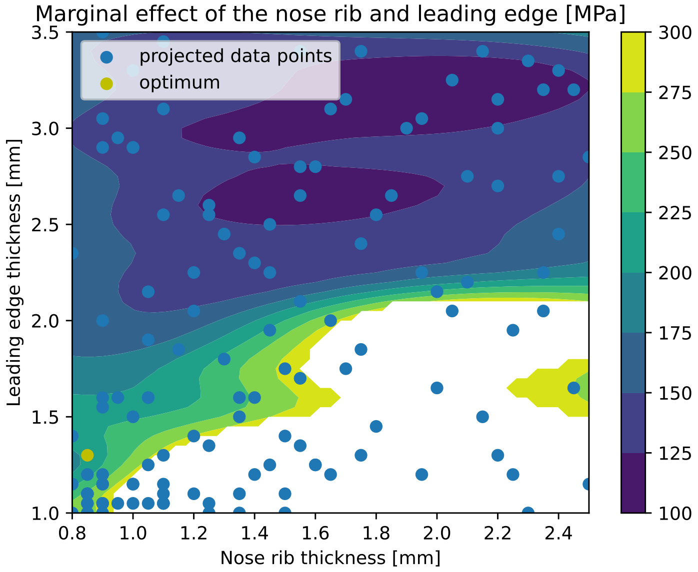
Component	Thickness [mm]
Stringer 1	1.15
Stringer 2	0.8
Stringer 3	1.55
Stringer 4	0.95
Stringer 5	0.8
Stringer 6	0.8
Nose rib	0.85
Wingbox rib	0.85
Front spar	1.5
Rear spar	1.75
Skin panel zone 1	1.3
Skin panel zone 2	1.0
Skin panel zone 3	1.0
Skin panel zone 4	1.0
Skin panel zone 5	1.05
Skin panel zone 6	1.0
Skin panel zone 7	1.0
Skin panel zone 8	1.0
Skin panel zone 9	1.0

Table 6.4: Thicknesses of the optimal design.

Just as a final note on the employed surrogate models, figures 6.21a and 6.21b present the marginal effects of the leading edge and nose rib thicknesses for the depth and strength constraints, respectively. Recall that these variables have been identified as the most significant for both constraints, and that their marginal effect can be seen as a mean value to be expected when considering only a reduced set of variables. The optimization process should be given some credit, as it has found an optimum in a region where the depth constraint is violated, as an expected value when taking into account these two variables. Moreover, the two figures emphasize that in the upper-central region, the two constraints seem to contradict each other, as a large penetration is expected, whereas a low stress on the front spar is predicted. As a consequence, the idea that both constraints should be simultaneously taken into account is reiterated.



(a) The marginal effects of the nose rib and leading edge thicknesses, for the depth constraint. Note: infeasible regions are masked.



(b) The marginal effects of the nose rib and leading edge thicknesses, for the residual strength constraint. Note: infeasible regions are masked.

Figure 6.21: The marginal effects of both constraints.

7

Conclusions and recommendations

The scope of the present study was to develop a methodology for the incorporation of bird strike crashworthiness requirements within a multidisciplinary optimization framework. The proposed methodology combines Bayesian optimization with a preliminary variable screening procedure. As outlined in chapter 1, if the design space involves many variables, as is the case in numerous crashworthiness optimization problems in the aerospace domain, the number of dimensions may prove prohibitive for the construction of a surrogate model, as the required number of training points may be too large. A variance-based variable screening method is thus proposed, which quantifies the main and interaction effects of the variables by conducting an analysis of variance on an initial Kriging surrogate model.

After the significant variables are discovered for each constraint function, the training data points are projected onto the reduced optimization space, and each surrogate model is retrained in its corresponding domain. It should be noted that the optimization space is the reunion of all the design spaces of the objective and constraint problems. The constrained Bayesian optimization procedure is finally conducted, with the constrained improvement serving as acquisition function.

In order to demonstrate the efficacy of the proposed methodology, a single-objective constrained optimization problem was formulated, the aim being to minimize the weight of a wing, while ensuring the feasibility of the design against two constraints related to bird strike crashworthiness. Significant weight reductions have been achieved, which is a remarkable outcome with respect to high-dimensional, multimodal, non-linear constraint functions.

Chapter 2 has presented recent literature on crashworthiness optimization, and culminated with a collection of research questions pertaining to the incorporation of bird strike requirements within a multidisciplinary optimization framework. As a consequence, the present chapter will first summarize the findings related to the nine proposed research questions, followed by a section comprising of various recommendations for future studies and improvements.

7.1. Answers to the proposed research questions

Because the present study combines various directions of research, the questions formulated in chapter 2 will be reordered based on each step of the proposed methodology. The most fundamental direction of research was related to the choice of constraints related to bird strike crashworthiness requirements:

> What is a sufficiently representative constraint function responsible for the feasibility of a design to ensure compliance with bird strike crashworthiness regulations?

Two constraint functions have been found. The first concerned the residual strength of the structure, and required the maximum Mises stress on the front spar not to exceed 80% of the yield strength of the material, when subjecting the damaged structure to a combined load case. The latter has been chosen such that the absolute minimum weight undamaged design can support the loads without yielding, and was formulated to emulate EASA CS 25.571(e), which requires that “the aeroplane must be capable of successfully completing a flight during which likely structural damage occurs as a result of bird impact as specified in CS 25.631. The AMC to 25.571 (in paragraph 2.7.2) specifies the loads associated with get home conditions that have to be met for this case". Similar constraints can be formulated for other components of the wing, and can be easily incorporated in the proposed methodology.

The second constraint enforced a maximum penetration depth corresponding to half the distance between the front spar and the leading edge. The latter may be important in the case of electric vertical take-off and landing aircraft, for instance, as battery packs may be situated in the leading edge, and a significant safety risk may arise if they are damaged. While the second constraint has been the subject of various optimization studies such as Ollar, Jones, and Toropov, 2017 and Pahange and Abolbashari, 2016, the residual strength of the damaged structure has never been taken into account in the context of multidisciplinary optimization, at least to the knowledge of the author.

It has been discovered that significant stresses can occur to the front spar, even if the penetration is minimal, and vice-versa. For instance, one of the initial designs, which boasted a weight of 100 kg, had only a penetration depth of 85 mm. However, the maximum Mises stress of the front spar was 353 MPa, with 369 MPa being the yield stress of the employed material, Aluminium 2024-T3(51). Therefore, as the considered load case is far from critical, both constraints should be taken into account, as plastic deformation of the front spar may occur even if the extent of damage does not appear to be significant. The incorporation of these constraints is targeted by the next research question:

> What is a suitable approach to incorporate the constraint functions related to bird strike crashworthiness in a multidisciplinary optimization framework?

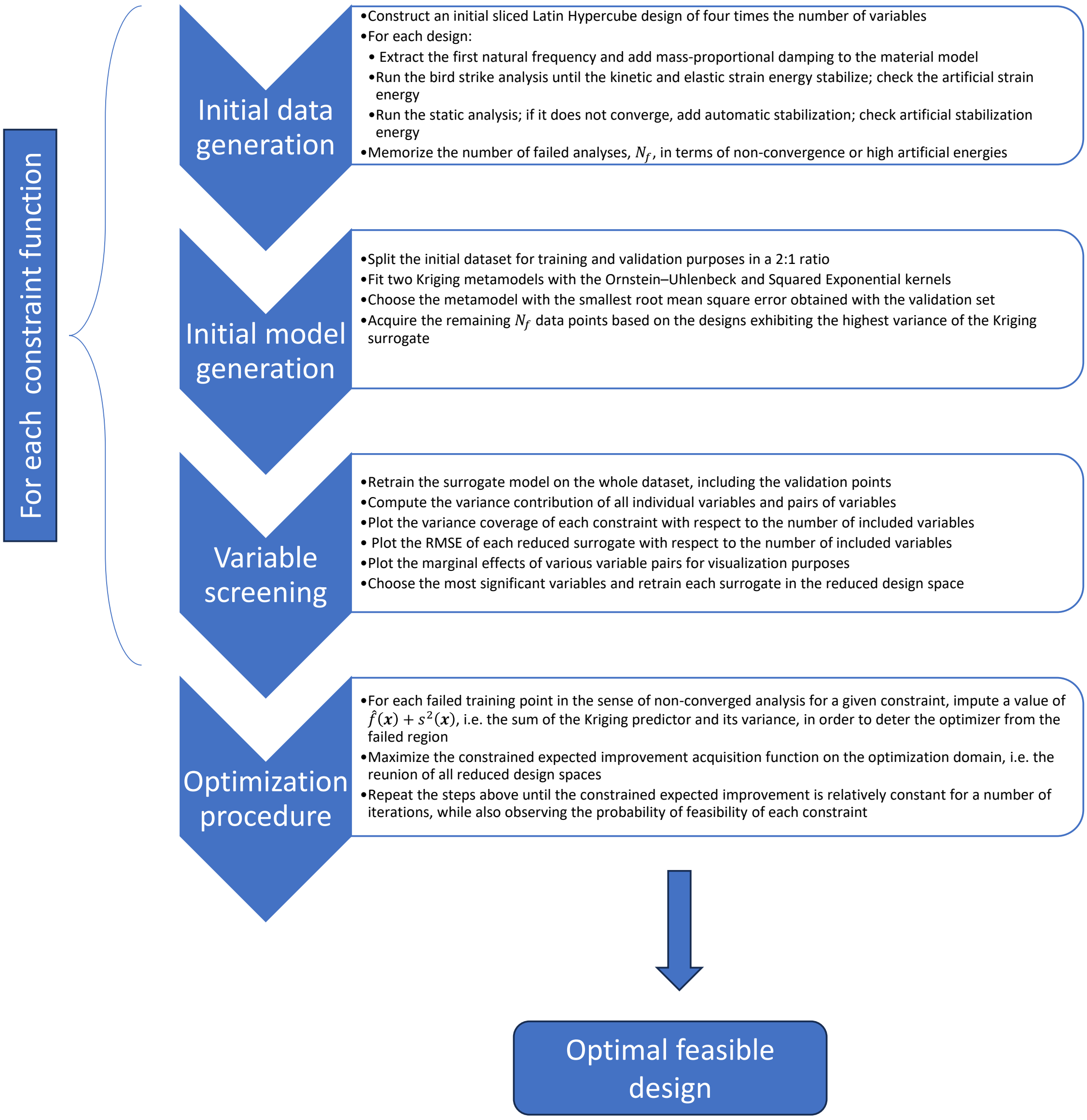
Chapter 2 has revealed that surrogate models are reliable in addressing computationally intensive, high-dimensional, multi-modal underlying functions, aligning well with the requirements of the chosen constraint functions. As Gaussian process-based metamodels constitute the foundation of Bayesian optimization, it was concluded that their use, combined with a variable-screening method, can prove dependable in the context of the present study. As a consequence, the following approach is proposed, which is also illustrated in figure 7.1.

Firstly, an initial sliced Latin Hypercube design is generated. As the proposed constraint functions are done sequentially in the present case study, for each point, the first natural frequency is extracted to include a mass-proportional damping to the bird strike analysis which follows. The impact simulation is terminated when the kinetic and elastic strain energies stabilize, in order for the static analysis to find equilibrium when importing the state of the model from the dynamic simulation. If the artificial strain energy is higher than 5% of the total internal energy of the model, the bird strike analysis is deemed as failed in the sense that no result can be extracted. Otherwise, the static analysis follows. If the latter does not converge, artificial stabilization is employed, and the artificial strain energy is checked not to exceed 5% of the total internal either. If the analysis still does not converge or is overdampened, the result of the second constraint cannot be extracted.

The initial dataset is then split for training and validation purposes in a 2:1 ratio. It is proposed to compare two kernels and select the most appropriate based on an error measure. Afterwards, the initial dataset is supplemented with additional samples which exhibit the largest Kriging variance. The variable screening procedure follows, which starts with the computation of all the contributions of individual and pairs of variables. Finally, the optimization procedure culminates with maximizing a Bayesian acquisition function until a convergence criterion is met, and extracting an optimal, feasible design.

Based on the steps proposed in the methodology, other research questions arise, such as the following, pertaining to surrogate models:

> - Given that global accuracy is not of great interest during the optimization process, what is a sufficient number of training points for the initial surrogate model?
> - Due to the large dimensionality of the problem, coupled with a high computational intensity of impact simulations, what is a suitable measure to assess the quality of the surrogate model of the proposed constraint?



- Construct an initial sliced Latin Hypercube design of four times the number of variables
- For each design:
  - Extract the first natural frequency and add mass-proportional damping to the material model
  - Run the bird strike analysis until the kinetic and elastic strain energy stabilize; check the artificial strain energy
  - Run the static analysis; if it does not converge, add automatic stabilization; check artificial stabilization energy
- Memorize the number of failed analyses, N_f, in terms of non-convergence or high artificial energies

- Split the initial dataset for training and validation purposes in a 2:1 ratio
- Fit two Kriging metamodels with the Ornstein–Uhlenbeck and Squared Exponential kernels
- Choose the metamodel with the smallest root mean square error obtained with the validation set
- Acquire the remaining N_f data points based on the designs exhibiting the highest variance of the Kriging surrogate

- Retrain the surrogate model on the whole dataset, including the validation points
- Compute the variance contribution of all individual variables and pairs of variables
- Plot the variance coverage of each constraint with respect to the number of included variables
- Plot the RMSE of each reduced surrogate with respect to the number of included variables
- Plot the marginal effects of various variable pairs for visualization purposes
- Choose the most significant variables and retrain each surrogate in the reduced design space

- For each failed training point in the sense of non-converged analysis for a given constraint, impute a value of $\hat{f}(\boldsymbol{x}) + s^2(\boldsymbol{x})$, i.e. the sum of the Kriging predictor and its variance, in order to deter the optimizer from the failed region
- Maximize the constrained expected improvement acquisition function on the optimization domain, i.e. the reunion of all reduced design spaces
- Repeat the steps above until the constrained expected improvement is relatively constant for a number of iterations, while also observing the probability of feasibility of each constraint

Figure 7.1: The proposed steps for incorporating bird strike crashworthiness requirements in a multidisciplinary design optimization framework.

Based on literature on Bayesian optimization, it was decided to generate at least four times the number of variables for the initial training set, such that a third of them can be used for validation purposes. Even if global accuracy is not the primary scope of the initial surrogate and that the constrained expected improvement approach balances the exploration and exploitation of the design space, the preliminary metamodel should be fit for screening purposes. A suggestion was to generate the proposed number of training points, evaluate the root mean square error, and decide if the model is of sufficient accuracy. In the case of the residual strength constraint, a sufficient accuracy had been reached, as its RMSE corresponded to just 20% of the yield strength. However, a rather large RMSE was discovered for the depth constraint, and another suggestion was given: to add sequential points in the areas of high predicted variance. However, after the addition of 12 such points, no significant improvement was found, and it was then proposed that, in such cases, several analyses should be conducted after the variable ranking procedure, in order to give additional decision criteria. These analyses will be explained later on, when the research questions corresponding to dimensionality reduction arise.

As for the suitable measure to assess the quality of the surrogate model, the mean square and cross-validation errors were compared, with the MSE serving as a baseline value. A disadvantage of the former is that additional data points should be generated for validation purposes, Viana et al., 2021 suggesting that the size of the validation set should be half the size of the training set. While the cross-validation error does not have this disadvantage, it was found that, when assessing the quality of the surrogate model in various reduced design spaces, it was not even close to being directly proportional to the MSE. In fact, in some cases, the two errors sometimes did not even follow the same trends. As a consequence, the MSE should be used as a global error measure, the cross-validation error being unreliable, especially since it was initially proposed to be used to judge the quality of the variable ranking procedure.

Moving on to the next step, the following research questions have also been formulated with respect to dimensionality reduction:

- What are the limitations of the proposed variance-based variable screening method with respect to the constraint functions?
- Which additional criteria can be formulated to safeguard against an erroneous identification of significant variables?
- How does the global accuracy of the surrogate model change with the reduction of the design space?

A downside of the proposed variable screening procedure is the necessity of an initial surrogate model. However, it was assumed that the quantification of the main and interaction effects among variables deserved this risk. Nevertheless, the analysis of variance on the residual strength constraint uncovered that higher-order interactions took place, which were not implemented in the developed Python procedure. For the depth constraint, though, higher-order effects did not occur. As 8 variables were identified as significant for the residual strength constraint versus 3 for the other one, it can be concluded that one of the limitations of the procedure would be the necessity to compute higher-order effects to arrive to a quantification of reasonable accuracy. However, it was chosen not to pursue this further development, mainly due to time constraints.

Another severe limitation is the lack of a recovery mechanism against erroneous variable identification, and it is recommended to pursue a future study related to this scope. This problem may be solved by adopting a different dimensionality reduction method altogether, such as the Kriging Partial Least Squares, KPLS, developed by Bouhlel et al., 2016. Nevertheless, the present work has been founded on the assumption that the significant variables do not change with the addition of more training points. The downside of this supposition is that the true optimum may never be found, although this risk is inherent to optimization based on surrogate models, and the optimization procedure will still manage to find a design of lower weight nonetheless.

The second research question aforementioned concerns additional criteria to safeguard against such an erroneous identification, especially since it was discovered that a variable was wrongly ranked for the depth constraint. Apart from engineering judgment and a heatmap illustrating all the main and interaction effects of the variables, two additional steps can be taken. Firstly, the construction of a plot where the variance coverage is depicted against an increasing number of variables can be useful. The variables are sequentially added such that the addition of the main and interaction effects of variables already included in the model maximizes the variance coverage. At each step of this addition, a list can be made with the next variable to be added, along with the resulting variance coverage. This step, combined with engineering judgment, raised a warning when choosing the most significant variables for the depth constraint, and it was decided to eliminate one of them.

Another method which was proven useful was to analyse the root mean square error of the surrogate model retrained on the training set of reduced dimension. In the case of the depth constraint, a significant accuracy increase has been registered when the first three erroneous variables were chosen, while an additional remarkable increase has been discovered with the three correct variables. Therefore, this method can also be used to assess the quality of the reduced design space.

The third research question is also related to this aspect. While it was expected that the higher sampling density of the reduced design space would always result in an increase in accuracy, the residual strength constraint has disproved this supposition. While the variance coverage seemed not to change between the inclusion of just 8 or all variables, the existence of higher-order effects is apparent from the fact that the root mean square error of the model did not change. Nevertheless, even if the accuracy did not improve, it did not deteriorate either, and the dimensionality reduction played a pivotal role during the optimization procedure, which could prove time-consuming on such a high-dimensional search space. As a consequence, the accuracy of the model seems to be improved by the variable screening procedure only if no higher-order interactions are uncovered, such that the variance coverage can be estimated correctly. However, this is just another assumption, which should be taken with a grain of salt.

Another research question related to reduced spaces is the following:

> How can the optimization problem be formulated if the objective and constraint functions are defined on different reduced design spaces?

As mentioned earlier, the optimization space is the union of all the design spaces of the objective and constraint problems. The Bayesian optimization procedure is conducted in the optimization space, while all the terms of the acquisition function values are evaluated in their corresponding design spaces. For instance, the expected improvement of the objective function, and probability of feasibility of each constraint are evaluated by projecting the variables of the optimization space on the reduced design space of the objective and constraint functions, respectively.

Finally, the following research question has been formulated with respect to the optimization process:

> What is a suitable convergence criterion for the optimization process?

It is acknowledged that time constraints are usually the decisive factor when stopping an optimization procedure. However, if these constraints are not taken into account, the optimization process can be safely stopped when the acquisition function value remains relatively constant over a number of consecutive iterations. During the preliminary exercises of the present study, the weight improvement, i.e. not multiplied with the probabilities of feasibility, was first considered for convergence. However, it was discovered that a surge in improvement is often accompanied by low probabilities of feasibility, which should have been expected. The algorithm is trapped in an explorative state in this case, indicating that it struggles to find any more feasible solutions. The value of the acquisition function itself, though, should stay consistently low in this case, thus constituting a reliable convergence criterion. In the present study, a rather fortunate event happened, as not only did the acquisition function decrease, but it also approached a null value during the last iterations, indicating either that the algorithm had converged, which was improbable, or that the optimizer struggled to maximize the acquisition function. Unfortunately, the latter proved to be true, and it was suggested to scale up the number of initial points recommended by the Trieste documentation and runs of the optimizer. However, due to time constraints, the optimization procedure was not redone, mainly because it took 13 days in the first place. Nevertheless, returning to the convergence criterion, the number of iterations should be assessed by the user, as one more iteration can mean 15 more hours spent on the optimization process in the present case study, for instance.

Based on the preceding discussion, all the research questions have been tackled, although the last one was only partially answered, as other problems arose. Nevertheless, as the proposed methodology spanned diverse fields of research, various problems were expected, and, regrettably, time constraints prevented their resolution.

7.2. Recommendations

First and foremost, the proposed variance-based screening procedure needs to be adapted in order to include higher-order effects, as their occurrence has been discovered for the residual strength constraint. Although a manner to incorporate categorical and discrete variables has been suggested in section 3.4, as Garrido-Merchán and Hernández-Lobato, 2020 proposed a different covariance function, its implementation is still necessary for future research. One more recommendation with respect to variable ranking would be the implementation of a recovery mechanism, although KPLS metamodels (Bouhlel et al., 2016) could also be employed.

The inclusion of variables such as various materials or front spar positions can yield an interesting case study, especially since even more failure modes can be uncovered, and that MDM, coupled with the

Abaqus interface, can readily generate the necessary input files without further development. Another aspect which should be included in future versions of the Abaqus interface, though, is the ability to model fasteners. Their inclusion is expected to improve the energy absorption capabilities of the leading edge, thanks to them taking over a portion of the kinetic energy of the bird and dissipating it through their failure.

Another interesting pursuit was the use of Variational Gaussian processes as classifiers, although they raised numerous other problems. The variance-based screening method would have needed to be adapted, which is no easy task. However, these classifiers could prove useful to predict various failure modes. A downside of this suggestion, though, is that the quantification of interaction effects is not possible anymore. A final recommendation is, of course, to give a more careful consideration of the acquisition function optimizer, at least by increasing substantially the number of initial optimization samples recommended by Trieste, which was not done in the present work.

7.3. Conclusions

As presented in the last section, the proposed methodology achieved remarkable weight savings, the final result being nearly 40% lighter than the minimum feasible design weight obtained after the initial 76-point training sample. The variable screening procedure, on its own, reduced the initial minimum weight by 33%, assuming that the observations did not change when the insignificant variables were fixed at their lowest corresponding values. This assumption is based on the fundamental premise that the initial surrogate was accurate enough for variable screening, and that the ranking procedure identified the correct significant variables, although it is acknowledged that it is not always true, especially in the present case study, when higher-order interactions were apparent.

The initial minimum weight considered after dimensionality reduction serves only as a starting weight for the Bayesian optimization procedure, and all subsequent infill points are validated. Therefore, the initial weight only promotes a more aggressive search. It is recognized that such remarkable weight savings are partly due to the reduced number of significant variables, but also due to the analytical formulation of the objective function. Indeed, the optimization procedure was terminated after only 14 batches totalling 47 optimization points, roughly 6 times the number of dimensions, with 3 individual consecutive analyses at the end to assess the convergence, although it was shown that the optimizer struggled and convergence was not attained. However, the optimization may have been sped up because, whenever the weight at a design point is higher than the minimum observed feasible weight, the improvement is considered null, yielding the acquisition function null. Moreover, as the starting weight was only 9 kilograms higher than the absolute minimum possible weight, which was not feasible, a large portion of the design space was null. As a consequence, even if the remaining 8 variables can constitute a rather large search space, a weight improvement could only be found in a rather small area. Another possible reason is the treatment of the thicknesses as discrete variables, thus forcing the acquisition function to have regions of constant value and accelerating the optimization procedure.

Even if the objective function was modelled with surrogates as well, the methodology is expected to be successful, although the optimization procedure might converge much slower. The inclusion of categorical variables has also been anticipated, and the chosen DoE, surrogate models and variable ranking method can accommodate them, with the small modifications proposed in chapter 3.

In conclusion, the use of Kriging models, variable ranking procedures, and Bayesian optimization all contributed to the achievement of significant weight savings, demonstrating the efficacy of the proposed methodology. Furthermore, it has been demonstrated that the incorporation of a residual strength requirement is necessary, as many cases were discovered where no significant penetration occurred despite the application of the considered load case, which is far from critical, resulting in high stresses to the front spar of the damaged structure. Moreover, the inclusion of this constraint also seems not to have been taken into account in any other optimization process, at least to the knowledge of the author, thus marking an inaugural attempt in introducing residual strength requirements with respect to bird strike crashworthiness.

References

Abaqus, Inc. (2009). *4.3 example: Hourglassing in a rubber block*. Retrieved November 6, 2023, from https://classes.engineering.wustl.edu/2009/spring/mase5513/abaqus/docs/v6.5/books/gsx/default.htm?startat=ch04s03.html

Acar, E., & Rais-Rohani, M. (2009). Ensemble of metamodels with optimized weight factors. *Structural and Multidisciplinary Optimization*, *37*(3), 279–294. https://doi.org/10.1007/s00158-008-0230-y

Acar, E. (2010). Various approaches for constructing an ensemble of metamodels using local measures. *Structural and Multidisciplinary Optimization*, *42*(6), 879–896. https://doi.org/10.1007/s00158-010-0520-z

Acar, E., & Solanki, K. (2009). Improving the accuracy of vehicle crashworthiness response predictions using an ensemble of metamodels. *International Journal of Crashworthiness*, *14*(1), 49–61. https://doi.org/10.1080/13588260802462419

Airlines Inform. (2012). *Fokker 50*. Retrieved November 27, 2023, from https://www.airlines-inform.com/commercial-aircraft/fokker-50.html

Astori, P., & Impari, F. (2013). Crash response optimisation of helicopter seat and subfloor. *International Journal of Crashworthiness*, *18*(6), 570–578. https://doi.org/10.1080/13588265.2013.815602

Ba, S. (2015). *SLHD: Maximin-Distance (Sliced) Latin Hypercube Designs*. Retrieved November 23, 2023, from https://CRAN.R-project.org/package=SLHD

Bao, Y. (2003). *Prediction of ductile crack formation in uncracked bodies* [Doctoral dissertation, Massachusetts Institute of Technology].

Barber, J., Taylor, H., & Wilbeck, J. (1975). *The Characterization of Bird Impacts on a Rigid Plate: Part I, Technical Report AFFDL-TR-75-5, ADA021142*.

Bisagni, C., Lanzi, L., & Ricci, S. (2002). Optimization of helicopter subfloor components under crashworthiness requirements using neural networks. *Journal of Aircraft*, *39*(2), 296–304. https://doi.org/10.2514/2.2927

Bishop, C. M. (1995). Neural networks for pattern recognition. Oxford University Press, Inc.

Bouhlel, M., Bartoli, N., Morlier, J., & Otsmane, A. (2016). Improving kriging surrogates of high-dimensional design models by Partial Least Squares dimension reduction. *Structural and Multidisciplinary Optimization*, *53*, 935–952. https://doi.org/https://doi.org/10.1007/s00158-015-1395-9

Budgey, R. (2000). *The development of a substitute artificial bird by the international Bird strike Research Group for use in aircraft component testing. International Bird Strike Committee ISBC25/WP-IE3, Amsterdam*. International Bird Strike Committee.

CEA/DEN, EDF R&D, & OPEN CASCADE. (2017). *2D Meshing Hypotheses*. Retrieved October 11, 2023, from http://trac.lecad.si/vaje/chrome/site/doc8.3.0/gui/SMESH/a2d_meshing_hypo_page.html#hypo_quad_params_anchor

Chuong, D. B. (2008). *The Multivariate Gaussian Distribution*. Retrieved February 15, 2023, from https://cs229.stanford.edu/section/gaussians.pdf

Craig, K. J., Stander, N., Dooge, D. A., & Varadappa, S. (2005). Automotive crashworthiness design using response surface-based variable screening and optimization. *Engineering Computations*, *22*(1), 38–61. https://doi.org/10.1108/02644400510572406

Cramer, E. J., Dennis, J. E., Jr., Frank, P. D., Lewis, R. M., & Shubin, G. R. (1994). Problem formulation for multidisciplinary optimization. *SIAM Journal on Optimization*, *4*(4), 754–776. https://doi.org/10.1137/0804044

Dassault Systèmes. (2022a). **PLASTIC*. Retrieved November 21, 2023, from https://abaqus-docs.mit.edu/2017/English/SIMACAEKEYRefMap/simakey-r-plastic.htm#simakey-r-plastic

Dassault Systèmes. (2022b). **STATIC*. Retrieved November 9, 2023, from https://docs.software.vt.edu/abaqusv2022/English/SIMACAEKEYRefMap/simakey-r-static.htm

Dassault Systèmes. (2022c). *About transferring results between Abaqus analyses*. Retrieved November 17, 2023, from https://docs.software.vt.edu/abaqusv2022/English/SIMACAEANLRefMap/simaanl-c-transferoverview.htm

Dassault Systèmes. (2022d). *Choosing a shell element*. Retrieved November 17, 2023, from https://docs.software.vt.edu/abaqusv2022/English/SIMACAEELMRefMap/simaelm-c-shellelem.htm

Dassault Systèmes. (2022e). *Damage evolution and element removal for ductile metals*. Retrieved November 21, 2023, from https://abaqus-docs.mit.edu/2017/English/SIMACAEMATRefMap/simamat-c-damageevolductile.htm

Dassault Systèmes. (2022f). *Damage initiation for ductile metals*. Retrieved November 16, 2023, from https://docs.software.vt.edu/abaqusv2022/English/SIMACAEMATRefMap/simamat-c-damageinitductile.htm

Dassault Systèmes. (2022g). *Equation of state*. Retrieved November 20, 2023, from https://docs.software.vt.edu/abaqusv2022/English/SIMACAEMATRefMap/simamat-c-eos.htm

Dassault Systèmes. (2022h). *Explicit dynamic analysis*. Retrieved November 17, 2023, from https://docs.software.vt.edu/abaqusv2022/English/SIMACAEANLRefMap/simaanl-c-expdynamic.htm#simaanl-c-expdynamic-bulkviscosity

Dassault Systèmes. (2022i). *Material damping*. Retrieved November 17, 2023, from https://docs.software.vt.edu/abaqusv2022/English/SIMACAEMATRefMap/simamat-c-dampingopt.htm

Dassault Systèmes. (2022j). *Mesh tie constraints*. Retrieved October 26, 2023, from https://docs.software.vt.edu/abaqusv2022/English/?show=SIMACAECSTRefMap/simacst-c-tiedconstraint.htm

Dassault Systèmes. (2022k). *Quasi-Newton solution technique*. Retrieved November 17, 2023, from https://docs.software.vt.edu/abaqusv2022/English/SIMACAETHERefMap/simathe-c-quasinewtsol.htm

Dassault Systèmes. (2022l). *Solving nonlinear problems*. Retrieved November 9, 2023, from https://docs.software.vt.edu/abaqusv2022/English/SIMACAEANLRefMap/simaanl-c-nonlineareqns.htm

Dassault Systèmes. (2022m). *Transferring results between Abaqus/Explicit and Abaqus/Standard*. Retrieved November 17, 2023, from https://docs.software.vt.edu/abaqusv2022/English/SIMACAEANLRefMap/simaanl-c-exptostd.htm

Dassault Systèmes. (2022n). *VUAMP*. Retrieved November 28, 2023, from https://docs.software.vt.edu/abaqusv2022/English/SIMACAESUBRefMap/simasub-c-vuamp.htm

den Hertog, D., Kleijnen, J. P. C., & Siem, A. Y. D. (2006). The correct Kriging variance estimated by bootstrapping. *The Journal of the Operational Research Society*, *57*(4), 400–409. https://doi.org/10.1057/palgrave.jors.2601997

Desboulets, L. D. D. (2018). A review on variable selection in regression analysis. *Econometrics*, *6*(4). https://doi.org/10.3390/econometrics6040045

Di Caprio, F., Cristillo, D., Saputo, S., Guida, M., & Riccio, A. (2019). Crashworthiness of wing leading edges under bird impact event. *Composite Structures*, *216*, 39–52. https://doi.org/10.1016/j.compstruct.2019.02.069

Efroymson, M. A. (1960). Multiple regression analysis. *Mathematical methods for digital computers*, 191–203.

Fang, J., Sun, G., Qiu, N., Kim, N. H., & Li, Q. (2017). On design optimization for structural crashworthiness and its state of the art. *Structural and Multidisciplinary Optimization*, *55*(3), 1091–1119. https://doi.org/10.1007/s00158-016-1579-y

Flight Global Image Archive. (2011). *Fokker F. 27 Maritime Enforcer Cutaway Drawing*. Retrieved November 14, 2023, from https://www.flightglobalimages.com/rights-managed/cutaways/civil-aviation-1949-present-cutaways/fokker-f-27-maritime-enforcer-cutaway-drawing-4577715.html

Forrester, A. I. J., & Jones, D. R. (2008). Global optimization of deceptive functions with sparse sampling. *12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, MAO*. https://doi.org/10.2514/6.2008-5996

Forrester, A. I. J., & Keane, A. J. (2009). Recent advances in surrogate-based optimization. *Progress in Aerospace Sciences*, *45*(1-3), 50–79. https://doi.org/10.1016/j.paerosci.2008.11.001

Forrester, A. I. J., Sóbester, A., & Keane, A. J. (2008). *Engineering design via surrogate modelling*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470770801

Friedman, J. H. (1991). Multivariate Adaptive Regression Splines. *The Annals of Statistics*, *19*(1), 1–141. https://doi.org/10.1214/aos/1176347963

Garrido-Merchán, E. C., & Hernández-Lobato, D. (2020). Dealing with categorical and integer-valued variables in Bayesian Optimization with Gaussian processes. *Neurocomputing*, *380*, 20–35. https://doi.org/10.1016/j.neucom.2019.11.004

Georgiadis, S., Gunnion, A. J., Thomson, R. S., & Cartwright, B. K. (2008). Bird-strike simulation for certification of the Boeing 787 composite moveable trailing edge. *Composite Structures*, *86*(1-3), 258–268. https://doi.org/10.1016/j.compstruct.2008.03.025

Goel, T., Haftka, R. T., Shyy, W., & Queipo, N. V. (2007). Ensemble of surrogates. *Structural and Multidisciplinary Optimization*, *33*(3), 199–216. https://doi.org/10.1007/s00158-006-0051-9

Goraj, Z., & Kustron, K. (2018). Review of current research trends in bird strike and hail impact simulations on wing leading edge. *Aircraft Engineering and Aerospace Technology*, *90*(4), 602–612. https://doi.org/10.1108/AEAT-02-2017-0053

Greg Lam Pak Ng. (2009). *Plane crash into Hudson River*. Retrieved November 1, 2023, from https://www.flickr.com/photos/22608787@N00/3200086900

Guida, M., Marulo, F., Belkhelfa, F. Z., & Russo, P. (2022). A review of the bird impact process and validation of the SPH impact model for aircraft structures. *Progress in Aerospace Sciences*, *129*. https://doi.org/10.1016/j.paerosci.2021.100787

Gutmann, H. M. (2001). A Radial Basis Function method for global optimization. *Journal of Global Optimization*, *19*, 201–227. https://doi.org/10.1023/A:1011255519438

Hedayati, R., & Sadighi, M. (2016). 6 - Finite element bird-strike modeling. In R. Hedayati & M. Sadighi (Eds.), *Bird strike* (pp. 113–159). Woodhead Publishing. https://doi.org/https://doi.org/10.1016/B978-0-08-100093-9.00006-6

Heimbs, S. (2011). Computational methods for bird strike simulations: A review. *Computers and Structures*, *89*(23-24), 2093–2112. https://doi.org/10.1016/j.compstruc.2011.08.007

Hwang, J. (2020). *Kriging with categorical or integer variables*. Retrieved February 16, 2023, from https://smt.readthedocs.io/en/latest/_src_docs/surrogate_models/krg.html

International Civil Aviation Organization. (2023). *2016-2021 Wildlife strike analyses (IBIS)*. Retrieved November 1, 2023, from https://www.icao.int/Aerodromes/ibis/Documents/03.%20EB%202023.30%20-%202016_2021%20WILDLIFE%20STRIKE%20ANALYSES/2016%20-%202021%20Wildlife%20Strike%20Analyses%20(IBIS)%20-%20EN.pdf

Jin, R., Chen, W., & Sudjianto, A. (2005). An efficient algorithm for constructing optimal design of computer experiments. *Journal of Statistical Planning and Inference*, *134*(1), 268–287. https://doi.org/10.1016/j.jspi.2004.02.014

Joby Aviation. (2023). *Joby Delivers First eVTOL Aircraft to Edwards Air Force Base Ahead of Schedule*. Retrieved November 1, 2023, from https://www.jobyaviation.com/news/joby-delivers-first-evtol-edwards/

Johnson, M., Moore, L., & Ylvisaker, D. (1990). Minimax and maximin distance designs. *Journal of Statistical Planning and Inference*, *26*(2), 131–148. https://doi.org/10.1016/0378-3758(90)90122-B

Jones, D. R. (2001). A taxonomy of global optimization methods based on response surfaces. *Journal of Global Optimization*, *21*, 345–383.

Jones, D. R., Schonlau, M., & Welch, W. J. (1998). Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global Optimization*, *13*, 455–492. https://doi.org/10.1023/A:1008306431147

Joseph, V. R., Hung, Y., & Sudjianto, A. (2008). Blind Kriging: A new method for developing metamodels [031102]. *Journal of Mechanical Design*, *130*(3). https://doi.org/10.1115/1.2829873

Khatouri, H., Benamara, T., Breitkopf, P., Demange, J., & Feliot, P. (2020). Constrained multi-fidelity surrogate framework using bayesian optimization with non-intrusive reduced-order basis. *Advanced Modeling and Simulation in Engineering Sciences*, *7*, 43. https://doi.org/10.1186/s40323-020-00176-z

Klosak, M., Santiago, R., Jankowiak, T., Bendarma, A., Rusinek, A., & Bahi, S. (2021). The Influence of Temperature in the Al 2024-T3 Aluminum Plates Subjected to Impact: Experimental and Numerical Approaches. *Materials*, *14*(15). https://doi.org/10.3390/ma14154268

Koch, P. N., Simpson, T. W., Allen, J. K., & Mistree, F. (1999). Statistical approximations for multidisciplinary design optimization: The problem of size. *Journal of Aircraft*, *36*(1), 275–286. https://doi.org/10.2514/2.2435

Lanzi, L., Bisagni, C., & Ricci, S. (2004). Crashworthiness optimization of helicopter subfloor based on decomposition and global approximation. *Structural and Multidisciplinary Optimization*, *27*(5), 401–410. https://doi.org/10.1007/s00158-004-0394-z

Lesuer, D. R. (2009). *Experimental Investigations of Material Models for Ti-6Al-4V Titanium and 2024-T3 Aluminum* (tech. rep.). U.S. Department of Transportation, Federal Aviation Adminitration.

Liu, J., Li, Y., Yu, X., Gao, X., & Liu, Z. (2018). Design of aircraft structures against threat of bird strikes. *Chinese Journal of Aeronautics*, *31*(7), 1535–1558. https://doi.org/10.1016/j.cja.2018.05.004

Liu, J., Li, Y., Yu, X., Tang, Z., Gao, X., Lv, J., & Zhang, Z. (2017). A novel design for reinforcing the aircraft tail leading edge structure against bird strike. *International Journal of Impact Engineering*, *105*, 89–101. https://doi.org/10.1016/j.ijimpeng.2016.12.017

Lualdi, P., Sturm, R., & Siefkes, T. (2023). A multi-fidelity successive response surface method for crashworthiness optimization problems. *Applied Sciences*, *13*(20). https://doi.org/10.3390/app132011452

Martins, J. R., & Lambe, A. B. (2013). Multidisciplinary design optimization: A survey of architectures. *AIAA Journal*, *51*(9), 2049–2075. https://doi.org/10.2514/1.J051895

Marulo, F., & Guida, M. (2014). Design criteria for birdstrike damage on windshield. *Advances in Aircraft and Spacecraft Science*, *1*, 2014. https://doi.org/10.12989/aas.2014.1.2.233

McCarthy, M. A., Xiao, J. R., McCarthy, C. T., Kamoulakos, A., Ramos, J., Gallard, J., & Melito, V. A. (2004). Modelling of bird strike on an aircraft wing leading edge made from fibre metal laminates – part 2: Modelling of impact with SPH bird model. *Applied Composite Materials*, *11*, 317–340. https://doi.org/10.1023/B:ACMA.0000037134.93410.c0

McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, *21*(2), 239–245. https://doi.org/10.2307/1268522

McNaughtan, I. I. (1972). *The design of leading edge and intake wall structure to resist bird impact*. Royal Aircraft Technical Report N. RAE-TR-72056, vol. 72056.

Mol, K., Salem, S. C., & Sadiq, A. (2022). Crashworthiness enhancement of aluminum alloy used for leading edges of wing and empennage structures. *Journal of Aerospace Engineering*, *35*(6), 04022079. https://doi.org/10.1061/(ASCE)AS.1943-5525.0001477

Naceur, H., Guo, Y. Q., & Ben-Elechi, S. (2006). Response surface methodology for design of sheet forming parameters to control springback effects. *Computers and Structures*, *84*(26-27), 1651–1663. https://doi.org/10.1016/j.compstruc.2006.04.005

Niu, C. (1988). *Airframe structural design: Practical design information and data on aircraft structures*. Conmilit Press.

O'Hagan, A. (2006). Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering and System Safety*, *91*(10-11), 1290–1300. https://doi.org/10.1016/j.ress.2005.11.025

Ollar, J., Jones, R., & Toropov, V. (2017). Sub-space metamodel-based multidisciplinary optimization of an aircraft wing subjected to bird strike. In *58th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. https://doi.org/10.2514/6.2017-1307

Ollar, J., Toropov, V., & Jones, R. (2017). Sub-space approximations for MDO problems with disparate disciplinary variable dependence. *Structural and Multidisciplinary Optimization*, *55*(1), 279–288. https://doi.org/10.1007/s00158-016-1496-0

Orban, F. (2011). Damping of materials and members in structures. *Journal of Physics: Conference Series*, *268*(1), 012022. https://doi.org/10.1088/1742-6596/268/1/012022

Pahange, H., & Abolbashari, M. H. (2016). Mass and performance optimization of an airplane wing leading edge structure against bird strike using Taguchi-based grey relational analysis. *Chinese Journal of Aeronautics*, *29*(4), 934–944. https://doi.org/10.1016/j.cja.2016.06.008

ParaPy B.V. (2022). *Parapy homepage*. Retrieved March 28, 2023, from https://parapy.nl/

Park, J. S. (1994). Optimal Latin-hypercube designs for computer experiments. *Journal of Statistical Planning and Inference*, *39*(1), 95–111. https://doi.org/10.1016/0378-3758(94)90115-5

Qian, P. Z. G. (2012). Sliced Latin Hypercube Designs. *Journal of the American Statistical Association*, *107*(497), 393–399. https://doi.org/10.1080/01621459.2011.644132

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

Sacks, J., Welch, W. J., Mitchell, T. J., & Wynn, H. P. (1989). Design and Analysis of Computer Experiments. *Statistical Science*, *4*(4), 409–435. https://doi.org/10.1214/ss/1177012413

Saltelli, A., & Sobol', M. (1995). Sensitivity analysis for nonlinear mathematical models: Numerical experience. *Mathematical models and computer experiment*, *7*(11), 16–28.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2007). Variance-based methods. In *Global sensitivity analysis. the primer* (pp. 155–182). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470725184.ch4

Saves, P., Lafage, R., Bartoli, N., Diouane, Y., Bussemaker, J. H., Lefebvre, T., Hwang, J. T., Morlier, J., & Martins, J. R. R. A. (2023). SMT 2.0: A surrogate modeling toolbox with a focus on hierarchical and mixed variables gaussian processes. *ArXiv preprint*. https://doi.org/https://doi.org/10.48550/arXiv.2305.13998

Schonlau, M., & Welch, W. J. (2006). Screening the input variables to a computer model via analysis of variance and visualization. In A. Dean & S. Lewis (Eds.), *Screening: Methods for experimentation in industry, drug discovery, and genetics* (pp. 308–327). Springer New York. https://doi.org/10.1007/0-387-28014-6_14

Schonlau, M., Welch, W. J., & Jones, D. R. (1998). Global versus local search in constrained optimization of computer models. *Lecture Notes-Monograph Series*, *34*, 11–25.

Schuhmacher, G., Murra, I., Wang, L., Laxander, A., O'Leary, O. J., & Herold, M. (2002). Multidisciplinary design optimization of a regional aircraft wing box. *9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*. https://doi.org/10.2514/6.2002-5406

Shan Ba, W. R. M., & Brenneman, W. A. (2015). Optimal Sliced Latin Hypercube Designs. *Technometrics*, *57*(4), 479–487. https://doi.org/10.1080/00401706.2014.957867

Siemann, M. H., & Ritt, S. A. (2019). Novel particle distributions for SPH bird-strike simulations. *Computer Methods in Applied Mechanics and Engineering*, *343*, 746–766. https://doi.org/10.1016/j.cma.2018.08.044

Simpson, T. W., Peplinski, J. D., Koch, P. N., & Allen, J. K. (2001). Metamodels for computer-based engineering design: Survey and recommendations. *Engineering with Computers*, *17*, 129–150. https://doi.org/10.1007/PL00007198

Sóbester, A., Leary, S. J., & Keane, A. J. (2004). A parallel updating scheme for approximating and optimizing high fidelity computer simulations. *Structural and Multidisciplinary Optimization*, *27*(5), 371–383. https://doi.org/10.1007/s00158-004-0397-9

Soch, J. (2020). *Proof: Conditional distributions of the multivariate normal distribution*. Retrieved February 23, 2023, from https://statproofbook.github.io/P/mvn-cond

The GPflow Contributors. (2023). *gpflow.optimizers.Scipy*. Retrieved January 1, 2024, from https://gpflow.github.io/GPflow/develop/api/gpflow/optimizers/index.html#module-gpflow.optimizers

The SciPy community. (2022). *scipy.stats.qmc.discrepancy*. Retrieved November 28, 2023, from https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.qmc.discrepancy.html

The SciPy community. (2023). *scipy.integrate.quad*. Retrieved December 26, 2023, from https://docs.scipy.org/doc/scipy/reference/generated/scipy.integrate.quad.html

The Trieste Contributors. (2020a). *Trieste documentation*. Retrieved November 2, 2023, from https://secondmind-labs.github.io/trieste/1.2.0/index.html

The Trieste Contributors. (2020b). *trieste.acquisition.optimizer*. Retrieved December 14, 2023, from https://secondmind-labs.github.io/trieste/1.2.0/autoapi/trieste/acquisition/optimizer/index.html#trieste.acquisition.optimizer.generate_continuous_optimizer

The Trieste Contributors. (2020c). *trieste.models.gpflow*. Retrieved January 1, 2024, from https://secondmind-labs.github.io/trieste/1.2.0/autoapi/trieste/models/gpflow/index.html

van den Berg, T., & van der Laan, T. (2021). A multidisciplinary modeling system for structural design applied to aircraft moveables. In *AIAA AVIATION 2021 FORUM*. https://doi.org/10.2514/6.2021-3079

van der Vaart, J., & Muhammad, H. (1983). *Report LR-394: Static Longitudinal Stability and Control Characteristics of the Fokker F27 'Friendship' Calculated by Simple Handbook Methods* (tech. rep.). Delft University of Technology, Department of Aerospace Engineering.

Vershinin, V. V. (2015). Validation of metal plasticity and fracture models through numerical simulation of high velocity perforation. *International Journal of Solids and Structures*, *67-68*, 127–138. https://doi.org/https://doi.org/10.1016/j.ijsolstr.2015.04.007

Viana, F. A., Gogu, C., & Goel, T. (2021). Surrogate modeling: Tricks that endured the test of time and some recent developments. *Structural and Multidisciplinary Optimization*, *64*(5), 2881–2908. https://doi.org/10.1007/s00158-021-03001-2

Viana, F. A., Haftka, R. T., & Steffen, V. (2009). Multiple surrogates: How cross-validation errors can help us to obtain the best predictor. *Structural and Multidisciplinary Optimization*, *39*(4), 439–457. https://doi.org/10.1007/s00158-008-0338-0

Viana, F. A., Simpson, T. W., Balabanov, V., & Toropov, V. (2014). Metamodeling in multidisciplinary design optimization: How far have we really come? *AIAA Journal*, *52*(4), 670–690. https://doi.org/10.2514/1.J052375

Wang, G. G. (2003). *Adaptive Response Surface Method Using Inherited Latin Hypercube Design Points*.

Welch, W. J., Buck, R. J., Sacks, J., Wynn, H. P., Mitchell, T. J., & Morris, M. D. (1992). Screening, predicting, and computer experiments. *Technometrics*, *34*(1), 15–25. https://doi.org/10.1080/00401706.1992.10485229

Wilbeck, J. S. (1978). *Impact Behaviour of Low Strength Projectiles, Technical Report AFML-TR-77-134*. Air Force Materials Laboratory, Wright-Patterson Air Force Base, Ohio, 45433, USA.

Yu, Z., Xue, P., Yao, P., & Zahran, M. S. (2019). Analytical determination of the critical impact location for wing leading edge under birdstrike. *Latin American Journal of Solids and Structures*, *16*(16), e152. https://doi.org/10.1590/1679-78255352

Zerpa, L. E., Queipo, N. V., Pintos, S., & Salager, J. L. (2005). An optimization methodology of alkaline-surfactant-polymer flooding processes using field scale numerical simulation and multiple surrogates. *Journal of Petroleum Science and Engineering*, *47*(3-4), 197–208. https://doi.org/10.1016/j.petrol.2005.03.002

Zhao, M., & Li, J. (2011). Efficiency metallic leading edge structure bird strike resistant design. *Advanced Materials Research*, *338*, 84–89. https://doi.org/10.4028/www.scientific.net/AMR.338.84

Zhou, X. J., Ma, Y. Z., & Li, X. F. (2011). Ensemble of surrogates with recursive arithmetic average. *Structural and Multidisciplinary Optimization*, *44*(5), 651–671. https://doi.org/10.1007/s00158-011-0655-6

Zhou, Y.-D., Fang, K.-T., & Ning, J.-H. (2013). Mixture discrepancy for quasi-random point sets. *Journal of Complexity*, *29*(3), 283–301. https://doi.org/https://doi.org/10.1016/j.jco.2012.11.006

A

ANOVA on the ordinary Kriging with the Ornstein-Uhlenbeck covariance function

The Ornstein-Uhlenbeck correlation function is nothing other than the exponential kernel with a unity power, $p_l = 1, \forall l \in \overline{1,k}$, i.e. it depends solely on the absolute distance between two points in a given direction. Similar to equation 3.31:

$$\hat{f}(\boldsymbol{x}) = \hat{\mu} + \sum_{i=1}^{n} b_i \prod_{j=1}^{k} \text{exp}\left(-\theta_j \left| x_j - x_j^{(i)} \right|\right) \tag{A.1}$$

The marginal effect becomes, paralleling 3.35:

$$\hat{f}_e(\boldsymbol{x_e}) = \hat{\mu} + \sum_{i=1}^{n} b_i \prod_{j \notin \boldsymbol{e}} \frac{1}{x_j^u - x_j^l} \int_{\chi_j} \text{exp}\left(-\theta_j \left| x_j - x_j^{(i)} \right|\right) dx_j \prod_{j \in \boldsymbol{e}} \text{exp}\left(-\theta_j \left| x_j - x_j^{(i)} \right|\right) \tag{A.2}$$

Where the integration is carried out as:

$$\begin{aligned}
\int_{\chi_j} \text{exp}\left(-\theta_j \left| x_j - x_j^{(i)} \right|\right) dx_j &= \int_{x_j^l}^{x_j^{(i)}} + \int_{x_j^{(i)}}^{x_j^u} \text{exp}\left(-\theta_j \left| x_j - x_j^{(i)} \right|\right) dx_j \\
&= \int_{x_j^l}^{x_j^{(i)}} \text{exp}\left(-\theta_j \left(x_j^{(i)} - x_j\right)\right) dx_j + \int_{x_j^{(i)}}^{x_j^u} \text{exp}\left(-\theta_j \left(x_j - x_j^{(i)}\right)\right) dx_j \\
\left(t \doteq -\theta_j \left(x_j^{(i)} - x_j\right) \Rightarrow dt = \theta_j dx_j\right) &= \frac{1}{\theta_j} \int_{-\theta_j \left(x_j^{(i)} - x_j^l\right)}^{0} \text{exp}(t) dt + \frac{1}{\theta_j} \int_{0}^{-\theta_j \left(x_j^u - x_j^{(i)}\right)} \text{exp}(-t) dt \\
&= \frac{1}{\theta_j} \left[2 - \text{exp}\left(-\theta_j \left(x_j^{(i)} - x_j^l\right)\right) - \text{exp}\left(-\theta_j \left(x_j^u - x_j^{(i)}\right)\right)\right]
\end{aligned} \tag{A.3}$$

The overall mean transforms in a similar manner to equation 3.37:

$$\mu_0 = \hat{\mu} + \sum_{i=1}^{n} b_i \prod_{j=1}^{k} \frac{1}{x_j^u - x_j^l} \int_{\chi_j} \text{exp}\left(-\theta_j \left| x_j - x_j^{(i)} \right|\right) dx_j \tag{A.4}$$

The corrected effects are unchanged. The only remaining term is the total variance of the predictor, which is calculated equivalent to equation 3.40:

$$Var[\hat{f}] = -(\hat{\mu} - \mu_0)^2 + \sum_{i,p=1}^{n} b_i b_p \prod_{j=1}^{k} \frac{1}{x_j^u - x_j^l} \cdot \int_{\chi_j} \text{exp}\left(-\theta_j \left[\left|x_j - x_j^{(i)}\right| + \left|x_j - x_j^{(p)}\right|\right]\right) dx_j \tag{A.5}$$

Supposing that $x_j^{(i)} \leq x_j^{(p)}$, the integral on the right is computed analytically as follows:

$$\begin{aligned}
&\int_{\chi_j} \text{exp}\left(-\theta_j \left[\left|x_j - x_j^{(i)}\right| + \left|x_j - x_j^{(p)}\right|\right]\right) dx_j = \\
&\quad = \int_{x_j^l}^{x_j^{(i)}} + \int_{x_j^{(i)}}^{x_j^{(p)}} + \int_{x_j^{(p)}}^{x_j^u} \text{exp}\left(-\theta_j \left[\left|x_j - x_j^{(i)}\right| + \left|x_j - x_j^{(p)}\right|\right]\right) dx_j \\
&\quad = \int_{x_j^l}^{x_j^{(i)}} \text{exp}\left(-\theta_j \left(x_j^{(i)} + x_j^{(p)} - 2x_j\right)\right) dx_j \\
&\quad + \int_{x_j^{(i)}}^{x_j^{(p)}} \text{exp}\left(-\theta_j \left(x_j^{(p)} - x_j^{(i)}\right)\right) dx_j \\
&\quad + \int_{x_j^{(p)}}^{x_j^u} \text{exp}\left(-\theta_j \left(2x_j - x_j^{(i)} - x_j^{(p)}\right)\right) dx_j \\
\left(t \doteq -\theta_j \left(x_j^{(i)} + x_j^{(p)} - 2x_j\right)\right) &= \frac{1}{2\theta_j} \int_{-\theta_j\left(x_j^{(i)} + x_j^{(p)} - 2x_j^l\right)}^{-\theta_j\left(x_j^{(p)} - x_j^{(i)}\right)} \text{exp}(t) dt \\
&\quad + \left(x_j^{(p)} - x_j^{(i)}\right) \text{exp}\left(-\theta_j \left(x_j^{(p)} - x_j^{(i)}\right)\right) \\
&\quad + \frac{1}{2\theta_j} \int_{-\theta_j\left(x_j^{(i)} - x_j^{(p)}\right)}^{-\theta_j\left(x_j^{(i)} + x_j^{(p)} - 2x_j^u\right)} \text{exp}(-t) dt \\
&\quad = \frac{1}{2\theta_j} \left[\text{exp}\left(-\theta_j \left(x_j^{(p)} - x_j^{(i)}\right)\right) - \text{exp}\left(-\theta_j \left(x_j^{(i)} + x_j^{(p)} - 2x_j^l\right)\right)\right. \\
&\quad \left. - \text{exp}\left(-\theta_j \left(2x_j^u - x_j^{(i)} + x_j^{(p)}\right)\right) + \text{exp}\left(-\theta_j \left(x_j^{(p)} - x_j^{(i)}\right)\right)\right] \\
&\quad + \left(x_j^{(p)} - x_j^{(i)}\right) \text{exp}\left(-\theta_j \left(x_j^{(p)} - x_j^{(i)}\right)\right) \\
&\quad = \frac{1}{2\theta_j} \left\{\left[2\theta_j \left(x_j^{(p)} - x_j^{(i)}\right) + 2\right] \text{exp}\left(-\theta_j \left(x_j^{(p)} - x_j^{(i)}\right)\right)\right. \\
&\quad - \text{exp}\left(-\theta_j \left(x_j^{(i)} + x_j^{(p)} - 2x_j^l\right)\right) \\
&\quad \left. - \text{exp}\left(-\theta_j \left(2x_j^u - x_j^{(i)} + x_j^{(p)}\right)\right)\right\}
\end{aligned} \tag{A.6}$$

B

One-stage Bayesian optimization approaches

An inherent disadvantage of two-stage approaches is that they can be easily tricked by deceiving functions or unfortunate DoEs, leading to the algorithm completely ignoring certain regions, thus locating local, not global optima. The main reason is that the variance can be severely underestimated, rendering the infill criterion useless. One-stage approaches aim to solve this issue. The infill points are no longer computed based solely on the prediction error estimates of the surrogate model; as Forrester and Keane, 2009 put it, in one-stage strategies, the infill criterion is used to calculate the surrogate model, not the other way around.

B.1. Goal seeking

The first approach mentioned by both Forrester and Keane, 2009 and Jones, 2001 aims at counteracting the deficiencies of the expected improvement in the case of a deceptive function, particularly when there are many responses with the same value. The Kriging predictor would see that region as having a constant value, which, indeed, is the most likely guess, but it would also estimate a null error. As mentioned earlier, the expected improvement strategy would completely disregard that region, and no infill points would be placed there.

The goal-seeking strategy is then based on the user supplying a "goal" y^g for the objective function. This hypothesis of the objective function yielding the response y^g is then tested, by computing the likelihood of the response surface passing through the point $(\boldsymbol{x}^g, y^g)$; a full derivation of the likelihood can be found in Soch, 2020:

$$\mathcal{L}_{cond} = \frac{1}{(2\pi\sigma^2)^{n/2}|\boldsymbol{C}|^{1/2}} \text{exp}\left[-\frac{(\boldsymbol{y}-\boldsymbol{m})^T\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{m})}{2\sigma^2}\right] \tag{B.1a}$$

$$\boldsymbol{C} = \boldsymbol{R} - \boldsymbol{r}\boldsymbol{r}^{\boldsymbol{T}} \tag{B.1b}$$

$$\boldsymbol{m} = \boldsymbol{1}\mu + \boldsymbol{r}(y^g - \mu) \tag{B.1c}$$

In the above equation, the notations from section 3.3 have been preserved: $\boldsymbol{R}$ is the correlation matrix, $\boldsymbol{r}$ is the vector of correlations between the unknown $\boldsymbol{x}^g$ and the other sampled points, while $\boldsymbol{R} - \boldsymbol{r}\boldsymbol{r}^{\boldsymbol{T}}$ is the *conditional* correlation matrix.

Taking a look at equation B.1, it becomes clear that $\boldsymbol{x}^g, \mu, \sigma^2, \theta_l, p_l, l \in \overline{1,k}$ in Kriging are all variables to be optimized in the maximization of the conditional likelihood. This optimization though is not very straightforward, and Forrester and Keane, 2009 mention that genetic algorithms should be used. The resulting $\boldsymbol{x}^g$ is the next infill point. A great advantage of this strategy is that the Kriging hyperparameters are not fixed, therefore even if the initial sampling is deceptive, a minimum can be found nonetheless. However, the goal, which is a user input, is often difficult to estimate even in the case of a more basic function such as the mass of a structure, especially if there are many constraints to be taken into account. An overly stringent goal would again promote excessive exploration. Finding a goal for a highly non-linear black-box function would become an even more arduous task.

As outlined by Forrester and Keane, 2009, if the goal is too low or optimistic, too much exploration will be conducted; conversely, if the goal is too high or pessimistic, too much exploitation is going to be carried out. One way to overcome this, inspired by the suggestion of Jones, 2001 with respect to the two-stage probability of improvement approach, would be to use several values for y^g of different magnitudes, which would balance exploration and exploitation. Gutmann, 2001 proposes an even better method: starting again from multiple goals, the infill points are chosen where there seem to be clusters of $\boldsymbol{x}^g$. While Gutmann used it for RBF surrogate models, Forrester and Keane, 2009 and Jones, 2001 also consider it useful for Kriging. After all, both models are based on Gaussian processes. However, a strategy that conducts this search automatically would solve this problem, and this is precisely what is going to be discussed in the next section.

B.2. The conditional lower bound

The present approach is quite similar to the goal-seeking method, but it tries to get rid of the user input y^g. As mentioned before, if the value of the goal is too optimistic and it is much lower than an appropriate statistical lower bound of $\hat{f} - s$, for example, the conditional likelihood would become extremely low. The conditional lower bound approach, however, abandons the goals which are too unlikely by conducting a so-called *likelihood ratio test* (Forrester and Jones, 2008). These ratio tests are usually used in statistics to test whether hypotheses with more degrees of freedom are significantly better than those with less. The distribution of the likelihood ratios can be approximated via the chi-square χ^2 cumulative distributions, which are the distributions of the square root sum of squares of a set of normally-distributed independent random variables.

To simplify things in the Kriging model, Forrester and Jones, 2008 have fixed the hyperparameters $p_l = 2, \forall l \in \overline{1,k}$; only $\boldsymbol{\theta}$ remains unknown. If $\mathcal{L}_0$ denotes the likelihood of the response y^g being achieved with the $\boldsymbol{\theta}$ parameters obtained from the estimation only with the current set of training points, the likelihood ratio test will be:

$$\Lambda = 2 \ln \frac{\mathcal{L}_0}{\mathcal{L}_{cond}} < \chi^2_{critical}(limit, k) \tag{B.2}$$

In the equation above, k is the number of dimensions, and the limit is tied to the desired confidence limit. For instance, Forrester and Keane, 2009 illustrate this method working on a one-dimensional problem with $k = 1$, the limit being chosen as 0.975 to yield a confidence limit of 0.95, which is the lower bound. The values of the distribution for different DOFs and the confidence limits can be easily retrieved from a look-up table. The only input supplied by the user is this limit, which is not problem dependent like the value of y^g. It could again be quite complicated to find this (although not as difficult as guessing a goal), but one could again try several values in a similar manner as Gutmann, 2001. If the likelihood ratio test fails, then the chosen y^g is dismissed.

Forrester and Jones, 2008 propose to couple this method with the expected improvement approach, thus retaining the useful properties of the latter. The optimization problem, which essentially yields the next infill point, will then be:

$$\begin{aligned} \text{maximize} \quad & E[I(\boldsymbol{x}, y^g)] \\ \text{with respect to} \quad & \boldsymbol{x}, y^g, \boldsymbol{\theta} \\ \text{subject to} \quad & 2 \ln \frac{\mathcal{L}_0}{\mathcal{L}_{cond}(\boldsymbol{x}, y^g, \boldsymbol{\theta})} < \chi^2(\text{erf}(1/\sqrt{2}), k) \end{aligned} \tag{B.3}$$

What this optimization essentially does is to find the next sampling point x such that a *likely* goal y^g is achieved and the expected improvement is maximized, while hopefully eliminating the problems associated with deceptive functions. The error function term simply implies that the confidence limit is set to one standard deviation. However, the implementation is rather problematic: the direct implementation of equation B.3 has highly multi-modal objective and constraint landscapes. To counteract this problem, Forrester and Jones, 2008 propose solving the following simplified, nested problem:

$$\underset{\boldsymbol{x}}{max} \qquad E\left[I\left(\boldsymbol{x}, \left\{\begin{array}{cc} \underset{y^g,\boldsymbol{\theta}}{max} & \hat{f}(\boldsymbol{x}) - y^g \\ \text{subject to} & 2\,\text{ln}\,\frac{\mathcal{L}_0}{\mathcal{L}_{cond}(\boldsymbol{x},y^g,\boldsymbol{\theta})} < \chi^2(\text{erf}(1/\sqrt{2}),k) \end{array}\right\}\right)\right] \tag{B.4}$$

Both optimization functions have been solved by multi-start sequential quadratic programming (SQP). While genetic algorithms have also been cited as being possible, the SQP algorithm has been found to yield better optima from the experience of Forrester and Jones, 2008.

All in all, one-stage approaches seem to solve the inherent problems of two-stage strategies, which are dependent on good estimates of the variance, which is most often an underestimator. Two-stage methods can also be easily tricked by deceptive functions and unfortunate initial samplings, which cannot be foreseen in the case of expensive, black-box problems. The last approach presented is the most promising one, but it is also the most computationally intensive: at each iteration of the optimization problem, the conditional likelihood computation requires the inverse of an $n \times n$ matrix, and $3k+1$, or $2k+1$, if the $\boldsymbol{p}$ hyperparameter is fixed, variables need to be found. However, in the case of expensive functions to be modelled, the number of training points n will not be too large, and the optimization problem would probably be much faster than obtaining the response of at a single data point. Moreover, if parallel computing is available, the responses at multiple infill points can be obtained at the same time, based, for instance, on the Kriging believer method.

C

Variational Gaussian processes. Binary classification constraints

As mentioned in section 4.2, a simple constraint enforcing no rupture of the leading edge skin has been attempted in the first weeks of study as an exercise, after the five-bay wing model had been decided on. A rupture is assumed to occur once an element is deleted, although it is recognized that this procedure is mesh dependent. In order to accelerate the optimization process, the bird strike dynamic analysis is terminated immediately once this happens, while a design is deemed feasible if no element is deleted until the total internal energy of the wing structure becomes relatively constant. The stabilization of the total internal energy, which comprises the energies related to elastic and artificial strains, and plastic and damage dissipation, is needed to ensure that the initial kinetic energy of the bird has been transferred to the structure. Once stabilization occurs, it can be safely assumed that damage will no longer occur, as the energy transfer is complete without dissipating any through damage.

This formulation accelerates the data collection process as the analyses are terminated immediately once rupture occurs. However, the extent of the damage cannot be assessed, yielding a binary response: the structure is either feasible or not. Due to the nature of the chosen constraint function, the Kriging metamodel, which assumes a continuous response, is no longer feasible. Fortunately, an adaptation of the Kriging surrogate to allow classification exists, which can be found under the name of "Variational Gaussian processes". In the present appendix, the binary classifier will be briefly presented theoretically, followed by an application to the no rupture constraint.

C.1. Binary classifiers

Gaussian process classification is covered in Rasmussen and Williams, 2005, p. 39, which is an excellent introduction to Gaussian processes and their applications in machine learning. Preserving the notations from section 3.3, let $\boldsymbol{X}$ be a vector of sampling points with corresponding $\boldsymbol{y}$ observations, which can either be 1 (pass) or 0 (fail). The goal of the binary classifier is to predict whether a new point is feasible or not, which hints that the surrogate model should evaluate the likelihood of the new point belonging to a category. However, if the Kriging metamodel was given binary data, the observations would give responses outside of the $[0, 1]$ interval. What is possible, instead, is to consider a Kriging/GP model, $\hat{f}(\boldsymbol{x})$, and to pass it through a sigmoid function g such that the resulting composition is guaranteed to take values between 0 and 1. For instance, Rasmussen and Williams, 2005 identifies the logit function for this scope, $g(x) = \frac{e^x}{1+e^x}$.

However, by passing the Kriging metamodel through the sigmoid function, not only is the surrogate altered, but the introduction of a new variable becomes necessary. Therefore, let $\boldsymbol{z}$ be a vector containing the n observations passed to the underlying model. Note that the values $z_i, \forall i \in \overline{1, n}$ are no longer 0 or 1; as a matter of fact, they do not have any practical meaning as they are not *direct* observations, and for this reason Rasmussen and Williams, 2005 coins them as *latent* variables. Nevertheless, Z will denote the random variable associated with $\boldsymbol{z}$. The underlying Kriging model still assumes a prior multivariate normal distribution, $Z \sim \mathcal{N}(\mu, \sigma^2)$; thus, its probability density function is going to be the same as equation 3.11:

$$p\left(\boldsymbol{z}, \mu, \sigma^2, \theta_l, p_l, l \in \overline{1,k}\right) = \frac{1}{(2\pi)^{n/2}\sigma^n |\boldsymbol{R}|^{1/2}} \text{exp}\left(-\frac{(\boldsymbol{z} - \boldsymbol{1}\mu)^T \boldsymbol{R}^{-1}(\boldsymbol{z} - \boldsymbol{1}\mu)}{2\sigma^2}\right) \tag{C.1}$$

The probability density function notation has been chosen in expression C.1 instead of the likelihood notation used in section 3.3 because Bayes' theorem will be invoked a bit later.

The correlation between $\boldsymbol{z}$ and $\boldsymbol{y}$ should now be discussed. As the final surrogate should be a binary classifier, the probability $p(\boldsymbol{y}|\boldsymbol{z})$, which describes the probability of the observations $y_i \in \{0,1\}, \forall i \in \overline{1,n}$ taking place given the underlying observations $\boldsymbol{z}$ which are passed to the Kriging model, should correspond to a Bernoulli distribution. Moreover, because the final surrogate should give the probability of a given sample to belong to one of the two categories, the probability should be:

$$p\left(y_i \middle| z_i\right) = \begin{cases} g(z_i), & y_i = 1 \\ 1 - g(z_i), & y_i = 0 \end{cases} \Rightarrow p\left(\boldsymbol{y}\middle|\boldsymbol{z}\right) = \prod_{i=1}^{n} p\left(y_i\middle|z_i\right) \tag{C.2}$$

However, the underlying model is not known yet. Therefore, in a manner similar to the one employed in section 3.3, the model hyperparameters should be estimated by maximizing the likelihood function in equation C.1. However, it is again reiterated that there are no direct observations of the Z variable, therefore it is desired to compute the likelihood of the model hyperparameters with respect to the direct observations $\boldsymbol{y}$, which, according to Bayes' theorem, is the following:

$$p(\boldsymbol{z}|\boldsymbol{y}) = \frac{p\left(\boldsymbol{y}\middle|\boldsymbol{z}\right)p(\boldsymbol{z})}{\int p\left(\boldsymbol{y}\middle|\boldsymbol{z}\right)p(\boldsymbol{z})d\boldsymbol{z}} \tag{C.3}$$

Both $p(\boldsymbol{z}|\boldsymbol{y})$ and $p(\boldsymbol{z})$ are a function of the model hyperparameters, although they were not included for conciseness. Theoretically, now, by maximizing the likelihood in equation C.3, the underlying Kriging model would be found and could be used to start making binary predictions by passing the prediction through the logit function. Unfortunately, the integration is not analytically tractable, therefore the likelihood cannot be computed analytically.

To solve this problem, variational inference is used, hence the name of the classifier surrogate model. It is proposed to approximate the likelihood $p(\boldsymbol{z}|\boldsymbol{y}, \mu, \sigma^2, \theta_l, p_l, l \in \overline{1,k})$ by another multivariate Gaussian distribution with mean μ_{var} and covariance matrix $\boldsymbol{R}_{var}$, $q(\mu_{var}, \boldsymbol{R}_{var})$. A measure of how close the two distributions are is given by the Kullback–Leibler divergence, and, by maximizing this similarity criterion, the posterior Gaussian distribution can be found out.

Suppose that the approximation has already been carried out, and let $\boldsymbol{x}^*$ denote a design point where a prediction, y^*, is needed. As stated before, the probability of the design falling in one of the classes is:

$$\begin{cases} p(y^* = 1) = & g(\hat{f}(\boldsymbol{x}^*)) \\ p(y^* = 0) = & 1 - g(\hat{f}(\boldsymbol{x}^*)) \end{cases} \tag{C.4}$$

On the other hand, the prediction of the underlying Kriging predictor, $z^* = \hat{f}(\boldsymbol{x}^*)$, is given by the posterior $z^* \sim p(z^*|\boldsymbol{y})$. However, only $p(\boldsymbol{z}|\boldsymbol{y})$ is known by its Gaussian approximate $q(\mu_{var}, \boldsymbol{R}_{var})$, and $p(z^*|\boldsymbol{y})$ results from taking the joint distribution $p(z^*, \boldsymbol{z}|\boldsymbol{y})$ and integrating out all latent observations:

$$\begin{aligned} p(z^*|\boldsymbol{y}) &= \int p(z^*, \boldsymbol{z}|\boldsymbol{y})d\boldsymbol{z} \\ \text{(chain rule)} &= \int p(z^*|\boldsymbol{z}, \boldsymbol{y}) \cdot p(\boldsymbol{z}|\boldsymbol{y})d\boldsymbol{z} \\ \text{(using the Gaussian approximate)} &= \int p(z^*|\boldsymbol{z}, \boldsymbol{y}) \cdot q(\mu_{var}, \boldsymbol{R}_{var})d\boldsymbol{z} \end{aligned} \tag{C.5}$$

The first term under the integral can be computed in a manner similar to equation 3.11, as the random variable Z has been assumed to be normally distributed. As a consequence, $\boldsymbol{z}^*$ and $\boldsymbol{z}$ follow the same multivariate normal distribution, paralleling $\boldsymbol{y}^*$ and $\boldsymbol{y}$ in section 3.3. However, no further details are going to be given about this predictor, as the purpose of this section is only to provide a brief introduction to Variational Gaussian processes.

While it is clear that such surrogate models are powerful tools, it is even more obvious that they cannot be used for variable ranking in the manner exposed in section 3.4. An interesting research study would be to investigate whether the analysis of variance procedure can be used on the underlying Kriging model, working directly with the latent variables. However, no progress has been made in this direction.

C.2. Application: no rupture constraint

As mentioned before, a two-variable Bayesian optimization problem has been formulated during the early stages of the study, which concerned the minimization of the weight of a wing structure, while ensuring that no rupture occurs in the leading edge. The two design variables were the thicknesses of the leading edge and nose ribs, $t_{LE}, t_{nose\ rib}$. The optimization problem can then be formulated as:

$$\begin{array}{cc} \text{minimize} & weight_{wing} \\ \text{with respect to} & t_{LE}, t_{nose\ rib} \\ \text{subject to} & n_{deleted\ elements} = 0 \end{array} \tag{C.6}$$

The ranges of the design variables have been adjusted for this problem, mainly because it was expected that a rather thick leading edge would be needed: $t_{nose\ rib} \in [0.8, 2.0], t_{LE} \in [4.0, 6.0]$. It is acknowledged, though, that the leading edge range could have been left smaller. All other thicknesses are given in the table below:

Component	Thickness [mm]
All other skin panel material zones	1.6
Front spar	2.0
Rear spar	1.8
Wingbox ribs	1.6
Stringers	1.2

Thanks to the rather inexpensive computations, a generous initial data of 20 points has been generated, and the exponential kernel has been chosen as the best option after splitting the set into training and validation samples. Figure C.1 shows the predictions of the VGP, which correspond to the probability of no rupture. Interestingly, the landscape is neither multimodal nor non-linear, and it seems that, no matter how thin the nose rib is, there is always a leading edge thickness which results in no rupture.

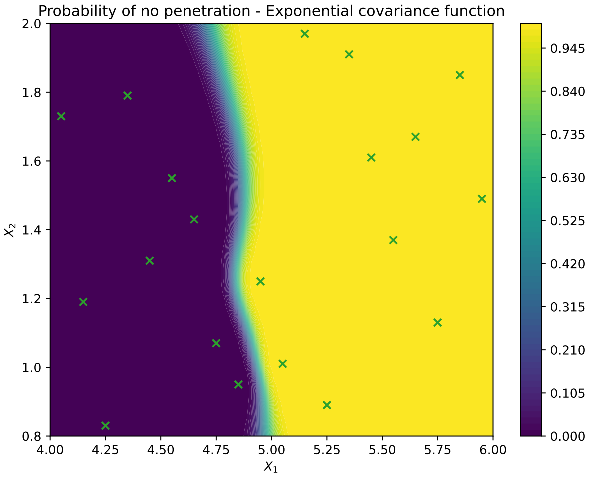


Figure C.1: Probability of no rupture versus leading edge and nose rib thicknesses.

Thus, the optimization process consists simply of finding this corresponding leading edge thickness. As the weight is known, the acquisition function for the Bayesian optimization procedure will be as follows:

$$\begin{aligned} \text{maximize} \quad & [weight_{min} - weight(t_{LE}, t_{nose\ rib})] \cdot P_{no\ rupture}(t_{LE}, t_{nose\ rib}) \\ \text{with respect to} \quad & t_{LE}, t_{nose\ rib} \end{aligned} \tag{C.7}$$

The first term is simply the weight improvement, while the second one is given by the VGP predictor itself, resulting in the same constrained expected improvement acquisition function proposed in section 3.5. A number of 40 sequential optimization steps has been requested. The final model predictions, together with the optimization points, illustrated with round dots, are depicted in figure C.2. As a first observation, the upper-central region has been altered due to the hyperparameter variation with additional training points.

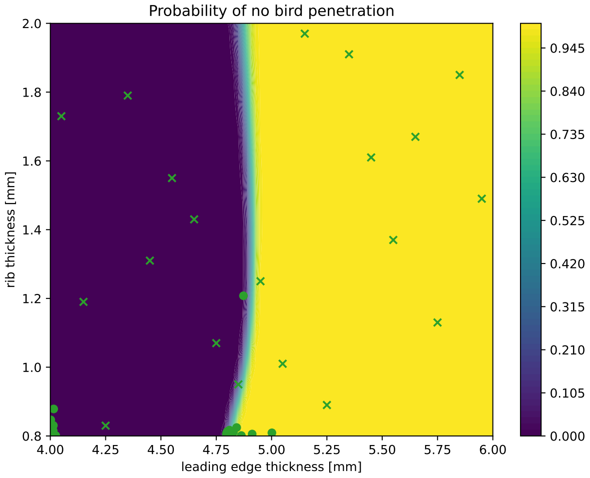


Figure C.2: Final model predictions for the no rupture constraint, after the optimization process.

The boundary between the feasible and unfeasible regions has been first exploited. As expected, all optimization points are gathering towards the line of minimum nose rib thickness. After exploiting that region, they switch to the lower-left corner, which is an unexplored region of minimum weight. However, the probability of no rupture is nearly null, resulting in a nearly null constrained improvement acquisition function, as depicted in figure C.3.

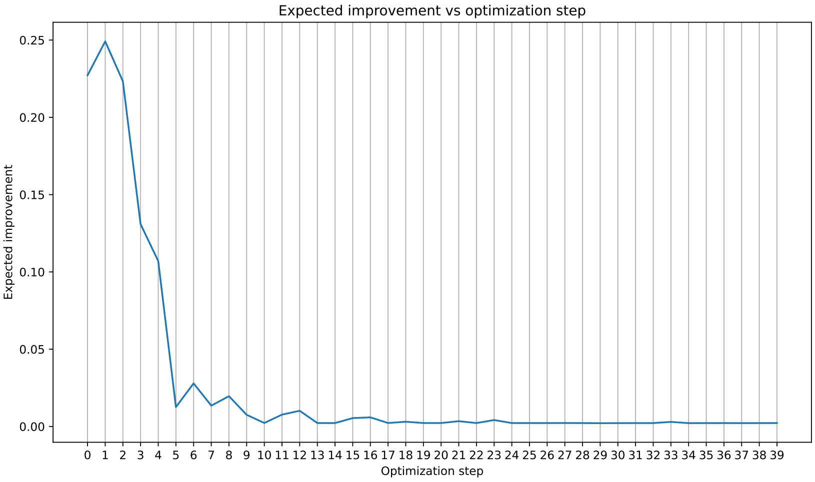


Figure C.3: The values of the acquisition function for every sequential optimization step.

It is obvious that this constraint alone cannot be used for the optimization process, as the nose rib thickness plays no important role. However, it did shed light on an important aspect: compliant nose ribs alleviate the energy absorption capabilities of the leading edge, and they accommodate large deformations of the skin without causing it to rupture. However, such thin nose ribs may have an adverse effect on the static strength of the damaged wing structure, which is another reason that promotes the inclusion of the residual strength requirement formulated in section 4.2.