



Delft University of Technology

Incorporating bird strike crashworthiness requirements within the design of wing structures

Ciobotia, R.I.; van der Laan, Ton; van de Waerdt, Wydo; Peeters, D.M.J.; Castro, Saullo G.P.

DOI

[10.1007/s00158-025-04158-w](https://doi.org/10.1007/s00158-025-04158-w)

Publication date

2025

Document Version

Final published version

Published in

Structural and Multidisciplinary Optimization

Citation (APA)

Ciobotia, R. I., van der Laan, T., van de Waerdt, W., Peeters, D. M. J., & Castro, S. G. P. (2025). Incorporating bird strike crashworthiness requirements within the design of wing structures. *Structural and Multidisciplinary Optimization*, 68(11), Article 236. <https://doi.org/10.1007/s00158-025-04158-w>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.



Incorporating bird strike crashworthiness requirements within the design of wing structures

Raluca-Ioana Ciobotia¹ · Ton van der Laan² · Wydo van de Waerd³ · Daniël Peeters¹ · Saullo G. P. Castro¹

Received: 5 June 2025 / Revised: 22 August 2025 / Accepted: 15 September 2025
© The Author(s) 2025

Abstract

The present study introduces an automated multidisciplinary optimization (MDO) workflow that, for the first time, couples an explicit dynamic bird strike analysis with a post-impact static stress check. This joint problem is solved during preliminary wing sizing by integrating batch Bayesian optimization on Kriging surrogates with a variance-based variable screening procedure. The optimization problem comprises 19 thickness design variables and two highly non-linear constraints, imposing a maximum leading edge penetration and a maximum post-impact front spar stress while minimizing wing mass. The workflow is demonstrated on a five-bay metallic wing segment, yielding a 43% weight saving over the best-performing design during initial data generation while respecting CS 25.631 crashworthiness limits. Results demonstrated substantial computational savings by variable screening and highlighted the necessity of the stress constraint, as designs satisfying only the penetration depth requirement could still experience critical post-impact stress levels.

Keywords Crashworthiness · Bird strike · Design · Bayesian optimization · Kriging · Variable screening

1 Introduction

Often perceived as rare anomalies, bird strikes are in reality a significant threat to aviation safety. Goraj and Kustron (2018), who conducted a review on research trends in both bird strike and hail impact simulations on wing leading edges, estimate that one bird strike occurs every 2,000 flights. 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. Between 2016 and 2021, the International Civil Aviation Organization (2023) concluded

that 3% of such incidents resulted in damage to aircraft components, 14% of which affected the wings.

Taking into account these three estimates, the chance of wing damage from a bird strike stands at a seemingly negligible 0.00021%. However, the sheer volume of air traffic amplifies this risk. In 2024 alone, Europe recorded 10.7 million flights (Eurocontrol 2025), suggesting that approximately 22 aircraft may have experienced wing damage due to bird strikes. These approximate calculations underscore that bird strikes are not mere statistical outliers, but persistent threats requiring robust design.

Acknowledging this hazard, the European Union Aviation Safety Agency (EASA) has established stringent structural requirements to safeguard aircraft integrity, such as CS 25.631:

“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 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.”

Communicated by Lei Wang.

✉ Raluca-Ioana Ciobotia
R.I.Ciobotia@tudelft.nl

✉ Saullo G. P. Castro
S.G.P.Castro@tudelft.nl

¹ Department of Aerospace Structures and Materials, Delft University of Technology, 2629HS Delft, The Netherlands

² Fokker Aerostructures B.V, Anthony Fokkerweg 4, 3351NL Papendrecht, The Netherlands

³ Fokker Aerostructures B.V, Edisonstraat 1, 7903AN Hoozeveen, The Netherlands

Designing aircraft structures to withstand bird strikes while simultaneously meeting diverse performance criteria requires a careful equilibrium of competing design factors. Compliant structures are ideal for absorbing the kinetic energy of an impact through large deformations, but fall short on static strength requirements which favour rigidity under normal operational loads. These conflicting demands call for multidisciplinary design optimization (MDO), which was also used by Schuhmacher et al. (2002) in the preliminary sizing of the wingboxes for the Fairchild Dornier regional jet family. The authors highlighted that focusing on a limited number of load cases in aircraft design can neglect essential criteria, leading to costly revisions when the full set of requirements is assessed later. Therefore, the preliminary design process should ideally consider as many requirements as possible.

Yet, incorporating crashworthiness into optimization is no small feat. Crash analyses are computationally intensive due to their highly non-linear nature, involving large deformations, material non-linearities, multiple contacts and progressive fracture. Furthermore, high strain rates require knowledge about additional material parameters that might not be readily available. As a consequence, gradient information cannot be acquired easily, limiting the applicability of many traditional optimization strategies.

To navigate these challenges, surrogate models emerge as a powerful solution, capable of approximating highly non-linear functions. Although optimization can be carried out on such approximations directly, coupling them with Bayesian optimization offers a systematic approach to efficiently balance exploration and exploitation within the design space. Specifically, Bayesian optimization leverages the uncertainty quantification provided by surrogate models such as Kriging to guide sampling, enabling the solution of constrained optimization problems without gradient information, while simultaneously reducing the required number of computationally expensive crash analyses compared to using surrogate models alone.

Additionally, the preliminary design process often involves a high number of variables, leading to the infamous ‘curse of dimensionality’, which was first put forward by Bellman and Kalaba (1959). While their study was related to a mathematical framework for adaptive control processes using dynamic programming, it was acknowledged that increasing the dimensionality of a design space leads to an exponential increase in the number of necessary data points and to difficulties in model fitting. The same is reiterated by Viana et al (2021), who conducted a review on surrogate modelling in the context of multidisciplinary structural optimization, recommending to reduce

the dimensionality of the problem if there are more than 10 variables.

Surrogate models have been identified as a viable solution by other studies related to crashworthiness optimization as well. Bisagni et al (2002) conducted structural optimization on a helicopter subfloor structure using Artificial Neural Networks. A sensitivity study was carried out by performing multiple finite element analyses, and the most influential variables were kept for the optimization process. Building on this work, Lanzi et al (2004) decomposed the entire structure into smaller interconnected substructures, leading to a lower-fidelity model. The behaviour of each substructure was approximated using Artificial Neural Networks utilized within a genetic algorithm optimization. While crash efficiency improved less (12% vs. 25%), significant CPU time savings enhanced the optimization process, as only 93 finite element analyses were required, while employing genetic algorithms without surrogate models was expected to necessitate 73,500 analyses for 23 design variables.

With respect to dimensionality reduction, Craig et al (2005) conducted variable screening in the crashworthiness field by integrating analysis of variance (ANOVA) with linear approximations of crash-related parameters, effectively ranking variables based on their influence. This innovative strategy led to a computational time reduction exceeding 30% compared to tackling the same problem without variable screening.

Turning to bird strike crashworthiness requirements specifically, Pahange and Abolbashari (2016) conducted a multi-objective optimization aiming to reduce both the mass and leading edge intrusion of a riveted metallic wing structure under bird strike conditions. Beginning with a Taguchi design of experiments, they identified the most significant variables using ANOVA as well. The multiple objectives were combined into a single response function through Grey Relational Analysis. However, this study did not utilize surrogate models, which could have enabled the exploration of a wider range of designs and potentially a more efficient optimization.

Another study conducted by Ollar et al (2017) started from the assumption that a bird strike is a local event, allowing significant components to be identified through engineering judgment. Surrogate models for the structural response of the wing impacted at various locations were constructed, enforcing a maximum intrusion constraint in the leading edge region.

The present optimization framework for bird strike crashworthiness brings two advances. First, it embeds an explicit dynamic impact simulation and a subsequent static stress

check inside the same optimization loop. Second, it solves the resulting 19-variable, two-constraint sizing problem with a Bayesian surrogate optimizer that incorporates the variance-based variable screening procedure of Schonlau and Welch (2006). One constraint limits the penetration depth of the leading edge during impact, while the other caps the peak stress in the damaged front spar under representative ‘get-home’ loads. Because the static analysis is run on the deformed geometry and material state produced by the impact simulation, the link between dynamic damage and residual strength remains fully consistent. This combination represents a unique approach of taking bird strike crashworthiness requirements into consideration during an optimization procedure.

The paper is structured as follows: section 2 details the case study and how the objective and constraint data are gathered. The proposed methodology is then outlined in section 3, including initial data generation, model fitting, variable ranking, and Bayesian optimization techniques. The results are presented in section 4, where each step of the methodology is discussed, with a particular focus on the impact of the variable ranking procedure on optimization efficiency. Finally, section 5 offers brief conclusions and recommendations for future research.

2 Case study

2.1 Problem description

A five-bay metallic wing segment impacted by a 1.81 kg bird at 150 m/s is examined. The wing structure is adapted from the Fokker F27 Friendship and Fokker 50 aircraft, which are selected due to their relatively low cruise velocities of approximately 150 m/s, simplifying bird strike analyses thanks to the lower kinetic energy the leading edge must absorb and the lower anticipated strain rates. Table 1 summarizes all relevant geometric data.

As illustrated in Fig. 1, 6 stringers are placed on the lower and upper skin panels, respectively. Six nose ribs are present, positioned in the same planes as the corresponding wingbox ribs. For simplicity, it is assumed that the material of all components is Aluminium 2024-T3, whose properties are given in Sect. 2.2.1. The optimization problem comprises 19 design variables, each representing specific thicknesses within the wing structure, as depicted in Fig. 1, with their respective ranges and labels detailed in Table 2.

The objective of this case study is to minimize the weight of the wing while accounting for two constraints. The first constraint (c_1) relates to bird strike crashworthiness, imposing a maximum penetration depth of 240 mm, corresponding to

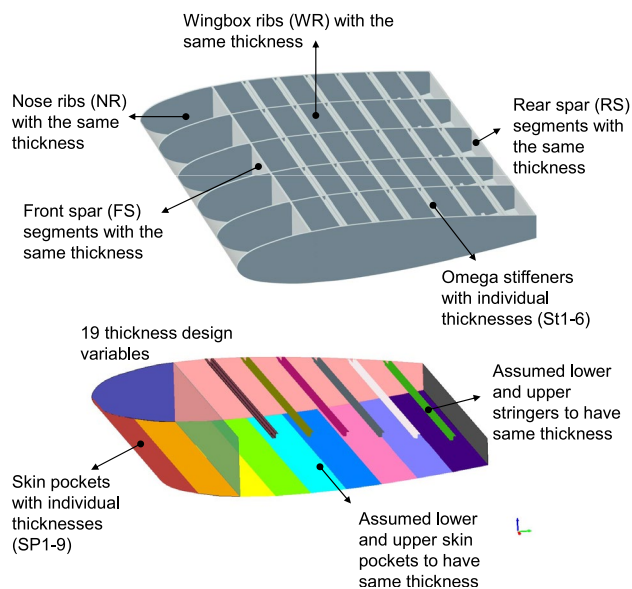


Fig. 1 The geometry of the wing model used for methodology demonstration along with explanatory notes on the design variables. Note: upper skin panels not included for clarity

Table 1 Summary of the geometrical features of the wing

Property or component	Geometrical features
Airfoil	NACA 0015
Span	2250 mm
Chord	2400 mm
Spars	Flange width: 14 mm Front spar at 20% chord Rear spar at 80% chord
Omega stringers	Lower flange width: 15 mm Upper flange width: 20 mm Height: 25 mm
Ribs	Flange width: 14 mm

Table 2 Summary of all design variables and their respective ranges. Note: ‘SP’, ‘WR’, ‘RS’, ‘FS’, ‘NR’ and ‘St’ stand for skin panel, wingbox rib, rear spar, front spar, nose rib and stringer, respectively

Design variable	Thickness range [mm]
Spars: 2 variables, FS & RS	1.5 – 3.5
Stringers: 6 variables, St1-6	0.8 – 2.0
Ribs: 2 variables, NR & WR	0.8 – 2.5
Skin panels: 9 variables, SP1-9	1.0 – 3.5

half the distance between the leading edge and the front spar. c_1 is therefore computed as the difference between the actual and desired maximum penetration depth. It is assumed that analysing this limited portion of the entire wing is sufficient for demonstrating the proposed methodology, and that the critical bird strike location is at the middle bay leading edge. The second constraint (c_2) relates to the remaining static strength of the damaged wing, emulating the ‘get-home’ flight conditions following a bird strike, as specified in CS 25.631, herein implemented as a maximum allowable stress of 80% of the yield stress on all the elements of the front spar, i.e., 295 MPa as will be given in Sect. 2.2.1 concerning material properties. Similarly, c_2 is computed as the difference between the actual and desired maximum stress. Denoting by $\mathcal{D}$ the 19-dimensional design space $[1.5, 3.5]^2 \times [0.8, 2.0]^6 \times [0.8, 2.5]^2 \times [1.0, 3.5]^9$ as explained back in Table 2, the optimization problem can be written as follows:

$$\begin{aligned} & \text{minimize} && \text{weight}(\mathbf{x}) \\ & \text{with respect to} && \mathbf{x} \in \mathcal{D} \\ & \text{subject to} && c_{1,2}(\mathbf{x}) \leq 0, \end{aligned} \quad (1)$$

2.2 Analysis workflow

Figure 2 illustrates the proposed analysis workflow, consisting of the following steps:

1. Modal analysis: Extract the first natural frequency of the model to apply Rayleigh damping for the dynamic relaxation of the bird impact simulation, detailed in Appendix A.
2. Explicit dynamic bird strike simulation to extract c_1 , the maximum penetration depth constraint. Dynamic relaxation is applied to reach a quasi-static state within a reasonable analysis time, as detailed in Appendix A. The analysis is terminated once stabilization of the kinetic and strain energies occurs. Automatic checks are made

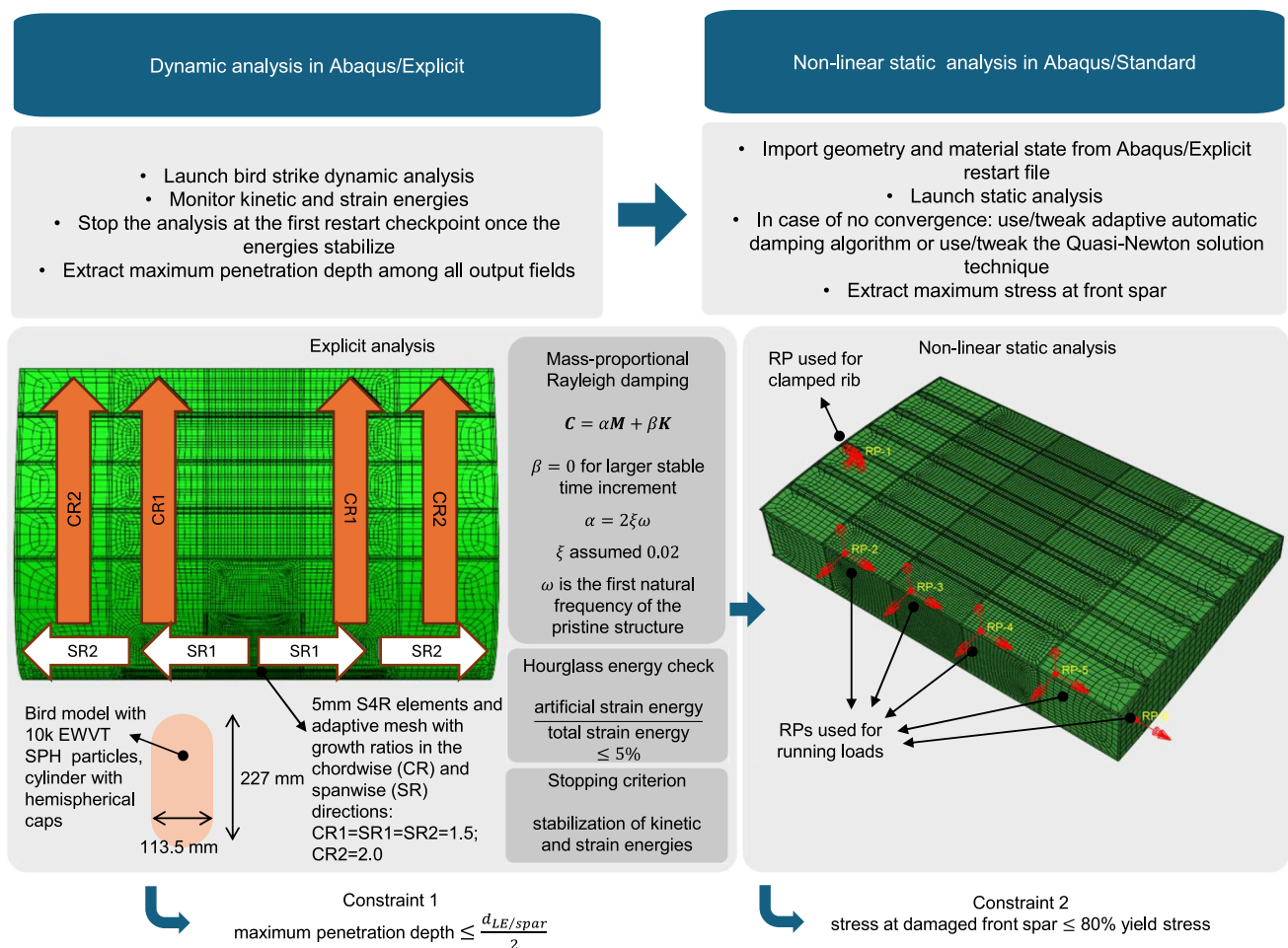


Fig. 2 Analysis workflow and details to obtain the constraints: 1) maximum penetration depth; 2) stresses at the front spar of the damaged wing

to ensure that the hourglass strain energy remains below 5%. Hourglass is present due to the use of the computationally efficient S4R elements.

3. Static analysis to evaluate the remaining static strength and extract the maximum stress experienced on the front spar.

In all cases, the boundary condition consists of the root rib being clamped.

2.2.1 Material properties

The Johnson-Cook model with strain rate dependency is employed for the plasticity model, as defined in Eq. 2. A , B , n , and C are material parameters, where σ_y represents the yield stress, ϵ_p is the equivalent plastic strain, and $\dot{\epsilon}_p$ denotes the equivalent plastic strain rate. All material parameters are provided in Table 3, based on data from Lesuer (2009, p. 8).

$$\sigma_y = [A + B(\epsilon_p)^n] [1 + C \ln \dot{\epsilon}_p] \quad (2)$$

Due to the deletion of elements upon failure, a progressive damage criterion is necessary. Although Lesuer (2009) provides parameters for the Johnson-Cook damage initiation criterion, it is applicable only to Abaqus/Explicit (Dassault Systèmes 2023a, Johnson-Cook criterion section). Since the static analysis is carried out in Abaqus/Standard based on the results from the explicit simulation, this criterion cannot be employed. Consequently, a ductile damage criterion in tabular format will be used, with data provided by Vershinin (2015). As for damage evolution, the plastic displacement formulation has been chosen, as its value of $\bar{u}_f^{pl} = 0.001$ mm was readily found in Klosak et al (2021).

Table 3 Aluminium 2024-T3 material constants used in the analysis, as given by Lesuer (2009)

Property	Value
Density	2,770 kg/m ³
Young's modulus	74.66 GPa
Poisson ratio	0.3
A	369 MPa
B	684 MPa
n	0.73
C	0.0083

Table 4 Data for the bird equation of state, as given by Marulo and Guida (2014)

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

2.2.2 Explicit dynamic bird strike analysis

To model the bird impact, a Smoothed Particle Hydrodynamics (SPH) formulation was employed for the bird, generated internally within GKN Fokker following the Extended Weighted Voronoi Tessellation (EWVT) algorithm proposed by Siemann and Ritt (2019). The resulting bird model comprises 10,000 particles and is shaped as a cylinder with hemispherical caps, maintaining a length-to-diameter ratio of 2 and material density of 950 kg/m³. The equation of state, relating pressure to volumetric strain, was defined using tabular data provided by Marulo and Guida (2014), as given in Table 4.

Abaqus/Explicit's General Contact Algorithm was used, which handles efficiently node-into-face and edge-to-edge contact across an all-inclusive contact surface for the entire model.

Assuming a perfectly rigid target, a rough calculation based on the bird velocity and length yields an impact duration of approximately $L_{bird}/V_{initial} \approx 1.5$ ms. However, the actual simulation time needs to be extended beyond this estimate due to the flexibility of the target and the necessity for the stabilization of elastic strain and kinetic energies within the wing structure, which is crucial to ensure that the subsequent static analysis can reach equilibrium. The time required for the dynamic relaxation may vary for each specific case, influenced by factors such as different damping ratios and structural failure modes. To address this variability, an Abaqus Python script was implemented to monitor the elastic strain and kinetic energies of the structure in real-time. The simulation automatically terminates once the energies have remained relatively constant. Specifically, the relative differences between consecutive history outputs are monitored, and when they reach less than 3% over 10 outputs, amounting to 0.02 s, the analysis is terminated at the next restart checkpoint.

Assuming that a bird strike is a localized event, the mesh density is increased in both chordwise and spanwise directions starting from the impacted bay, as illustrated in Fig. 2. The mesh size for the impacted bay was fixed after conducting a convergence study on an isolated leading edge with a clamped front spar and subject to the same bird velocity, keeping all design variables at their minimum values to ensure convergence throughout the entire design space. A small element size of 5 mm was needed to prevent large hourglass effects. Further studies investigated different mesh growth ratios in the spanwise and chordwise directions, settling to the ones depicted in Fig. 2.

2.2.3 Static analysis for remaining strength

For the static analysis, only the main wingbox is considered, as illustrated in Fig. 2, ignoring the leading edge skin and nose ribs. An adaptive automatic damping algorithm is used, as presented in Dassault Systèmes (2022). A maximum allowable ratio of 5% of the stabilization energy to the total strain energy is imposed, although it was found that the automatic stabilization algorithm overestimates the maximum ratio, and that, usually, a ratio of approximately 2% is obtained.

Lift and pitching moment coefficients have been used from van der Vaart and Muhammad (1983, pp. 31–34) to derive the static loads. To simplify the analysis, the entire wing is modelled as a beam with a constant line load acting perpendicular to it, representing the aerodynamic lift distributed upward force and an associated bending moment that varies along the spanwise direction, approximated as concentrated forces and moments acting at specific reference points located on each wingbox rib, positioned at a quarter of the chord length. These reference points then distribute the loads to the nodes on all four rib flanges, as depicted in Fig. 2. The resulting concentrated forces and moments are summarized in Table 5.

3 Proposed methodology

The proposed methodology consists of several key steps as illustrated in Fig. 3. Firstly, an initial dataset is acquired and divided into training and validation sets. Two Kriging surrogate models are then constructed using exponential and squared-exponential kernels, selected for their suitability with functions of varying smoothness. The model demonstrating the best performance based on the root mean squared error (RMSE) is chosen. Subsequently, a variable ranking procedure is conducted, and the surrogate model is retrained on the reduced design space identified through

Table 5 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

Wingbox rib number	Torsion moment [Nm]	Transverse 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

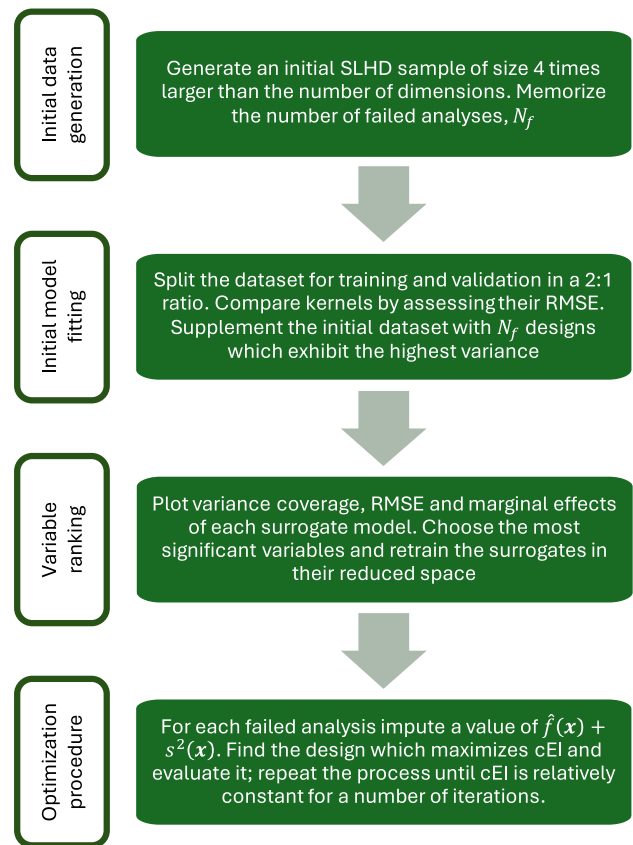


Fig. 3 Flowchart of the proposed methodology

this process. These steps are repeated individually for each constraint. Finally, the refined surrogate models are incorporated into a Bayesian optimization framework. Each of these steps is detailed in the subsequent subsections.

3.1 Initial data generation

Constructing the initial surrogate model for variable ranking requires a careful balance between ensuring enough training points to accurately identify significant variables, while also minimizing computational effort due to the intensive nature of data acquisition. While some recommend using 10 times the number of variables for initial sampling (Schonlau et al 1998), studies by the same authors have shown that 3–4 times the number of variables can suffice even for non-linear problems (Jones et al 1998; Welch et al 1992). Since optimization prioritizes the ability of the surrogate model to find minima over global accuracy (Viana et al 2009), an initial dataset of 4 times the number of variables is proposed. Approximately one-third of these samples are reserved for validation, following the recommendation by Viana et al (2021).

The samples are generated using the Sliced Latin Hypercube Design (SLHD) algorithm proposed by Ba et al (2015),

which has a space-filling nature that prevents two data points from sharing the same coordinate in any dimension.

The constraints c_1 and c_2 are obtained with the analysis workflow in section 2. However, in certain instances, extracting these constraints was not feasible. For example, during the bird strike analysis, if there is a surge in total energy or if the ratio of artificial hourglass to total strain energy is high, the penetration depth cannot be determined. Similarly, the static analysis might fail to converge, or the ratio between stabilization energy and total strain energy may also be too high, preventing data extraction once again. Denoting by N_f the number of failed analyses where constraint data could not be extracted during the initial data generation, it is proposed to fit a Kriging surrogate model as detailed in the next subsection and to supplement it with N_f additional data points that maximize the variance of the model, enhancing its global accuracy (Wang et al 2020, 2022). The maximum acceptable ratio of artificial hourglass to total strain energy in this study is 5%.

3.2 Initial model fitting

A comprehensive explanation on the derivation of the Kriging surrogate models is presented in Forrester et al (2008, Chapter 2) or Jones (2001). The final form of the ordinary Kriging model is given in Eq. 3, where $\hat{f}$ is the predictor at an unknown point $\mathbf{x}$:

$$\hat{f}(\mathbf{x}) = \hat{\mu} + \mathbf{r}^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu}) \quad (3)$$

The first term stems from the prior assumption of the ordinary Kriging that the observational random variables Y which models the data are normally distributed with mean μ and variance σ^2 . Their maximum likelihood estimates, $\hat{\mu}$ and $\hat{\sigma}^2$, are obtained by maximizing the likelihood of the observational random variable yielding the responses in the training set, $\mathbf{y}$, based on the normal distribution aforementioned. $\mathbf{r}$ denotes the correlation vector between the new point and the training points, and $\mathbf{R}$ is the correlation matrix among the training points. The correlation $\text{Corr}[Y(\mathbf{x}^{(1)}), Y(\mathbf{x}^{(2)})]$ between the observational random variables of two points $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ is assumed to have the form given in Eq. 4, where $\theta_l > 0, p_l \in [1, 2]$ are hyperparameters:

$$\exp\left(-\sum_{l=1}^k \theta_l |x_l^{(1)} - x_l^{(2)}|^{p_l}\right) \quad (4)$$

As the design space has k dimensions, the training process involves estimating the $2k + 2$ hyperparameters. However, in the present study the exponential or squared-exponential functions will be used, fixing $p_l = 1$ and $p_l = 2, \forall l \in \overline{1, k}$, respectively, because of their ability to model functions of different smoothness and their ease of implementation

in the variable screening process that is detailed next. The most suitable kernel is chosen by computing the root mean squared error using the validation dataset. To tune the hyperparameters, the chosen package, Trieste, maximizes the logarithmic marginal likelihood, which is found by integrating the likelihood over the prior, using the gradient descent Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm.

Last but not least, the variance $s^2(\mathbf{x})$ at an unknown point $\mathbf{x}$ is given by Eq. 5, which is a prediction error estimate:

$$\hat{\sigma}^2 \left[1 - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} \frac{(1 - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r})^2}{\mathbf{1}^T \mathbf{R}^{-1} \mathbf{1}} \right] \quad (5)$$

3.3 Variable ranking

Once the initial model is trained, the significance of the variables can be ranked using the analysis of variance (ANOVA) on the Kriging surrogate. In this manner, not only the main effects of the variables, but also the higher-order interactions among them can be captured. This approach was previously introduced in conjunction with Universal Kriging surrogates by Schonlau and Welch (2006). However, for simplicity, the present study employs ordinary Kriging models instead. For a more detailed derivation of all the terms presented in this section, the reader is referred to Ciobotia (2024).

Intuitively, this method quantifies how much of the total variance exhibited by the surrogate model across the entire design space is attributable to a specific set of variables. Let χ denote the entire design space, which is assumed to be the Cartesian product of the one-dimensional domains of the individual variables:

$$\chi = \bigotimes_{j=1}^k \chi_j \quad (6)$$

If $\mathbf{e}$ is a subset of $\overline{1, k}$, the notation $\mathbf{x}_{\mathbf{e}}$ will be used to represent the variables whose effects are under investigation, while $\mathbf{x}_{-\mathbf{e}}$ denotes the remaining variables in the set $\{1, \dots, k\} - \mathbf{e}$. Accordingly, the input vector can be partitioned as $\mathbf{x} = (\mathbf{x}_{\mathbf{e}}, \mathbf{x}_{-\mathbf{e}})$. The notation $\mathbf{x}_{\mathbf{e}}$ is essentially the canonical projection of $\mathbf{x} \in \chi$ onto the subspace $\bigotimes_{j \in \mathbf{e}} \chi_j$. Note that for the study of main effects, $\mathbf{e}$ will contain only a single element.

The so-called marginal effect $\hat{f}_{\mathbf{e}}$ of $\mathbf{x}_{\mathbf{e}}$ is computed by integrating out all other variables, effectively yielding the mean value of the predictor given only these selected dimensions. From a statistical perspective, this effect corresponds to the expected value of the predictor $\hat{f}$ conditioned on the variables $x_i, i \in \mathbf{e}$, i.e., $\mathbb{E}[\hat{f} | X_i, i \in \mathbf{e}]$, which will be denoted by $\hat{f}_{\mathbf{e}}(\mathbf{x}_{\mathbf{e}})$ and has the following form:

$$\hat{f}_e(\mathbf{x}_e) = \int_{\otimes_{j \in e} \mathcal{X}_j} \hat{f}(\mathbf{x}_e, \mathbf{x}_{-e}) \prod_{j \notin e} w_j(x_j) dx_j \quad (7)$$

In Eq. 7, the weight $w_j(x_j)$ is the probability of the value x_j occurring in the j^{th} dimension of the design space, $\mathcal{X}_j$. It is assumed that the thicknesses acting as design variables follow the simple uniform distribution given in Eq. 8:

$$w_j(x_j) = \frac{1}{x_j^u - x_j^l}, \forall j \in \overline{1, k}, \quad (8)$$

where $x_j^u = \max(\mathcal{X}_j)$, $x_j^l = \min(\mathcal{X}_j)$. The product of the weights in Eq. 8 will be comprised in $w(\mathbf{x}) = \prod_{j=1}^k w_j(x_j)$. The variance of the predictor in the whole design space is:

$$\begin{aligned} \text{Var}[\hat{f}] &= \int_{\mathcal{X}} [\hat{f}(\mathbf{x}) - \mathbb{E}[\hat{f}]]^2 w(\mathbf{x}) d\mathbf{x}, \text{ where} \\ \mathbb{E}[\hat{f}] &= \int_{\mathcal{X}} \hat{f}(\mathbf{x}) w(\mathbf{x}) d\mathbf{x} \end{aligned} \quad (9)$$

To estimate the significance of a set of variables, the ANOVA decomposition utilizes the so-called corrected marginal effects. In particular, the first-order and second-order corrected marginal effects are defined as follows:

$$\begin{aligned} \mu_j(x_j) &= \hat{f}_j(x_j) - \mathbb{E}[\hat{f}], \forall j \in \overline{1, k} \\ \mu_{jl}(x_j, x_l) &= \hat{f}_{jl}(x_j, x_l) - \mu_j(x_j) \\ &\quad - \mu_l(x_l) - \mathbb{E}[\hat{f}], \forall j, l \in \overline{1, k} \end{aligned} \quad (10)$$

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:

$$\begin{aligned} \text{Var}[\hat{f}] &= \sum_{e \subseteq \{1, 2, \dots, k\}} \text{Var}[\hat{f}|X_e] \\ &= \sum_{j=1}^k \int_{\mathcal{X}_j} \mu_j^2(x_j) w_j(x_j) dx_j \\ &\quad + \sum_{j=1}^{k-1} \sum_{l=j+1}^k \int_{\mathcal{X}_j} \int_{\mathcal{X}_l} \mu_{jl}^2 w_j(x_j) w_l(x_l) dx_j dx_l \\ &\quad + \dots + \int_{\mathcal{X}} \mu_{1\dots k}^2(x_1, \dots, x_k) \prod_{j=1}^k w_j(x_j) dx_j \end{aligned} \quad (11)$$

The contribution of the main effects or joint interactions to the total variance represents the importance of the variable(s) within the model. This importance is quantified by expressing these terms as percentages of the total variance. Accordingly, the importance of variable x_j will be $\text{Var}[\hat{f}|X_j]/\text{Var}[\hat{f}] \cdot 100\%$, with the importance of the interaction between two variables x_j and x_l defined as $\text{Var}[\hat{f}|X_{\{j,l\}}]/\text{Var}[\hat{f}] \cdot 100\%$ in a similar manner. Note that

the symbol X_j refers to a random variable, while the lower-case notation x_j is used to indicate a specific realization of the random variable X_j .

In addition to these significance measures, evaluating the total variance coverage of a subset can help determine whether the reduced design space reliably captures the information from the entire domain. To further ensure accurate variable screening, plotting the marginal effects of various variable pairs can be used for visualization purposes. Moreover, assessing error metrics such as the RMSE or cross-validation errors of the surrogates within the reduced design space is advisable. As the sample density increases in this reduced-space, the accuracy of the surrogate model is expected to improve accordingly.

3.4 Optimization procedure

The constrained expected improvement (cEI) approach proposed by Schonlau et al (1998) is used as an acquisition function during the Bayesian optimization procedure, which has the following form:

$$\mathbb{E}[I(\mathbf{x})]_c = \mathbb{E}[I(\mathbf{x})] \prod_{i=1}^2 \mathbb{P}[c_i(\mathbf{x}) \leq 0] \quad (12)$$

In Eq. 12, $\mathbb{E}[I(\mathbf{x})]$ is the expected improvement of the objective function. However, in this specific case, the weight function can be expressed analytically due to the choice of design variables, i.e., the thicknesses of various structural components. As a result, the wing does not undergo any architectural changes, and all components maintain a constant area. Therefore, the expected improvement can be replaced by a known improvement, as shown in Eq. 13.

$$\mathbb{E}[I(\mathbf{x})] \rightarrow \text{weight}_{\min} - \text{weight}(\mathbf{x}) \quad (13)$$

The term $\mathbb{P}[c_i(\mathbf{x}) \leq 0]$ represents the probability of feasibility for one of the constraint functions, $c_i(\mathbf{x})$. For ordinary Kriging surrogate models utilizing squared-exponential correlation functions, this probability can be derived as shown in Eq. 14:

$$\int_{-\infty}^0 \frac{1}{\sqrt{2\pi} s_i(\mathbf{x})} \exp \left[-\frac{1}{2} \left(\frac{t - \hat{f}_i(\mathbf{x})}{s_i(\mathbf{x})} \right)^2 \right] dt, \quad (14)$$

$$\forall i \in \{1, 2\}$$

The optimization procedure aims to maximize the acquisition function, balancing exploration and exploitation within the design space. When the expected improvement is significant but the probabilities of feasibility are low, the optimizer explores new regions, seeking potential optimal solutions in uncertain areas. Conversely, when the probabilities of feasibility are high but the expected improvement is minimal,

the optimizer exploits known feasible regions to refine the solution.

However, difficulties may arise if the optimization algorithm selects a design that results in a non-converged analysis or a failed design in this context. This issue is addressed by Forrester et al (2008, p. 133), who propose a method to impute data for such failed designs to steer the optimizer away from these problematic regions. The approach involves initially training the surrogate model using all data points corresponding to successful (converged) analyses. For each failed design point $\mathbf{x}$, an imputed observation is generated using $\hat{f}(\mathbf{x}) + s^2(\mathbf{x})$. The surrogate model is then retrained on the entire dataset, including these imputed values. Subsequently, the constrained expected improvement function is maximized to determine the next sampling point. This process is iteratively repeated until a suitable convergence criterion is achieved for the Bayesian optimization procedure.

Using $\hat{f}(\mathbf{x}) + s^2(\mathbf{x})$ for imputed data ensures that the smoothness of the surrogate model remains unaffected. If a failed data point $\mathbf{x}^f$ is very close to a training point $\mathbf{x}^{(i)}$, i.e., $\|\mathbf{x}^f - \mathbf{x}^{(i)}\| \rightarrow 0$, the variance $s^2(\mathbf{x}^f)$ approaches zero. Since the predictor interpolates the data, it follows that $\hat{f}(\mathbf{x}^f) + s^2(\mathbf{x}^f) \rightarrow y_i$, where y_i is the observation at the training point. Conversely, if the variance is high at $\mathbf{x}^f$, the imputed value becomes significantly larger, discouraging the optimizer from sampling in that region due to the increased likelihood of violating the constraint function.

An important aspect to consider is the convergence criterion of the Bayesian optimization procedure. In practical applications, time and computational resources are often the most decisive factors. However, setting these aside, a straightforward convergence criterion can be formulated: the optimization process can be terminated when the acquisition function remains relatively constant over a specified number of iterations.

In order to accelerate the optimization process, two approaches are compared:

1. **Sequential approach:** Each new data point is obtained by maximizing the acquisition function only after the Abaqus workflow for the previous data point has completed. This method ensures that each point is the true maximizer of the acquisition function at that iteration.
2. **Batch approach:** More data points are obtained simultaneously by maximizing the acquisition function while ‘fantasizing’ unknown results via the ‘Kriging believer’ method (Ginsbourger et al 2008; Sun et al 2020), where the value is fantasized to be the predictor itself, similar to imputing data for unresolved analyses. The number of data points in the batch depends on available computational resources, such as licenses or processing power. Although this approach is expected to minimize total

wall time, the data points may not be the exact maximizers of the acquisition function due to the fantasized data.

3.5 Analytical validation problem

The proposed methodology will first be validated on the non-linear Rosenbrock function proposed by Rosenbrock (1960), constrained to a disk:

$$\begin{aligned} &\text{minimize} && (1 - x_1)^2 + 100(x_2 - x_1^2)^2 \\ &\text{with respect to } \mathbf{x} \in [-1.5, 1.5]^2 \\ &\text{subject to} && x_1^2 + x_2^2 - 2 \leq 0 \end{aligned} \quad (15)$$

The known global solution of Eq. 15 is $\mathbf{x}^* = (1.0, 1.0)$ and the objective function value is 0. The challenge of this function is therefore given by the optimum lying on the feasibility border, inside a long valley, which is very flat along its centreline but steep across it. Moreover, in order to test the variable screening, the variables will be normalized and the search space will be extended to a 20-dimensional unit hypercube.

A Latin Hypercube Sampling (LHS) of 80 initial points was created. The exponential and squared-exponential kernels were compared in terms of RMSE using a 27-point validation dataset. This validation set was selected by generating 10^6 random samples and choosing the one with the smallest discrepancy, which is a measure of how uniformly samples fill the design space, in order to ensure that the accuracy of the model is assessed on the whole design space. The squared-exponential kernel demonstrated superior performance for the objective function and constraint.

The variable screening process identified correctly only the first two variables as important for the objective and constraint functions, with more than 99.95% variance coverage from their main and joint effects for the objective function, as presented in Table 6. No interaction was identified for the constraint function, as expected. The other effects for both functions did not exceed 0.0014%.

Both the sequential and the batch strategies were left to run until the acquisition function was smaller than 10^{-6} for 10 consecutive iterations. The surrogate model seems to be unable to approximate well the steep valley near the true

Table 6 Variance coverage for the Rosenbrock function

Function	Variables	Variance coverage
Objective	x_1	34.57%
	x_2	40.49%
	x_1, x_2	24.90%
Constraint	x_1	50.10%
	x_2	49.86%
	x_1, x_2	0.002%

Table 7 Sequential and batch approaches comparison for the Rosenbrock function

Approach	Sequential	Batch
Solution	(0.832099, 0.831337)	(0.830477, 0.828131)
Number of iterations	143	31 batches of 4 points each
Objective value	0.000211	0.000286
Constraint value	−0.019340	−0.048041

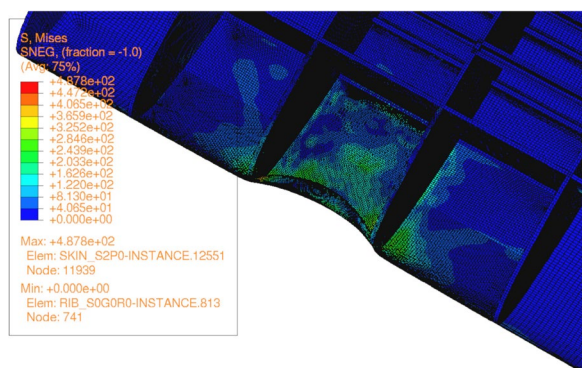
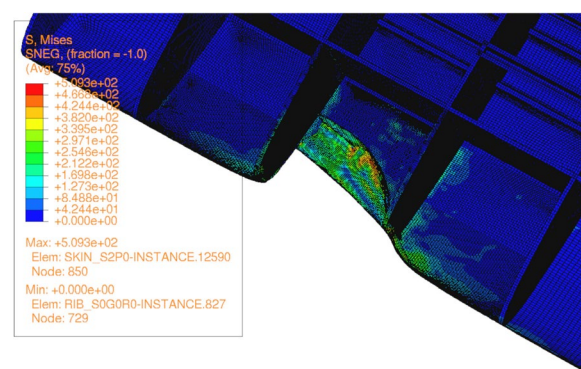
optimum. However, the solutions were feasible and rendered a satisfactory objective value, the batch approach converging in a smaller number of iterations, as presented in Table 7.

4 Results and discussion

4.1 Initial data generation and model fitting

To generate the initial dataset, an LHS of 76 samples was created using the package developed by Ba (2015). Due to the extensive design space, a wide range of failure modes was observed in the analyses. These modes varied from designs exhibiting minimal penetration and slight deformation of the nose ribs as depicted in Fig. 4, to more severe cases where the leading edge skin completely ruptured and the nose ribs were significantly crushed, as shown in Fig. 5.

Out of the 76 initial data points generated, two bird strike analyses failed due to high artificial hourglass strain energies, and an additional nine static analyses did not converge. To evaluate the performance of the surrogate models, the exponential and squared-exponential kernels were compared in terms of RMSE using a 25-point validation dataset. This validation set was generated in the same manner as for the Rosenbrock benchmark function.

**Fig. 4** Small penetration and small nose rib distortion**Fig. 5** Large penetration, leading edge skin rupture and severe nose rib crushing

Ultimately, the squared-exponential kernel demonstrated superior performance for both constraints. However, the depth surrogate model exhibited an RMSE of approximately 125 mm, which exceeds half of the maximum acceptable penetration depth, indicating insufficient global accuracy. In contrast, the front spar stress surrogate model achieved an RMSE of less than 52 MPa, corresponding to less than 15% of the yield strength of the spar material, and was therefore considered acceptable. Due to the high RMSE of the depth surrogate, the remaining 11 data points were added by selecting locations that maximized the variance of the model, thereby improving the global accuracy of the surrogate.

4.2 Variable ranking

The models are retrained on the entire dataset, also incorporating the validation points. This retraining ensures that the surrogate models capture more reliably the underlying functions, improving the effectiveness of the variable ranking process.

4.2.1 Spar stress constraint

Figure 6 illustrates the main and interaction effects as percentages of the total variance of the model. Main effects are positioned along the diagonal, while interaction effects are located in the lower-right section of the plot. For simplicity, effects below 0.1% are masked. Interestingly, despite the exclusion of the leading edge skin and nose ribs from the Abaqus static analysis model, the variable screening method identified them as the most significant variables affecting the stresses in the front spar, including their interaction.

Furthermore, the wingbox rib was determined to be important, even though it does not interact with other variables. Components of the wingbox, such as stringers and skin panels, were likely identified as significant due to their

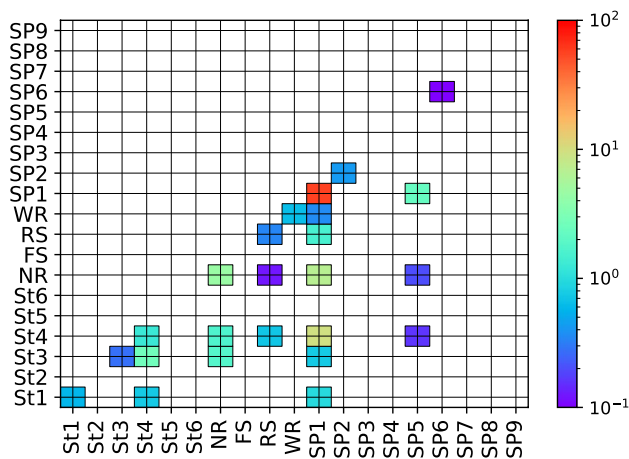


Fig. 6 Main and interaction effects, in percentages, for the maximum Mises stress constraint on the front spar. Note: all effects with a value of less than 0.1% have been masked

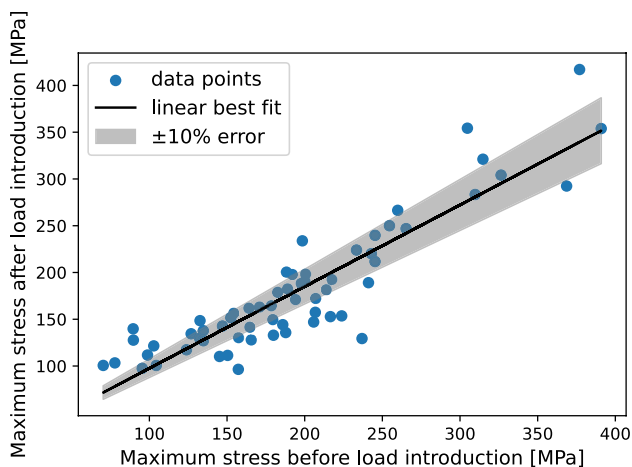


Fig. 7 Scatter plot between the maximum stress before and after load introduction, along with the best linear predictor and a $\pm 10\%$ error area

contribution to the torsional stiffness of the wingbox, given the substantial magnitude of the torsion moment. This rationale will be detailed shortly by plotting the marginal effect of the wingbox rib.

It is noteworthy that the front spar does not contribute significantly to the response despite the maximum Mises stress being derived from it. This may be attributed to the transverse loads applied not being sufficient to substantially increase the stress on the spar. Figure 7 demonstrates a strong correlation between maximum stresses before and after load introduction, with a Pearson correlation coefficient of 0.90 and an R-squared value of 0.82, indicating a linear relationship between the two. However, since these correlations are not perfect, the stress after load introduction is not entirely explained by the initial stress. This

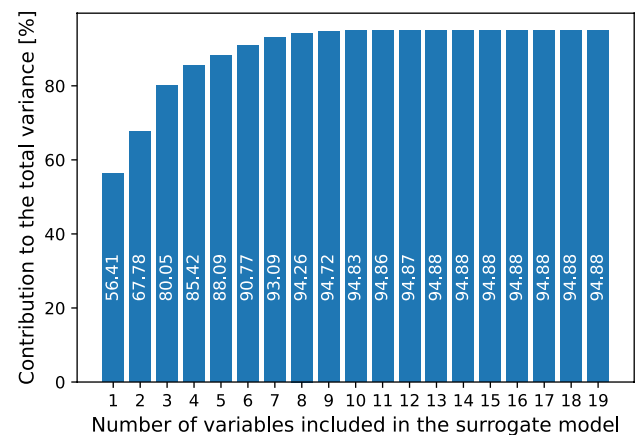


Fig. 8 Variance coverage versus number of included variables for the maximum stress constraint. The order of variable addition: SP1, NR, St4, St3, SP5, RS, St1, WR, SP2, SP6, St6, St2, SP4, SP7, St5, SP9, SP3, SP8, FS

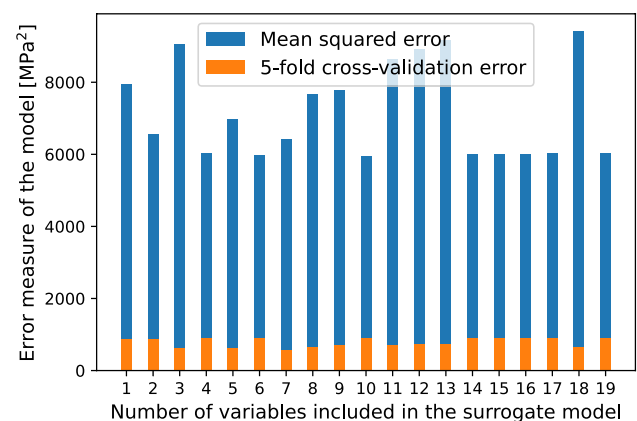


Fig. 9 The mean squared and 5-fold cross-validation errors for reduced-space surrogates for the stress constraint

partial explanation underscores the importance of other design variables influencing the overall response.

Figure 8 presents a bar plot illustrating the evolution of variance coverage as the number of included variables increases. Note that the variables are added in the specific order which maximizes the variance coverage with that particular number of variables. It is evident that higher-order interactions are present, as the variance coverage remains below 95% despite the inclusion of all variables. Beyond eight variables, the benefit of adding additional dimensions in terms of variance coverage decreases to less than 1%, indicating that further inclusion of variables yields minimal gains in the variance coverage of the model.

Another aspect aiding in decision-making involves refitting surrogate models within the reduced design space using the initial training dataset and evaluating their mean

squared and cross-validation errors, illustrated in Fig. 9. The latter significantly underestimates the MSE and does not align with its trend. Curiously, both errors exhibit a sudden decrease at the five-variable threshold. This reduction is disregarded due to only 88% of the total variance being covered, which is insufficient. Beyond eight variables, both error metrics stabilize, except for a slight anomaly in the MSE at fourteen to seventeen variables. It is important to recognize that MSE estimates global accuracy and may be influenced by artefacts or fortunate hyperparameter optimization within the specific reduced design space. Nonetheless, the error measures of the eight-dimensional surrogate model remain comparable to those of the nineteen-variable model, indicating that the reduced metamodel maintains adequate accuracy despite the presence of higher-order interactions. This suggests that the variable screening procedure effectively retains the most significant parameters without compromising the reliability of the surrogate model.

In total, eight significant variables were identified: the leading edge skin, nose ribs, wingbox ribs, three stringer stations, one material zone of the skin panel, and the rear spar. The exclusion of certain stringer stations and additional skin panel material zones may appear odd. This limitation suggests that higher-order interactions with other variables were not fully captured, exposing a potential weakness in the current methodology.

A final safeguard against the erroneous identification of significant variables involves evaluating the marginal effect plots, illustrated with two examples. The significance of the wingbox rib, attributed to its contribution to torsional stiffness, is confirmed by the linear, decreasing relationship shown in Fig. 10, where a thicker rib will lead to lower stresses on the front spar. The marginal effect of the leading edge skin, the most significant variable, is depicted in Fig. 11, exhibiting a parabolic shape: a very thin skin results in substantial penetration and rupture of the front spar, while an excessively thick skin may become overly rigid, reducing

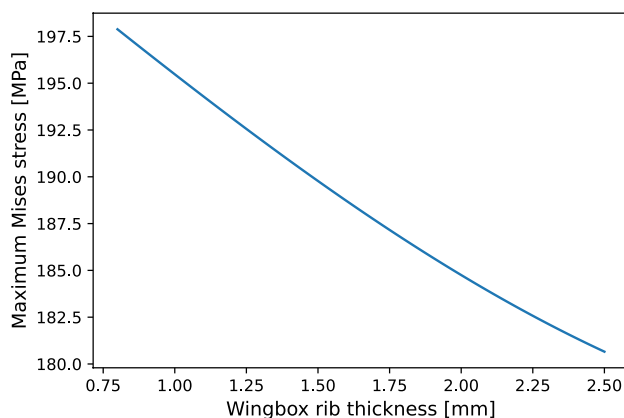


Fig. 10 Marginal effect of the wingbox rib

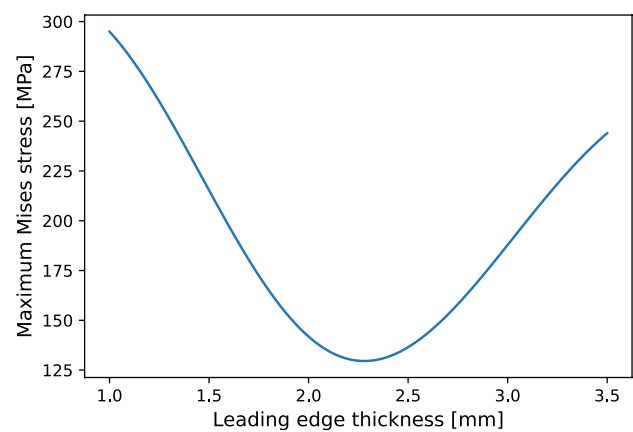


Fig. 11 Marginal effect of the leading edge skin

its energy absorption capacity and transferring more kinetic energy to the rest of the structure, including the front spar.

4.2.2 Penetration depth constraint

Despite the high RMSE associated with the penetration depth constraint, only a few significant variables are identified, as illustrated in Fig. 12. The leading edge skin stands out as the most influential factor, accounting for 88.60% of the total variance. It interacts with the nose rib, contributing an estimated 1.35%, even though the nose rib itself has a modest main effect of 0.78%. It is acknowledged that the variance coverage for the nose rib seems low, and that the values may be erroneous due to the high RMSE of the surrogate.

The wingbox rib also plays a notable role, with a main effect of 1.31%, although it does not interact with other

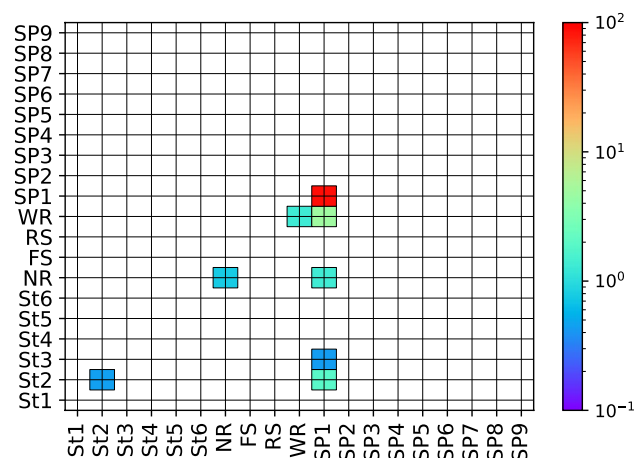


Fig. 12 Main and interaction effects for the penetration depth constraint. Note: all effects with a value of less than 0.1% have been masked

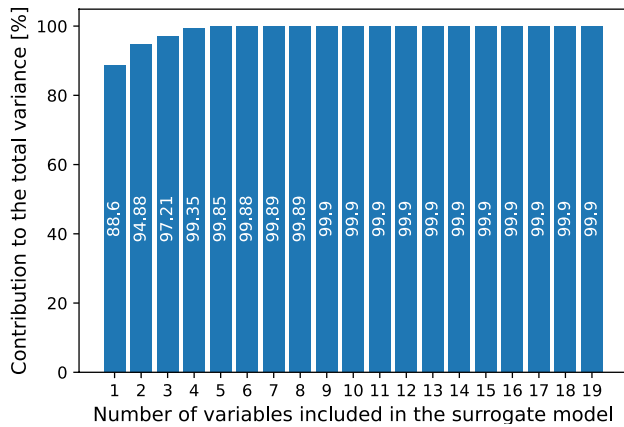


Fig. 13 Variance coverage versus number of included variables for the penetration depth constraint. The order of variable addition: SP1, WR, St2, NR, St3, SP3, St6, St4, SP6, SP4, SP9, St1, St5, SP5, SP8, FS, SP7, SP2, RS

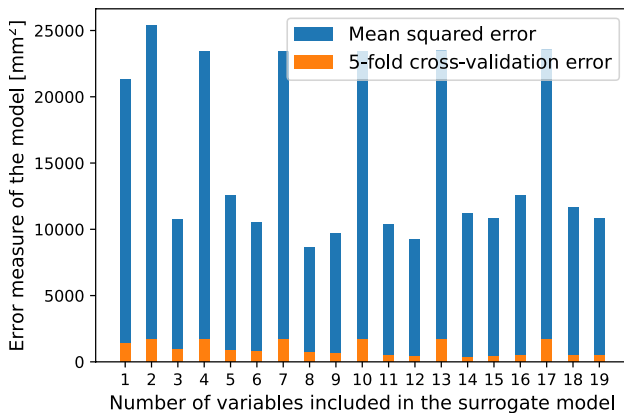


Fig. 14 The mean squared and 5-fold cross-validation errors for reduced-space surrogates the penetration depth constraint

variables. Interestingly, stringer 2 has a main effect of 0.44% and is estimated to interact with the leading edge skin by 1.89%. However, this interaction may be inaccurately identified due to the poor global accuracy of the surrogate model.

Figure 13 illustrates the variance coverage, while Fig. 14 displays the mean squared and cross-validation errors. It is evident that increasing the number of variables beyond four yields minimal improvements in variance coverage. Notably, the three-dimensional design space achieves substantial variance coverage and an impressive 80% reduction in MSE compared to the full-dimensional model. However, the cross-validation errors significantly overestimate the global accuracy of the model and do not align with the trend observed in the MSE. This discrepancy suggests that the cross-validation metric may not reliably reflect the true accuracy of the surrogate model.

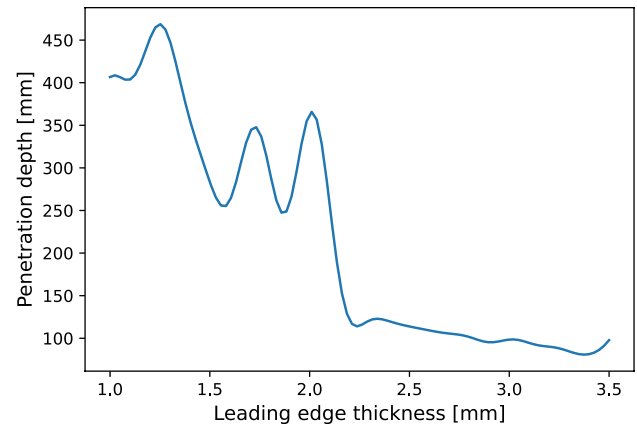


Fig. 15 Marginal effect of the leading edge skin for the penetration depth constraint

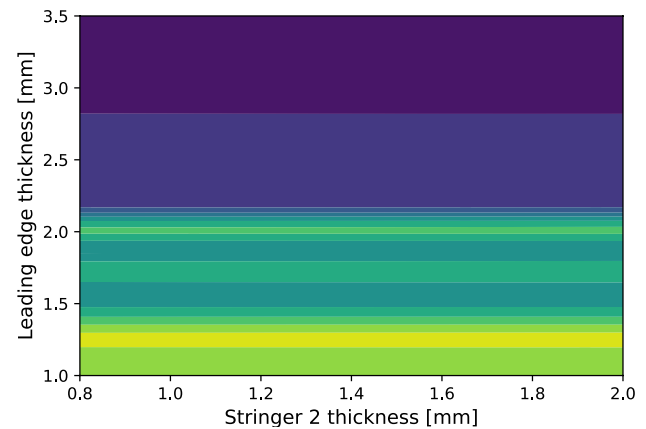


Fig. 16 Marginal effect of the second station stringer and the leading edge skin for the penetration depth constraint

The importance of examining the marginal effects of individual variables becomes clearer in the case of the maximum penetration depth constraint. While it is apparent from the marginal effect illustrated in Fig. 15 that the leading edge has a significant contribution, a highly multimodal landscape is predicted. Figure 16 showcases the marginal effect of the leading edge and the second stringer station, whose interaction was predicted rather high. However, the presence of bands of constant values suggests that the metamodel may lack sufficient accuracy.

Based on engineering judgment, the leading edge, nose ribs, and wingbox ribs are selected as significant variables, while the second stringer station is excluded from further consideration. Following this selection, a re-evaluation of the error metrics within the design space shows an improvement in the MSE, decreasing from 7532.61 to 6269.87 mm². On the other hand, the cross-validation error increased from 192.54 to 643.8 mm², reinforcing the necessity of

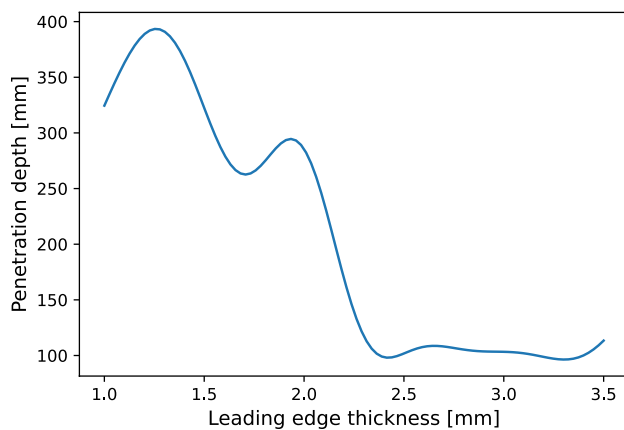


Fig. 17 Marginal effect of the leading edge skin for the penetration depth constraint in the reduced design space

maintaining a separate validation set to ensure the reliability of the surrogate models.

The fact that the variables have been well chosen is also supported by the marginal effect of the leading edge skin in the reduced design space: the high modality is no longer present, as depicted in Fig. 17.

4.3 Optimization procedure

After the variable screening procedure, all training points are projected onto the reduced eight-dimensional design space, with insignificant variables maintained at their minimum values. It is assumed that this projection would not alter the constraint observations. As a consequence, it is assumed that the weight of the optimal design among the initial dataset decreases from 78.83 kg before projection to 52.58 kg after, which amounts to a 33.31% reduction. Figure 18 depicts the damage extent of the optimal design among the initial dataset, showing moderate penetration and rib crushing with a depth of 161.64 mm. The maximum Mises stress on the front

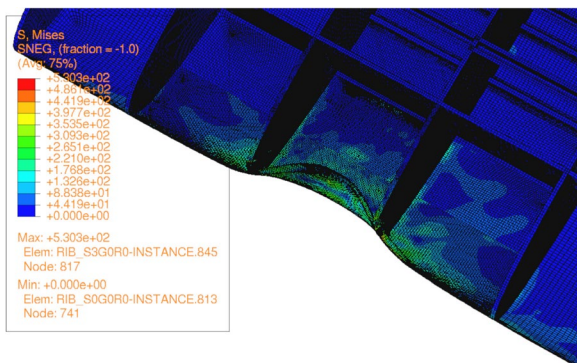


Fig. 18 The damaged structure of the initial optimal design. Note: the upper skin has been hidden

spar after load introduction reached 239.72 MPa, highlighting opportunities for further improvement.

4.3.1 Sequential approach

For this approach, no fantasizing is conducted, and the constrained improvement acquisition function is maximized at each iteration to determine the next design to be evaluated. Note that another strategy to reduce the optimization search space involved utilizing the minimum observed feasible weight. Since the initial optimal weight was rather low right after the variable screening process, much of the search space had zero constrained improvement due to having guaranteed larger weights. Consequently, the maximum thicknesses for these variables were adjusted to exclude regions with guaranteed null improvement. This adjustment was feasible due to the analytical formulation of weight, which ensured null improvement in these regions.

The optimization process proceeded for 64 iterations, out of which 12 were feasible. Figure 19 presents the evolution of the optimal weight over the optimization iteration, the final feasible design having a weight of 44.53 kg, representing a 15.31% reduction compared to the initial optimum after projection. The optimizer effectively exploited increased wing compliance to enhance kinetic energy absorption of the leading edge, thereby reducing the stresses on the spar while balancing the maximum penetration depth, the final design being depicted in Fig. 20. The final weight and constraint values, along with the weight reductions, are summarized in Table 8.

4.3.2 Batch approach

The unresolved or ongoing observations of optimization points were estimated by fantasizing data according to the Kriging believer method. In batch optimization, analyses

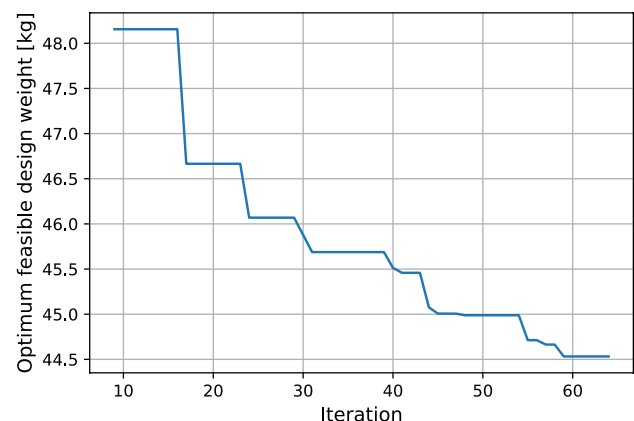


Fig. 19 Evolution of the optimal weight versus optimization iteration for the sequential approach

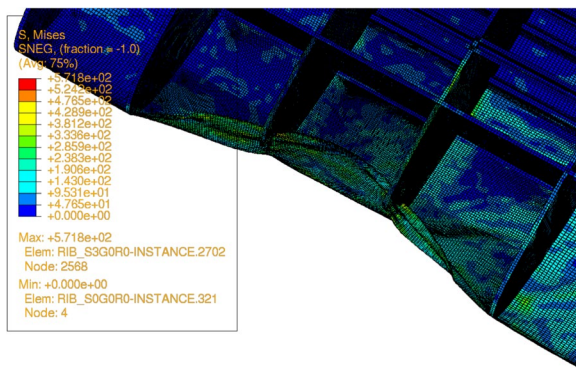


Fig. 20 Final optimal design in the sequential approach. Note: upper skin hidden

Table 8 Summary of optimal designs

	Sequential	Batch
Optimal weight [kg]	44.53	44.59
Weight reduction before projection [%]	43.51	43.44
Weight reduction after projection [%]	15.31	15.20
Penetration constraint value [mm]	235.32	235.61
Stress constraint value [MPa]	283.76	285.04
Cumulative wall time [hours]	1,338	802

typically do not complete simultaneously. Additionally, 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 the automatic stabilization parameters, thus increasing the wall time of a single simulation. To maximize resource utilization, results were saved immediately upon completion of each Abaqus workflow step, allowing the submission of additional design points based on available data and fantasized values whenever possible. The optimization search space was reduced in the same manner as for the sequential approach.

The optimization process comprised 89 points, with 14 feasible designs. However, the cumulative wall time was significantly smaller than for the sequential approach, with 33 days and 10 h versus 55 days and 18 h. Table 8 presents the improvements of the batch approach over the sequential. Moreover, Fig. 21 illustrates the evolution of the optimal weight with respect to the cumulative wall time. Convergence is reached faster, as depicted in Fig. 22, which presents the evolution of the acquisition function with respect to cumulative wall time.

In the end, the final optimal design exhibits a similar response to the one depicted in Fig. 20. Its weight is 44.59 kg, which represents a 15.20% reduction compared to the initial optimum after the variable screening process.

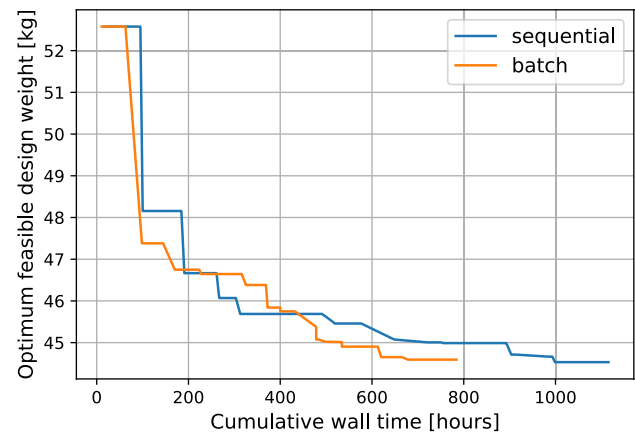


Fig. 21 Evolution of the optimal weight versus cumulative wall time

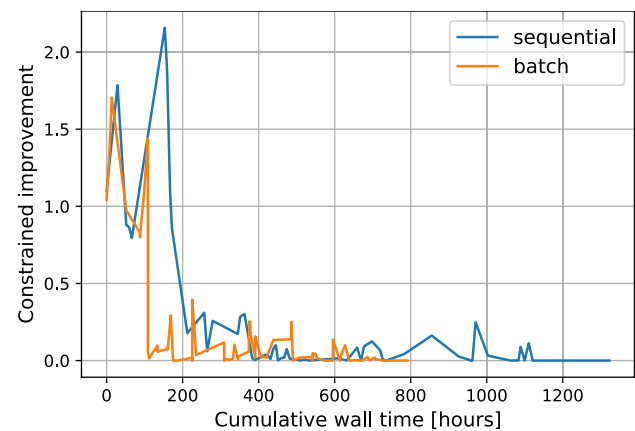


Fig. 22 Evolution of the constrained improvement acquisition function versus cumulative wall time

These findings substantiate the use of the ‘Kriging believer’ method in the case of computationally intensive analyses instead of the traditional sequential approach, as the optimum has nearly the same weight, but at only 60% of the wall time.

5 Conclusions and recommendations

This study successfully integrated bird strike crashworthiness requirements into a multidisciplinary optimization (MDO) framework for aircraft wing design, achieving significant weight savings while satisfying critical safety constraints which require highly non-linear analyses. By employing Bayesian optimization in conjunction with Kriging surrogate models and a variance-based variable ranking procedure, the methodology effectively managed the high-dimensional design space comprising 19 variables. The variable screening process reduced the

dimensionality to eight parameters, improving computational efficiency without compromising the reliability of the surrogate models.

Applied to a five-bay metallic wing segment, the optimization process attained a weight reduction of 43.44% compared to the lightest feasible design in the initial dataset before variable ranking. However, the study identified limitations in the surrogate modelling process, particularly concerning the penetration depth constraint, which exhibited a high RMSE. This indicates a need for improved surrogate accuracy for certain constraints. Additionally, the variable screening procedure, while effective in reducing dimensionality, cannot capture higher-order interactions among all design variables.

Recommendations for future research include:

1. **Include higher-order interactions:** The variable ranking procedure should be enhanced to identify and account for higher-order interactions among design variables.
2. **Develop a recovery mechanism:** The variable screening procedure would greatly benefit from a recovery mechanism against erroneous identification of the significant variables. In this particular case study, investigating the marginal effects revealed that the surrogate model was clearly inaccurate in a particular dimension, and engineering judgment could be applied. However, in more black-box type of functions, this may not be possible.
3. **Validation on other crashworthiness requirements:** The generalizability of the methodology should be validated across various design challenges.
4. **Validation on geometrical parameters:** The methodology should be tested against other design variables which would also change the architecture of the wing structure, such as spar positions, number of nose ribs, and different materials. Another valuable adjustment would be to vary the thickness along the wing span, as the farther the location is from the bird impact site, the smaller the effect is expected to be.

Dynamic relaxation of the explicit dynamic bird strike analysis

Dynamic relaxation of the explicit dynamic analysis is necessary when coupling the outcome of the explicit analysis, i.e., the damaged wing, with a subsequent static analysis. In the present work, the static analysis aims to evaluate the remaining static strength of the damaged wing. The dynamic relaxation consists of reducing the kinetic energy and reaching a quasi-static equilibrium after the bird strike event, preserving all plastic deformation and damaged state.

During the initial phases of this work, it was observed that without incorporating damping in the explicit dynamic simulations, the kinetic energy of the wing decreased very slowly or sometimes not at all. While extending the simulation time could mitigate this issue, the absence of energy dissipation is unrealistic since metallic structures inherently exhibit some damping. Therefore, a damping ratio of 0.02 was sought to be enforced, considered a reasonable estimate for continuous metallic structures as recommended by Orban (2011).

However, defining a constant damping ratio directly in Abaqus/Explicit is not feasible. As a result, Rayleigh damping can be employed, which assumes that the damping matrix is a linear combination of the mass and stiffness matrices, as given in Eq. A1, where α and β are the mass and stiffness-proportional damping coefficients, respectively. For a given mode with natural frequency ω , the corresponding damping ratio ξ is related to the Rayleigh coefficients by Eq. A2.

$$C = \alpha M + \beta K \quad (A1)$$

$$\xi = \frac{\alpha}{2\omega} + \frac{\beta\omega}{2} \quad (A2)$$

However, as demonstrated by Dassault Systèmes (2023b), the stiffness-proportional damping coefficient β can significantly reduce the stable time increment in explicit dynamic simulations. To circumvent this issue, only mass-proportional damping is applied by setting $\beta = 0$ and calculating the mass damping coefficient α using the first natural frequency from Eq. A2. For higher frequencies, this approach results in a reduced damping factor, preventing overdamping.

Tools

Addressing the time-consuming and repetitive nature of exploring multiple structural concepts, complicated by the involvement of various disciplines, GKN Fokker's Center of Competence in Design developed a Knowledge-Based Engineering (KBE) multidisciplinary design system (van den Berg and van der Laan 2021). The Multidisciplinary Modeller (MDM) is a Python package that automates the generation of products like flaps, wingboxes, and movables.

By allowing users to define these products through Python dictionaries, MDM facilitates efficient design space exploration for trade studies and enables rapid, consistent generation of analysis models. Moreover, MDM encompasses various modules such as an automatic mesh generator 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, and was used extensively here for data generation.

In addition to the tools previously mentioned, two open-source Python packages were utilized throughout this study: the Surrogate Modeling Toolbox (SMT) (Saves et al 2024), and Trieste (Picheny et al 2023). Both packages provide continuous and mixed-integer surrogate models, benchmark functions, and optimization algorithms. SMT was employed for the variable ranking procedure, while Trieste was used for Bayesian optimization on the reduced design space. Apart from Abaqus, which was necessary for finite element analyses, no other commercial software was used in this project, ensuring that the methodologies developed are accessible and reproducible.

Author contributions R.I. Ciobotia: Methodology, software, formal analysis, writing of the original draft. T. van der Laan, W. van de Waerdt and D. Peeters: supervision, methodology, review and editing of the original draft. S.G.P. Castro: Conceptualization, supervision, methodology, writing of the original draft, resources, funding acquisition, data curation. All authors have read and agreed to the published version of the manuscript.

Funding This research is conducted within the research and innovation programme Luchtvaart in Transitie, which is co-funded by the Netherlands National Growth Fund. It is part of the 'H2Crash' project, and part of The CrashProof Knowledge Centre (Castro2022, CrashProofLab).

Data availability All models, optimization algorithms and datasets generated and analysed during this study are publicly available under a BSD-3 license (Ciobotia et al 2025).

Declarations

Replication of results Results-only reproduction: Using the archived dataset, scripts are available to account for every figure and table in the paper (variable screening heatmaps, error metrics, acquisition function values). For full FEM and optimization reproduction: The repository contains scripts and job templates to (a) generate the initial designs, (b) run Abaqus jobs, (c) extract results, (d) fit ordinary Kriging surrogates, (e) perform variable screening, and (f) run constrained Bayesian optimization in either sequential or batch mode.

Conflict of interest On behalf of all authors, the corresponding authors state that there is no Conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ba S (2015) SLHD: Maximin-Distance (Sliced) Latin Hypercube Designs. <https://cran.r-project.org/package=SLHD>. Accessed 12 November 2024
- Ba S, Myers WR, Brenneman WA (2015) Optimal sliced latin hypercube designs. *Technometrics* 57(4):479–487. <https://doi.org/10.1080/00401706.2014.957867>
- Bellman R, Kalaba R (1959) A mathematical theory of adaptive control processes. *Proc Natl Acad Sci* 45(8):1288–1290. <https://doi.org/10.1073/pnas.45.8.1288>
- 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*, pp 2021–3079. <https://doi.org/10.2514/6.2021-3079>
- Bisagni C, Lanzi L, Ricci S (2002) Optimization of helicopter sub-floor components under crashworthiness requirements using neural networks. *J Aircr* 39(2):296–304. <https://doi.org/10.2514/2.2927>
- Castro S (2022). The CrashProof Knowledge Centre (CrashProofLab). <https://doi.org/10.5281/zenodo.11500722>
- Ciobotia RI (2024) Incorporating bird strike crashworthiness requirements within an MDO framework. Master's Thesis, Delft University of Technology. <https://resolver.tudelft.nl/uuid:d8d7f2a0-d1ba-46b6-affd-c2903a527fc1>
- Ciobotia RI, Peeters D, van der Laan T, van de Waerdt W, Castro S (2025). Dataset for incorporating bird strike crashworthiness requirements within the design of wing structures. <https://doi.org/10.5281/zenodo.15411557>
- Craig KJ, Stander N, Dooge DA, Varadappa S (2005) Automotive crashworthiness design using response surface-based variable screening and optimization. *Eng Comput* 22(1):38–61. <https://doi.org/10.1108/02644400510572406>
- Dassault Systèmes (2022) *STATIC. <https://docs.software.vt.edu/abaqusv2023/English/SIMACAEKEYRefMap/simakey-r-static.htm>. Accessed 11 November 2024
- Dassault Systèmes (2023a) Damage initiation for ductile metals. <https://docs.software.vt.edu/abaqusv2023/English/SIMAC AEMATRefMap/simamat-c-damageinitductile.htm>. Accessed 11 November 2024
- Dassault Systèmes (2023b) Material damping. <https://docs.software.vt.edu/abaqusv2023/English/SIMACAEMATRefMap/simamat-c-dampingopt.htm>. Accessed 9 November 2024
- Eurocontrol (2025) European Aviation Overview 2024. <https://www.eurocontrol.int/publication/eurocontrol-european-aviation-overview-archive-2024>. Accessed 6 March 2025
- Forrester AIJ, Sobester A, Keane AJ (2008) *Engineering design via surrogate modelling*. Wiley
- Ginsbourger D, Le Riche R, Carraro L (2008) A Multi-points criterion for deterministic parallel global optimization based on gaussian processes. <https://hal.science/hal-00260579>
- Goraj Z, Kustron K (2018) Review of current research trends in bird strike and hail impact simulations on wing leading edge. *Aircr Eng Aerosp Technol* 90(4):602–612. <https://doi.org/10.1108/AEAT-02-2017-0053>
- International Civil Aviation Organization (2023) 2016–2021 Wildlife strike analyses (IBIS). <https://www.icao.int/Aerodromes/ibis/Documents/03>
- Jones DR (2001) A taxonomy of global optimization methods based on response surfaces. *J Global Optim* 21:345–383. <https://doi.org/10.1023/A:1012771025575>
- Jones DR, Schonlau M, Welch WJ (1998) Efficient global optimization of expensive black-box functions. *J Global Optim* 13:455–492. <https://doi.org/10.1023/A:1008306431147>

- 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*. <https://doi.org/10.3390/ma14154268>
- Lanzi L, Bisagni C, Ricci S (2004) Crashworthiness optimization of helicopter subfloor based on decomposition and global approximation. *Struct Multidiscip Optim* 27(5):401–410. <https://doi.org/10.1007/s00158-004-0394-z>
- Lesuer DR (2009) Experimental Investigations of Material Models for Ti-6Al-4V Titanium and 2024-T3 Aluminum. Tech. rep., U.S. Department of Transportation, Federal Aviation Administration, <https://apps.dtic.mil/sti/citations/ADA384431>
- Marulo F, Guida M (2014) Design criteria for birdstrike damage on windshield. *Adv Aircr Spacecr Sci* 1:2014
- Ollar J, Toropov V, Jones R (2017) Sub-space approximations for mdo problems with disparate disciplinary variable dependence. *Struct Multidiscip Optim* 55(1):279–288. <https://doi.org/10.1007/s00158-016-1496-0>
- Orban F (2011) Damping of materials and members in structures. *J Phys: Conf Ser* 268(1):012022. <https://doi.org/10.1088/1742-6596/268/1/012022>
- Pahange H, Abolbashi MH (2016) Mass and performance optimization of an airplane wing leading edge structure against bird strike using taguchi-based grey relational analysis. *Chin J Aeronaut* 29(4):934–944. <https://doi.org/10.1016/j.cja.2016.06.008>
- Picheny V, Berkeley J, Moss HB, Stojic H, Granta U, Ober SW, Artemev A, Ghani K, Goodall A, Paleyes A, Vakili S, Pascual-Diaz S, Markou S, Qing J, Loka NRBS, Couckuyt I (2023) Trieste: Efficiently exploring the depths of black-box functions with TensorFlow. 10.48550/ARXIV.2302.08436
- Rosenbrock HH (1960) An automatic method for finding the greatest or least value of a function. *Comput J* 3(3):175–184. <https://doi.org/10.1093/comjnl/3.3.175>
- Saves P, Lafage R, Bartoli N, Diouane Y, Bussemaker J, Lefebvre T, Hwang JT, Morlier J, Martins JR (2024) SMT 2.0: A Surrogate Modeling Toolbox with a focus on hierarchical and mixed variables Gaussian processes. *Advances in Engineering Software* 188:103571. <https://doi.org/10.1016/j.advengsoft.2023.103571>
- Schonlau M, Welch WJ (2006) Screening the input variables to a computer model via analysis of variance and visualization. Springer, NY
- Schonlau M, Welch WJ, Jones DR (1998) Global versus local search in constrained optimization of computer models. *Inst Math Stat*. <https://doi.org/10.1214/lnms/1215456182>
- Schuhmacher G, Murra I, Wang L, Laxander A, O’Leary OJ, Herold M (2002) Multidisciplinary design optimization of a regional aircraft wing box. 9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization. 10.2514/6.2002-5406
- Siemann MH, Ritt SA (2019) Novel particle distributions for sph bird-strike simulations. *Comput Methods Appl Mech Eng* 343:746–766. <https://doi.org/10.1016/j.cma.2018.08.044>
- Sun G, Tian Y, Wang R, Fang J, Li Q (2020) Parallelized multiobjective efficient global optimization algorithm and its applications. *Struct Multidiscip Optim* 61(2):763–786. <https://doi.org/10.1007/s00158-019-02417-1>
- 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. <https://resolver.tudelft.nl/uuid:3a4a1c1b-5780-49f9-beed-5c26b255b34b>
- Vershinin VV (2015) Validation of metal plasticity and fracture models through numerical simulation of high velocity perforation. *Int J Solids Struct* 67–68:127–138. <https://doi.org/10.1016/j.ijsolstr.2015.04.007>
- Viana FA, Haftka RT, Steffen V (2009) Multiple surrogates: how cross-validation errors can help us to obtain the best predictor. *Struct Multidiscip Optim* 39(4):439–457. <https://doi.org/10.1007/s00158-008-0338-0>
- Viana FA, Gogu C, Goel T (2021) Surrogate modeling: tricks that endured the test of time and some recent developments. *Struct Multidiscip Optim* 64(5):2881–2908. <https://doi.org/10.1007/s00158-021-03001-2>
- Wang Z, Almeida JHS Jr, St-Pierre L, Wang Z, Castro SG (2020) Reliability-based buckling optimization with an accelerated kriging metamodel for filament-wound variable angle tow composite cylinders. *Compos Struct* 254:112821. <https://doi.org/10.1016/j.compstruct.2020.112821>
- Wang Z, Almeida JHS, Ashok A, Wang Z, Castro SGP (2022) Lightweight design of variable-angle filament-wound cylinders combining kriging-based metamodels with particle swarm optimization. *Struct Multidiscip Optim*. <https://doi.org/10.1007/s00158-022-03227-8>
- Welch WJ, Buck RJ, Sacks J, Wynn HP, Mitchell TJ, Morris MD (1992) Screening, predicting, and computer experiments. *Technometrics* 34(1):15–25. <https://doi.org/10.1080/00401706.1992.10485229>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.