

Surrogate-based design optimization of the binder cover combining performance and production cost

Pavel Eremeev ^{1,2,✉}, Hendrik Devriendt ^{1,2}, Alexander De Cock ³ and Frank Naets ^{1,2}

¹ KU Leuven, Belgium, ² Flanders Make@KU Leuven, Belgium, ³ Flanders Make, Belgium

✉ pavel.eremeev@kuleuven.be

Abstract

This study integrates surrogate models into combined design optimization of a binder cover, considering production cost and performance constraints. Results reveal that models trained on substantial datasets achieve designs close to the global optimum. Incorporating model variance into constraints prediction in surrogate-based optimization improves robustness and accuracy, especially with noisy functions. This modification enhances the likelihood of obtaining feasible designs, reducing computational demands and showcasing the potential of smaller datasets in predicting local optima.

Keywords: artificial intelligence (AI), surrogate-based design optimisation, product design

1. Introduction

Design optimization strategies are essential tools in engineering product development and modernization. Although performance-based design optimization is common, the integration of production aspects such as cost, manufacturing, and assembly processes, is rarely observed in design optimization and is seldom combined with performance metrics. This scarcity is primarily attributed to the challenges in quantitatively evaluating production aspects and the high computational demands associated with combined design optimization. This arises from the computational demand needed to evaluate performance and production metrics based on the product's design. The inherent unknown, typically non-linear behaviour of functions, and the possible presence of discrete functions and variables further compound these challenges (Nocedal and Wright 1999).

Incorporating state-of-the-art advancements in machine learning offers a promising approach to mitigate the complexities of the described design optimization problem through surrogate modelling techniques (Roy, Hinduja and Teti 2008). By leveraging machine learning capabilities, complex and resource-intensive functions used in direct design evaluations can be replaced with surrogates constructed from limited datasets, offering a more efficient alternative (Forrester and Keane 2009). Surrogate-based design optimization (SBDO) significantly alleviates the computational burden inherent in the design optimization process (Alizadeh and Mistree 2020).

Performance-based design optimization, utilizing surrogate modeling techniques, has proven successful in various domains, including electrical (Li, et al. 2022) and mechanical applications (Jin, Lu and Guan 2022). Despite these successes, the integration of manufacturing aspects, along with production cost considerations, remains a relatively unexplored dimension in regular and surrogate-based design optimization. In this study, we introduce the application of Gaussian Process (GP) and Artificial Neural Networks (ANN) regression models to optimize the design of a binder cover. Our approach takes into account both performance requirements and production cost considerations, incorporating the intricacies of the manufacturing process. The complexity of the optimization

challenge is further compounded by the presence of discrete design variables (Stork, et al. 2020) and noisy functions (Palar, et al. 2019).

We demonstrate the impact of the training dataset size on the accuracy of constructed surrogates and the quality of potential optimal designs obtained in the presence of noisy functions, highlighting the limitations of static sampling techniques. Additionally, we illustrate how GP prediction variance can be used to construct confidence intervals for the predictions, enhancing the likelihood of discovering feasible local optima and accounting for potential prediction errors, particularly in scenarios with a limited training dataset.

2. Test case description

To illustrate the complexities outlined in the introduction regarding the challenges of combined design optimization, integrating computationally expensive production and performance metrics, we turn our attention to a practical example. The test case involves the design optimization of a plastic cover for an agriculture binder machine, considering both performance and production cost.

The initial design of the cover, presented in Figure 1, consists of eight parts. This design has been deliberately simplified, excluding auxiliary elements such as fasteners and handles, given their negligible impact on the performance and cost metrics considered in this study. Two plastic elements – an inner layer with stiffening elements (beads) and an outer layer – are produced using thermoforming with 10% fiberglass-reinforced ABS. Additionally, two steel channels are positioned between the inner and outer layers to reinforce areas where gas lifts attach to the cover. At the bottom of the inner layer, two steel plates are attached to each of the channels to connect the gas lifts to the cover. This elements are depicted in the center part of Figure 1.

The primary objective in this case is to optimize the inner and outer layers of the plastic cover. These parts are parameterized with four design variables, detailed in Table 1. It is important to note that the thickness of both plastic elements remains constant at 5 mm throughout the optimization process.

Table 1. Description of the design parameters

Design parameters		Type	Initial value	Lower bound	Upper bound
x_1	Length of the bead base, L_{bead} (mm)	Continuous	120	50	150
x_2	Width of the bead base, W_{bead} (mm)	Continuous	70	50	150
x_3	Number of vertically aligned beads (columns), N_c	Discrete	8	{4,6,8,10}	
x_4	Number of horizontally aligned beads (rows), N_r	Discrete	7	{5,7,9,11}	

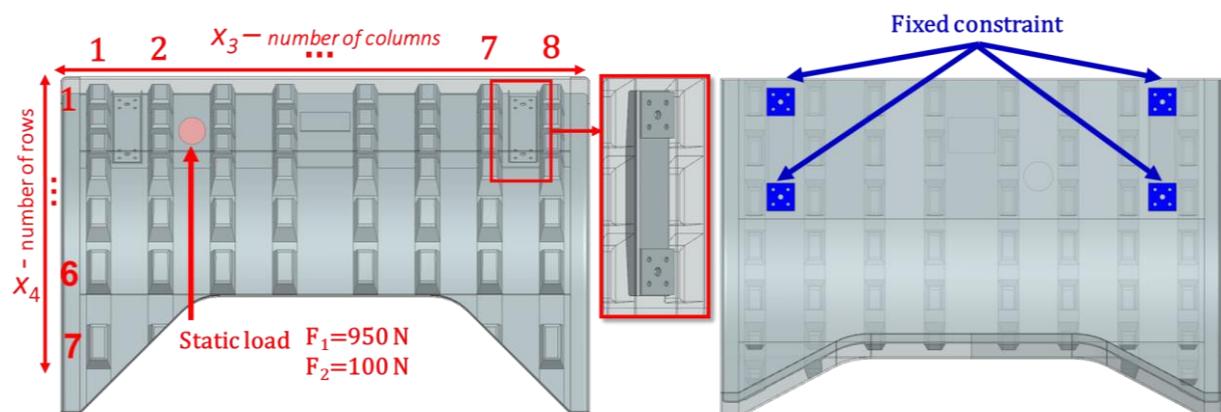


Figure 1. Initial test case design and boundary conditions used for performance evaluation

2.1. Performance requirements

The performance requirements for the binder cover predominantly adhere to the standards outlined in ISO 4254 Agricultural Machinery. Within the scope of this work, we specifically focus on two requirements that are influenced by the design of the plastic cover:

- The cover must withstand a 95 kg mass applied over a 100 mm diameter without damage.
- The maximum distortion under a 10 kg load, applied on a 100 mm diameter, is limited to 4 mm.

Mathematically, these performance requirements are expressed as constraints for design optimization:

$$\max(\sigma_{mis}) \leq \sigma_{ts} \quad (1)$$

$$\max(S) \leq 4 \text{ mm} \quad (2)$$

Here, σ_{mis} represents the von Mises stress in the plastic parts, with $\sigma_{ts} = 45$ MPa denoting the tensile stress of the outer layer (fiberglass-reinforced ABS), and S indicating the displacement in the plastic parts. Performance constraints are assessed using a Finite Element (FE) model. Stress and displacement of the plastic parts are determined by applying a static loads with amplitudes of 100 N and 950 N on the outer layer of the cover, targeting the circle with a 100 mm diameter. Fixed constraints are applied to the bottom face of the metal plates, which connect the gas lifts to the cover. The boundary conditions are illustrated in Figure 1.

The position of the area where the static load is applied is not explicitly defined by the standard. To account for the most challenging conditions, the load is applied to the area with the lowest stiffness. The stiffness of the outer plastic layer is influenced by the configuration of the beads, with the lowest stiffness found between the beads. Considering that the configuration of the beads is subject to change during the optimization process, the position of the area with the lowest stiffness, where the load is applied, also dynamically adjusts in response to these variations.

2.2. Production cost evaluation

The production process of the binder cover, including the manufacturing of inner and outer plastic layers and the assembly operations, is illustrated in Figure 2. As the optimization focuses solely on the design of the plastic elements, only the production steps related to these elements are impacted. The change of the characteristic dimensions and total mass of the plastic parts and respectively the cost of thermoforming do not significantly change during optimization. The primary influence of design optimization is on the cost of plasma treatment and glue positioning.

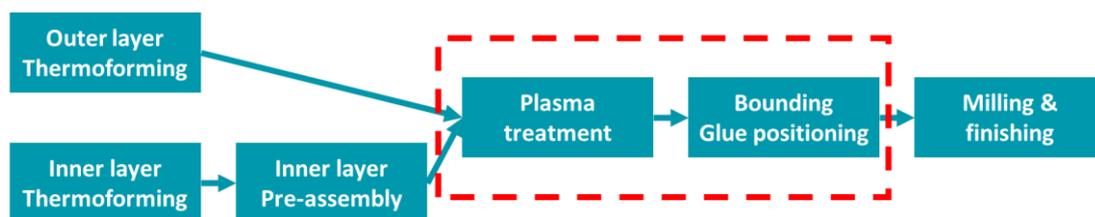


Figure 2. Production process of the binder cover

The total production cost (C_{total}) is defined as the sum of material and processing ($C_{processing}$) cost. Considering only gluing and plasma treatment, the material cost involves only the cost of glue (C_{glue}):

$$C_{total} = C_{glue} + C_{processing} \quad (3)$$

Glue is applied on the top flange of the beads (highlighted in Figure 3) along the bead's center, with fixed dimensions: width (W_g) and thickness (T_g) equal to 10 and 0.5 mm respectively.

The length of glue line equals the length of the bead (L_{bead} , parameter x_1). The volume (V_{glue}) and cost of the applied glue (C_{glue}), as functions of design parameters, are given by:

$$V_{glue} = ((N_r - 1) \cdot N_c + 2) \cdot L_{bead} \cdot W_g \cdot T_g \quad (4)$$

$$C_{glue} = V_{glue} \cdot p_{tv} \quad (5)$$

Here, $p_{tv} = 2 \text{ €/mm}^3$ represents the glue price per cubic mm, and the bottom row of beads always has only two elements.

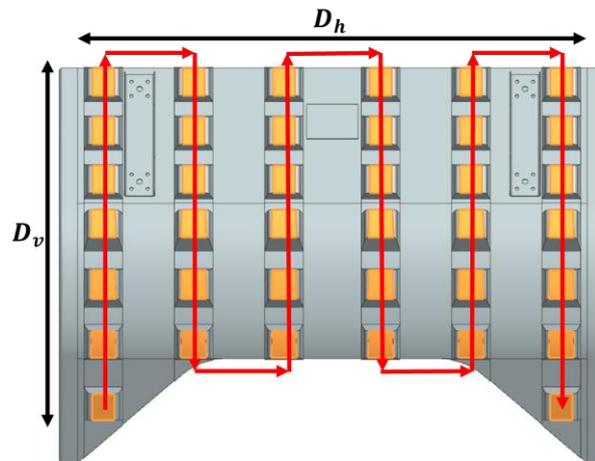


Figure 3. Processing cost estimation; path of plasma treatment and glue application

Both plasma treatment and glue application are executed using automatic robot systems, and the path for both operations is schematically depicted in Figure 3. The processing cost ($C_{processing}$) is estimated based on the total processing time, comprising movement time ($T_{movement}$) and positioning time ($T_{positioning}$) given by Equation 6. Here, $p_o = 10 \text{ €/hour}$ represents the total operating cost of the robotic systems employed for plasma treatment and glue application.

Movement time is computed using the total length of the operation path (L_{path}), determined by the characteristic dimensions of the part (vertical and horizontal dimensions D_v and D_h as illustrated in Figure 3) and the movement speed of the robot arm ($v = 2.5 \text{ mm/s}$). Positioning time accommodates a pause ($\tau = 10 \text{ s}$) for precise positioning at the edges of each bead. Details of the movement and positioning time calculations are presented in Equations 7-8.

$$C_{processing} = (T_{positioning} + T_{movement}) \cdot p_o \quad (6)$$

$$T_{positioning} = ((N_r - 1) \cdot N_c + 2) \cdot \tau \quad (7)$$

$$T_{movement} = L_{path}/v \quad (8)$$

$$L_{path} = D_v \cdot N_r + D_h \quad (9)$$

The presented approach for production cost evaluation is intended for rough estimation, serving as a proof of concept and providing an illustrative example of a function with mixed inputs (typical manufacturing and/or assembly process cost) suitable for SBDO. In this specific example, production cost can be directly evaluated using optimization variables as inputs. However, in more complex cases, it might be necessary to consider the entire CAD model of the product as an input, which can significantly increase problem complexity.

3. Single-objective design optimization

For the test case described in Section 2, a single-objective design optimization problem is formulated. The objective function is the total production cost, with two nonlinear constraints restricting maximum stress ($\sigma_{mis}(\mathbf{x})$) and maximum displacement ($S(\mathbf{x})$).

$$\begin{cases} \min_{\mathbf{x}}: C_{total}(\mathbf{x}) \\ s. t. : \max(S(\mathbf{x})) - 4 \leq 0 \\ \max(\sigma_{mis}(\mathbf{x})) - \sigma_{ts} \leq 0 \\ \mathbf{x} = \{x_1, \dots, x_4\} \end{cases} \quad (10)$$

The solution of the specified design optimization problem requires a function that defines a mapping between design variables and values of the performance and manufacturing process metrics. This mapping is performed with the help of the automated framework for design evaluation, which uses a set of design variables as an input and provides values of the preliminary specified performance and manufacturing process metrics as an output. Details of the framework functionality can be found in the previous research (Eremeev, et al. 2024). The framework is automated with the help of Python-based API – NX Open, which uses Siemens NX for CAD processing operations and FE calculation. Then, the framework can be used directly as an optimization function in direct optimization or for evaluation of the sample points in the design space for surrogate model construction. Evaluation of one point in design space (one function evaluation) involves an update of the CAD model, evaluation of the FE model for stress and displacement and estimation of the production cost.

In this work number of function evaluations or data points is used as a measure of computational efficiency or budget. One function evaluation, e.g. evaluation of one data point, takes 210-300 seconds (computations run on HP ZBook Power G9; Intel Core i7-12800H @2.4 GHz; 32 GB RAM).

3.1. Direct optimization

The single-objective optimization problem described in Equation 10 faces several challenges. It involves discrete and continuous optimization variables, and there is no information available on the behaviour of the objective and constraint functions, making their gradients unavailable. This restriction necessitates the use of derivative-free algorithms. To address the discrete nature of some variables, a genetic algorithm (GA) is employed, given its suitability for mixed-integer problems. In this example, the MATLAB implementation of the GA (Deep, et al. 2009) is utilized.

The results of direct optimization are presented in Table 2. Two different local optima were found depending on the maximum number of function evaluations used as the stopping criterion (2600 and 1000 evaluations). The best solution was obtained in the run restricted to 2600 function evaluations.

Direct optimization, while providing valid solutions, involves significant computational burden due to the high number of required function evaluations. This makes obtaining a global optimum within a reasonable time challenging. Additionally, each attempt to run the optimization algorithm requires careful consideration of computational resources.

Table 2. Results of direct optimization, genetic algorithm

Design parameters and optimization metrics	Optimal design variables and values of the metrics of interest	
	2600 function evaluations	1000 function evaluations
L_{bead} , mm	52.7	64.7
W_{bead} , mm	148.9	81.2
N_c	6	6
N_r	11	11
C_{total} , €	47.6	55.5
Max Displacement, mm	0.8	1.8
Max Stress, MPa	44.9	44.9

3.2. Surrogate-based design optimization

An alternative approach in dealing with the described design optimization problem is surrogate-based optimization. Rather than executing expensive function evaluations directly within the optimization routine, a computational budget is allocated to sample the design space and construct fast-to-compute surrogates. These surrogate models of the objective and constraint functions are then employed in the design optimization process. This approach helps significantly reduce the computational budget while improving the quality of the obtained solution (Alizadeh and Mistree 2020). The workflow of the SBDO is depicted in Figure 4.

3.2.1. Surrogate model construction

The design space is sampled using the Latin Hypercube sampling technique, a common method for random sampling of the design space (M. D. Mckay and Conover 2000). Three datasets, containing 100, 200, and 300 points, are considered. Gaussian Process (GP) and Artificial Neural Networks (ANN) regression models are utilized for surrogate construction, given their flexibility in dealing with functions of varying complexity and high accuracy with nonlinear functions (Jin, Chen and Simpson 2001).

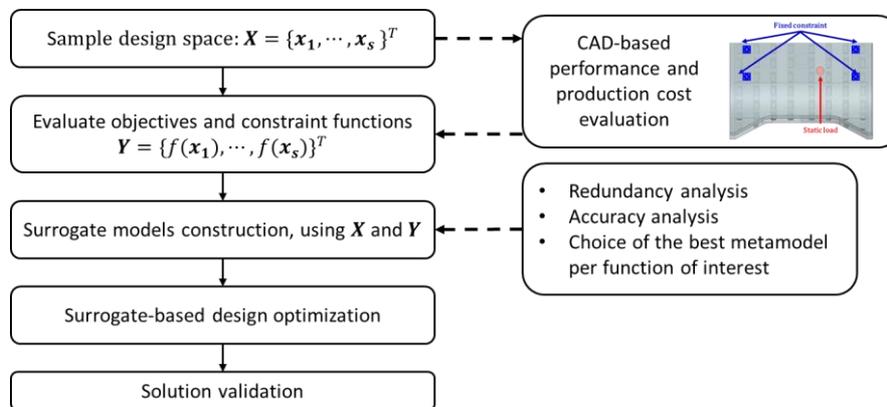


Figure 4. Workflow of the surrogate-based design optimization

In the GP regression, the function of interest is represented as a realization of a Gaussian process:

$$\hat{y} = \beta^T f(\mathbf{x}) + \sigma^2 Z(\mathbf{x}) \quad (11)$$

Where the first part is the mean of the Gaussian process, consisting of a collection of arbitrary functions, $f(\mathbf{x})$, with weighing coefficients β . The second part consists of GP variance σ^2 and zero-mean, unit variance stationary Gaussian process $Z(\mathbf{x})$. This stationary process is defined by a correlation function ($R(\mathbf{x}, \mathbf{x}^*, \theta)$), characterizing the correlation between two samples with respect to the hyperparameter θ . GP models were developed using the Matlab implementation of universal kriging with a linear mean ($\beta^T f(\mathbf{x})$) and Gaussian correlation function (Rasmussen 2004).

Meanwhile, ANN models utilized a feedforward, fully connected neural network with two hidden layers, each consisting of 50 neurons. Mathematically, the behaviour of a neuron is described as follows:

$$y_k = \phi\left(\sum_{j=1}^m w_{kj} x_j + b_k\right) \quad (12)$$

Here, x_1, \dots, x_m represent the inputs, w_{k1}, \dots, w_{km} are the weights of the neurons, b_k is the bias and $\phi(\cdot)$ denotes the activation function, in this case Rectified Linear Unit (ReLU) function (Hagan, et al. 2014). Redundancy analysis, using least absolute shrinkage and selection operator (LASSO) regression analysis (Lee, Shi and Gao 2022), was performed to identify possible redundant variables for all considered functions. In cases where redundant variables were detected, regression models were trained twice, with the full and reduced datasets, to avoid false detection. To further reduce the number of function evaluations, all generated sampling points were used in training. Since a separate testing dataset was not available, surrogate accuracy was compared using 10-fold cross-validation.

Predictive (pred.) and average cross-validation (c-val.) errors for all considered functions are presented in Table 3, in both cases mean square error is used as a loss function. Best models, chosen for each function of interest, are highlighted in bold.

For the objective function, the correct detection of the redundant variable (x_2 representing W_{bead}) was only feasible with larger datasets (200 and 300 points). However, in the case of the second constraint function (maximum stress), false detection of the redundant variable occurred. Nevertheless, cross-validation error analysis revealed it to be a false positive. The accuracy of the GP models surpassed that of all considered functions. In this study, hyperparameter optimization was not employed during the construction of regression models. Since the quality of ANN models is highly dependent on their architecture, the performance of ANN models with a fixed architecture was inferior to that of GA models, which are inherently more flexible.

Table 3. Accuracy of the established surrogate models

Model type	100 points		200 points		300 points	
	Pred. error	C-val. error	Pred. error	C-val. error	Pred. error	C-val. error
Objective function, C_{total}						
GP	1.62e-3	0.55	1.3e-3	6.71e-2	7.86e-4	9.25e-2
ANN	5.69	7.89	5.70	6.26	6.17	6.67
GPr (1 3 4)	–	–	3.41e-4	3.12e-3	2.58e-4	7.12e-2
ANNr (1 3 4)	–	–	5.61	5.77	5.97	4.58
Constraint #1 (maximum displacement)						
GP	2.17e-4	4.51e-3	4.46e-4	1.88e-3	4.89e-4	4.96e-3
ANN	1.42e-3	1.21e-2	2.98e-3	1.30e-2	8.88e-3	8.11e-3
Constraint #2 (maximum stress)						
GP	5.60e5	5.34e6	1.12e6	4.39e6	1.33e6	3.61e6
ANN	1.70e7	3.06e7	1.55e7	2.23e7	1.76e7	1.95e7
GPr (1 3 4)	1.12e7	1.78e7	1.19e7	1.55e7	1.23e7	1.50e7
ANNr (1 3 4)	1.76e7	2.69e7	1.80e7	1.99e7	1.83e7	2.05e7

GPr/ANNr – models trained with reduced dataset. Indexes of the considered design variables are in parenthesis.

3.2.2. Results of surrogate-based design optimization

The optimization results for the single-objective design problem, as formulated in Equation 10, are presented in Table 4. In this instance, a surrogate-based approach was employed, solving the optimization problem using the same GA configuration as in the direct optimization, discussed earlier. To ensure robustness, the GA was run 10 times for each set of models, and all distinct solutions were gathered for subsequent validation.

Table 4. Results of surrogate based optimization and solution validation

	Optimal design variables and predictions for metrics of interest					
	100 points		200 points		300 points	
L_{bead} , mm	67.78		53.47		54.86	
W_{bead} , mm	144.97		144.97		145	
N_c	6		6		6	
N_r	9		11		11	
	<i>Predicted</i>	<i>True</i>	<i>Predicted</i>	<i>True</i>	<i>Predicted</i>	<i>True</i>
C_{total} , €	49.15	49.06	48.08	48.08	48.99	49.00
Max Displacement, mm	0.97	0.91	0.85	0.86	0.83	0.85
Max Stress, MPa	45.00	46.59	45.00	45.49	45.00	44.62
Prediction feasibility	0/9		0/3		2/4	
	Mean prediction error, %					
C_{total} , €	0.17		7.99e-3		3.66e-3	
Max Displacement, mm	6.16		2.09e-1		1.67	
Max Stress, MPa	4.39		1.97		1.76	

The best solutions found using models trained on datasets with 200 and 300 points are relatively close to the best solution obtained with direct optimization. However, the solution found using the smallest dataset has a different bead configuration (N_r value). All distinct solutions obtained for each set of surrogate models were validated by direct design evaluation. In Table 4, true and predicted values are given for the best-predicted optimum, and prediction error and feasibility (number of feasible solutions compared to the total number of solutions) are evaluated for all obtained design configurations.

The use of surrogate models in single-objective design optimization has yielded valuable insights. Specifically, surrogates trained on a substantial dataset of 300 points successfully identified two feasible solutions near the best optimum obtained through direct optimization. In contrast, models trained with 200 and 100 points struggled to predict feasible solutions, violating performance constraints. Analysis of mean prediction errors underscored the superior accuracy of models trained with larger datasets. However, the transition from a 200 to a 300-point dataset did not yield a significant improvement in accuracy, suggesting that blindly increasing the sampling dataset size using a space-filling criterion might not be practical. On the other hand, the application of adaptive sampling procedures emerged as a promising approach for SBDO, potentially reducing the necessity for expensive function evaluations. Despite the success in achieving designs close to the location of the global optimum with reduced computational costs, there are notable drawbacks. Consistently achieving feasible designs necessitated sufficiently large training datasets. Surrogates built with smaller datasets effectively guided optimization toward the global optimum but struggled to accurately predict stress constraint values.

3.2.3. Improving design feasibility in surrogate-based optimization through constraints variance incorporation

An examination of the stress values in the validated designs (Table 4) suggests an irregular, non-linear behaviour in stress distribution. In Figure 5, the distribution of maximum stress concerning bead width is illustrated for a fixed bead configuration ($N_c = 6$; $N_r = 11$) with a bead length (L_{bead}) of 54.5 mm. While stress generally follows a decreasing trend with increasing bead width, periodic oscillations with an amplitude of approximately 1.5-2 MPa introduce noticeable noise.

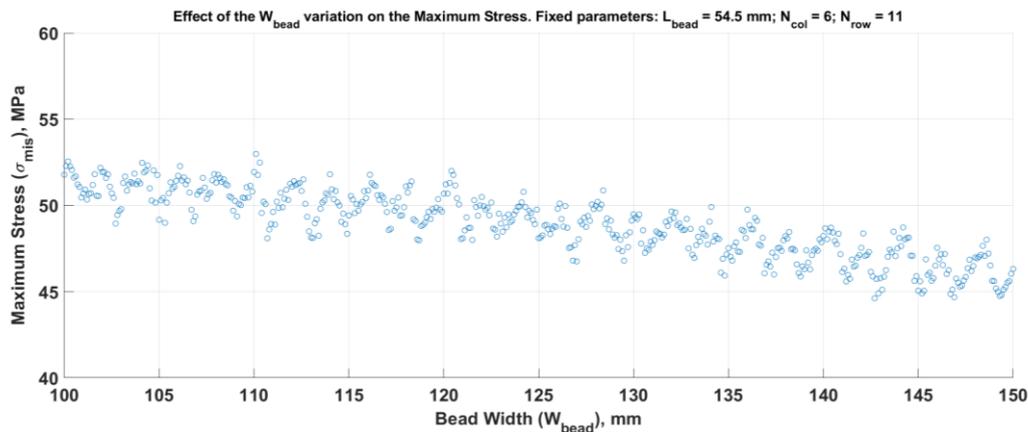


Figure 5. Distribution of the maximum stress at variable width of the beads (W_{bead})

These oscillations result from imperfections in modeling the contact between the inner and outer layers of the cover, simulated using gluing conditions. Despite applying finer mesh constraints around the load application area in the outer layer and the rounded edges of the beads in the inner layer, slight relative movement of mesh nodes between layers causes stress oscillations.

To address this noisy behaviour, alternative representations of gluing, such as mesh mating conditions, or adjustments to meshing constraints, could be considered. However, these methods impact computational time either indirectly, through increased model definition time, or directly, by affecting the computation time of the model evaluation due to finer meshing.

SBDO allows for a different approach to tackle such noisy functions, common in derivative-free optimization. The variance distribution from the GP model (Equation 11) can be incorporated into the prediction for the performance constraints. The modified optimization problem is expressed as follows:

$$\begin{cases} \min_x \hat{C}_{total}(\mathbf{x}) \\ s. t. : \hat{S}(\mathbf{x}) + \hat{Z}_S(\mathbf{x}) - 4 \leq 0 \\ \hat{\sigma}_{mis}(\mathbf{x}) + \hat{Z}_\sigma(\mathbf{x}) - \sigma_{ts} \leq 0 \\ \mathbf{x} = \{x_1, \dots, x_4\} \end{cases} \quad (13)$$

Here, \hat{C}_{total} , \hat{S} , and $\hat{\sigma}_{mis}$ represent the best model prediction for production cost, maximum displacement and maximum stress respectively; \hat{Z}_S and \hat{Z}_σ denote the GP model prediction variance for the maximum displacement and maximum stress, respectively. Incorporating the prediction variance into constraint functions serves to make optimization constraints more stringent than the original constraints. Essentially, it adds a dynamic margin to the constraint function. However, unlike a constant margin, the GP prediction variance varies based on the distance to the sample points. Consequently, near the sample points, the impact on constraint modification is minimal as confidence in the model prediction is high. Conversely, when constraints are evaluated far from the sampling points, a margin is added to ensure the feasibility of the solution. Solving this modified optimization problem produces predicted optima with a higher probability of feasibility. The results of this optimization, including constraints variance, are presented in Table 5.

Table 5. Results of surrogate based optimization with constraints variance

	Optimal design variables and predictions for metrics of interest					
	100 points		200 points		300 points	
L_{bead} , mm	61.59		61.87		60.39	
W_{bead} , mm	139.51		143.68		144.82	
N_c	6		6		6	
N_r	11		11		11	
	Predicted	True	Predicted	True	Predicted	True
C_{total} , €	53.41	53.44	53.62	53.62	52.63	52.65
Max Displacement , mm	0.89	0.89	0.81	0.83	0.80	0.82
Max Stress , MPa	43.20	43.71	43.19	43.91	43.13	43.53
Prediction feasibility	9/10		10/10		9/9	

In this scenario, all three sets of surrogates predict optima with the correct bead configuration ($N_c = 6$; $N_r = 11$). However, bead dimensions slightly deviate from the best optimum found with direct optimization due to a confidence interval, based on prediction variance, included in the optimization constraints. Solution validation, however, reveals that not only the best designs predicted by all sets of surrogates are feasible, but the majority of predicted optimal designs also lie in the feasible part of the design space. Additionally, even the model trained on the smallest dataset manages to predict a design in the vicinity of the best found optimum, superior to the results of direct design optimization after 1000 function evaluations. This approach enables a more economical use of the available computational budget, enhancing the quality of predictions provided by models trained on a limited dataset.

4. Conclusion

In this study, we explored the application of surrogate models for combined design optimization considering production cost along with performance constraints. The use of surrogate models, particularly Gaussian Process regression models, proved to be an effective strategy for approximating complex and computationally expensive functions. The single-objective design optimization of the binder cover demonstrated the capability of surrogate models to guide the optimization process efficiently. Additionally, we highlighted the importance of dataset size in training these models and the drawbacks of increasing the dataset size based solely on space-filling criterion.

The results of the surrogate-based optimization, including prediction variance in a form of confidence interval for optimization constraints, demonstrated a more economically efficient use of computational resources. The modified optimization problem has a higher probability of yielding feasible designs. This approach not only provided optimal design comparable with direct optimization with significantly smaller computational budget but also highlighted the potential of smaller datasets in predicting designs in the vicinity of the global optima. Moreover, the incorporation of variance information from surrogate models proved instrumental in enhancing the robustness and accuracy of the optimization process, particularly in scenarios with noisy function behaviours.

The methodology presented in the paper has a high level of automation to be applied to a wide spectrum of possible problems. There is no limitation on the type of product to be considered. The only requirement is the availability of a parametrized CAD model, along with precise models for production cost and performance evaluation. These models may be either data-driven or physics-based, varying in complexity. However, the software implementing these models must have a corresponding API to seamlessly integrate with the framework, enabling the utilization of design parameters from the CAD model as inputs. While these models are inherently case-dependent, the optimization procedure for both direct and surrogate-based design optimization can be fully automated once they are available.

Acknowledgement

The Research Fund KU Leuven is gratefully acknowledged for its support. This research was partially supported by Flanders Make, the strategic research centre for the manufacturing industry.

References

- Alizadeh, Reza., and Farrokh Mistree. 2020. "Managing computational complexity using surrogate models: a critical review." *Research in Engineering Design* 31: 275–298. <https://dx.doi.org/10.1007/s00163-020-00336-7>.
- Deep, Kusum, Krishna Pratap Singh, M. L. Kansal, and C. Mohan. 2009. "A real coded genetic algorithm for solving integer and mixed integer optimization problems." *Applied Mathematics and Computation* 212: 505–518. <https://doi.org/10.1016/j.amc.2009.02.044>.
- Eremeev, Pavel, Alexander De Cock, Hendrik Devriendt, and Frank Naets. 2024. "Framework for metamodel-based design optimization considering product performance and assembly process complexity." *SIMULATION* 0: 00375497231217301. <https://dx.doi.org/10.1177/00375497231217301>.
- Forrester, Alexander I. J., and Andy J. Keane. 2009. "Recent advances in surrogate-based optimization." *Progress in Aerospace Sciences* 45: 50–79. <https://doi.org/10.1016/j.paerosci.2008.11.001>.
- Hagan, M. T., H. B. Demuth, M. H. Beale, and O. De Jesús. 2014. *Neural Network Design (2nd Edition)*. United States: Martin Hagan. <https://books.google.be/books?id=4EW9oQEACAAJ>.
- Jin, R., Wei Chen, and Timothy Simpson. 2001. "Comparative Studies of Metamodeling Techniques Under Multiple Modeling Criteria." *Structural and Multidisciplinary Optimization* 23: 1–13. <https://dx.doi.org/10.1007/s00158-001-0160-4>.
- Jin, Xihong, Jun Lu, and Weiyuan Guan. 2022. "Crashworthiness analysis and multiobjective robust optimization of two-stage variable thickness expansion tube under impact loading." *Structural and Multidisciplinary Optimization* 65: 178. <https://dx.doi.org/10.1007/s00158-022-03267-0>.
- Lee, Ji Hyung, Zhentao Shi, and Zhan Gao. 2022. "On LASSO for predictive regression." *Journal of Econometrics* 229: 322–349. <https://doi.org/10.1016/j.jeconom.2021.02.002>.
- Li, Congbo, Mingli Huang, Wei Li, Ningbo Wang, and Jiadong Fu. 2022. "Optimization of an induction motor for loss reduction considering manufacturing tolerances." *Structural and Multidisciplinary Optimization* 65: 187. <https://dx.doi.org/10.1007/s00158-022-03276-z>.
- M. D. McKay, R. J. Beckman, and W. J. Conover. 2000. "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code." *Technometrics* (Taylor & Francis) 42: 55–61. <https://dx.doi.org/10.1080/00401706.2000.10485979>.
- Nocedal, Jorge, and Stephen J. Wright. 1999. *Numerical optimization*. Springer. <https://dx.doi.org/10.1007/b98874>.
- Palar, Pramudita Satria, Rhea Patricia Liem, Lavi Rizki Zuhail, and Koji Shimoyama. 2019. "On the use of surrogate models in engineering design optimization and exploration: the key issues." *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. New York, NY, USA: Association for Computing Machinery. 1592–1602. <https://dx.doi.org/10.1145/3319619.3326813>.
- Rasmussen, Carl Edward. 2004. "Gaussian Processes in Machine Learning." In *Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, February 2 - 14, 2003, Tübingen, Germany, August 4 - 16, 2003, Revised Lectures*, edited by Olivier Bousquet, Ulrike von Luxburg and Gunnar Rätsch, 63–71. Berlin, Heidelberg: Springer Berlin Heidelberg. https://dx.doi.org/10.1007/978-3-540-28650-9_4.
- Roy, Rajkumar, Srichand Hinduja, and Roberto Teti. 2008. "Recent advances in engineering design optimisation: Challenges and future trends." *CIRP Annals* 57: 697–715. <https://doi.org/10.1016/j.cirp.2008.09.007>.
- Stork, Jörg, Martina Friese, Martin Zaefferer, Thomas Bartz-Beielstein, Andreas Fischbach, Beate Breiderhoff, Boris Naujoks, and Tea Tušar. 2020. "Open Issues in Surrogate-Assisted Optimization." In *High-Performance Simulation-Based Optimization*, edited by Thomas Bartz-Beielstein, Bogdan Filipič, Peter Korošec and El-Ghazali Talbi, 225–244. Cham: Springer International Publishing. https://dx.doi.org/10.1007/978-3-030-18764-4_10.