Available online at www.eccomasproceedia.org
Eccomas Proceedia UNCECOMP (2023) 545-561

ECCOMAS
Proceedia

UNCECOMP 2023
5th ECCOMAS Thematic Conference on
Uncertainty Quantification in Computational Science and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou (eds.)
Athens, Greece, 12-14 June 2023

ROBUST DESIGN OPTIMIZATION UNDER EPISTEMIC UNCERTAINTY USING ADAPTIVE KRIGING AND EXTREME VALUE DISTRIBUTIONS

Augustin Persoons[1], Conradus Van Mierlo[1], Pierre Beaurepaire[2], David Moens[1]

[1] KU Leuven, Department of Mechanical Engineering, LMSD Div.
Jan De Nayerlaan 5, 2860, Sint-Katelijne-Waver, Belgium
e-mail: {augustin.persoons,koen.vanmierlo,david.moens}@kuleuven.be

[2] Clermont Auvergne INP, CNRS, Institut Pascal
Campus des Cézeaux, SIGMA Clermont, 63170 Aubière, France
e-mail: pierre.beaurepaire@sigma-clermont.fr

Abstract. *This article proposes a method to solve robust optimization problems under interval uncertainties. The goal of such a problem is to find the set of design variables minimizing the amplitude of performance variations due to the uncertain variables. It is proposed to solve it using the well-known framework of adaptive-Kriging optimization, e.g., EGO. In this context the complexity lies in the estimation of the Kriging counterpart of the quantity of interest, i.e., the width of the interval of performance for a given design, or more specifically its statistical moments. In this work an indicator of this quantity is explored based on the local extreme value distributions associated with the Gaussian process surrogate. This indicator allows us to reduce the computation burden compared to approaches based on confidence intervals. The performances are studied on analytical benchmark problems and show promising results.*

Keywords: Robustness, Design optimization, Adaptive Kriging, Interval uncertainty

ISSN:2623-3339
Peer-review under responsibility of the organizing committee of UNCECOMP 2023.
doi: 10.7712/120223.10357.19880

1 INTRODUCTION

Robustness is commonly defined as the ability of a system to maintain constant performances under variations of its inputs and is usually attributed to Taguchi [1, 2]. Together with reliability-based approaches, robustness consists in one of the major metrics used to include and alleviate the influence of uncertainties on engineering systems. While the definition is common, many metrics have been proposed to measure robustness [3], some deterministic [4, 5, 6, 7], some stochastic, i.e., using statistical moments [8, 9, 10]. This first kind is more suited for cases of large interval uncertainty. In this context computationally efficient approaches exist for particular system behaviours (e.g. linear or quadratic) based on interval algebra but in the general case the problem comes to a nested min-max problem.

Two-level optimization problems are very computationally intensive and are usually intractable when associated with an expensive system's model. Surrogate modelling techniques is an efficient way to alleviate this issue and many methods have been proposed in this framework [8]. The principle is to calibrate an emulator of the expensive numerical model using a limited number of evaluations called a Design of Experiments (DoE). While this calibration step is computationally intensive, once calibrated, the surrogate model can be used to solve the nested optimization problem with a very limited computation burden. One scheme have been particularly popular among surrogate-based technique is the scheme of adaptive Kriging. The principle is to take advantage of the Kriging's standard deviation, as an indicator of the localized prediction uncertainty. Using this quantity, schemes have been developed to iteratively build a near-optimal design of experiment aiming at minimizing the influence of the prediction uncertainty on a quantity of interest. EGO [11] for efficient global optimization is a very popular global optimization methods that falls in this category.

The main difficulty is how to efficiently solve the nested optimization problem with the adaptive-Kriging framework. The problem is that Kriging surrogates provides predictions in the shape of a Gaussian process and active learning schemes rely on the Gaussian nature of the prediction to guide the calibration. In the context of Robust design optimization the quantity of interest is not the prediction itself but its optimum on a given sub-space. While this optimum is a random variable its distribution is not available in closed form in the general case.

Several approaches have been proposed in the literature to implement EGO in this context, mainly relying on obtaining empirically its statistical moments [12], using metrics based on confidence intervals [13] or solving two nested EGO algorithms [14]. The authors have proposed a method of the second kind which, while efficiently converging, requires the evaluation of numerous metrics based on various optima. They induce a complex scheme of up to five nested optimization and is intractable using continuous global optimization. Therefore this scheme, as the ones based on empirical metrics, rely on finely sampling the input space reducing the problem to a discrete optimization. Such a discretization induces a limited error on small dimension problems taking advantage of the fast evaluation of the surrogate but scales very poorly in higher dimensions because of the memory that a dense sampling grid would require.

An alternative approach is proposed in this work where EGO is implemented on the outer optimization loop and the statistical information required for the active calibration are estimated in the inner loop using extreme value distributions. This approach relies on strong simplification hypothesis (see section 3) and provides biased statistical information. However, this metrics is only used to guide the training of the surrogate by self-comparison for multiple designs. This approach allows us to implement EGO in a very straightforward way in the outer loop without requiring any sampling. However, in the current state of the method, the inner-loop estimator

is still evaluated by sampling on the associated sub-space. The space covered by sampling is already reduced to the uncertain dimensions thus alleviating the dimensionality issues, but the goal is to ultimately use a standard global optimization algorithm and completely avoid fine sampling.

In this paper the formalization of the problem is first given in Section 2 with both the robust design optimization (RDO) and adaptive Kriging formalism. The proposed learning function is detailed in Section 3 followed by the description of the associated algorithm in section 4. The convergence and overall performances of the method on two analytical function are discussed and compared to a method based on confidence intervals in section 5. Conclusions and complementary discussed are given in Section 6

2 FORMULATION OF THE PROBLEM

2.1 Robust design optimization

This work is focused on robust optimization in cases of interval uncertainty corresponding for example to early design stages when only expert's judgement information is available on the uncertain parameters. This type of problem consider a large interval on input variation in the evaluation of the robustness. In addition the goal is to develop a method widely applicable so no assumptions are taken with respect to the behaviour of the system.

Consider a system of $d \in \mathbb{N}$ dimensions such that the variables affecting it can be sorted into two kinds, $D \in \mathbb{R}^{d_D}$ are the design parameters that are to be optimized; the design space is considered bounded by the interval I_D. $U \in \mathbb{R}^{d_U}$ are the uncertain parameters that can take any value within the associated interval I_U. Consider an accurate physical model M of the system such that the scalar performance of interest y can be numerically estimated as $y = M(D, U)$ for any $(D, U) \in (I_D, I_U)$.

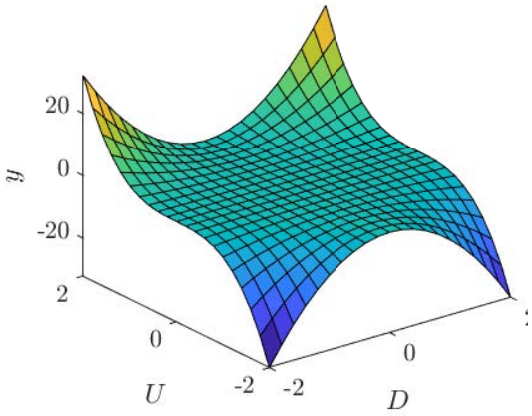
The robustness metric Δy of a design D is defined as the amplitude of performance over the interval of uncertain parameters I_U. The optimal design D^* is the one minimizing the amplitude of variation and is simply defined as:

$$\begin{aligned} D^* &= \underset{D \in I_D}{\operatorname{argmin}} \left(\Delta y(D) \right), \\ D^* &= \underset{D \in I_D}{\operatorname{argmin}} \left(\max_{U \in I_u} \left(M\left(D, U\right) \right) - \min_{U \in I_u} \left(M\left(D, U\right) \right) \right). \end{aligned} \tag{1}$$

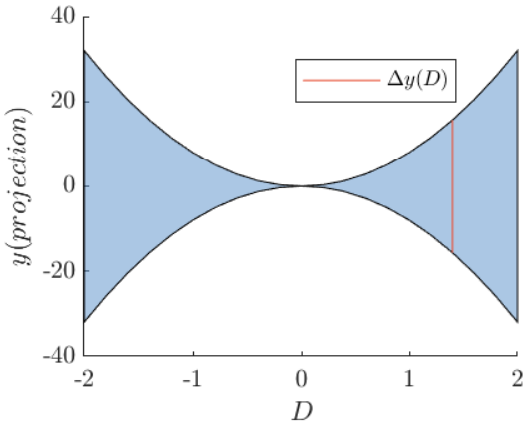
The general RDO problem and its main quantities are illustrated on a simple dummy problem in figure 1. The problem is intrinsically nested and with the considered definition of robustness it requires solving two optimization problems in the inner loop. The resolution using global optimization algorithm is usually intractable when associated with numerical models simulating the physics of the system. In this work this limitation is lifted using the surrogate modelling approach and more precisely the well known EGO [11] algorithm.

2.2 Adaptive Kriging approach

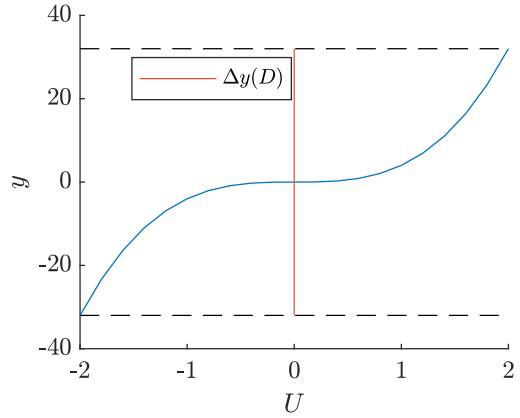
The principle of the surrogate modelling techniques is to calibrate an emulator $\hat{M}$ of the model M, which will mimic it's behaviour while being extremely fast to evaluate. Many mathematical object have been successfully used as surrogates e.g. support vector regression[15, 16], polynomial-chaos expansion [17] or response surface modelling [18]. The surrogate in this work is based on a Gaussian process, refereed to as Kriging, and its calibration can be seen as a Bayesian updating. For more details on Kriging models and their calibration the readers can



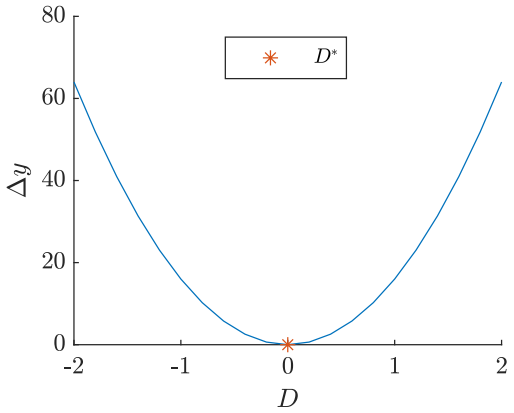
(a) Dummy system behavior



(b) Projection of performances on the design axis



(c) Performances as a function of U for an arbitrary design



(d) Robustness metric Δy as a function of the design

Figure 1: Illustration of a dummy robust design optimization problem and its solution

referred to [19]. Due to the Gaussian nature of the prediction, the estimated performance for a given set of inputs is a Gaussian random variable with parameters (μ, σ) dependent on the inputs:

$$\hat{M}(D,U) \sim N(\mu(D,U), \sigma(D,U)) \tag{2}$$

For the sake of clarity, after being first properly introduced once, the inputs dependency of random variables and related parameters or quantities will be noted with subscripts for the rest of the paper, e.g., $\hat{M}_{D,U}$. The popularity of Kriging surrogates can be partially attributed to the availability of the posterior standard deviation σ which can be seen as an estimation of the localized prediction uncertainty. This information have been exploited to develop so-called active-learning schemes where the design of experiment used for calibration is built iteratively in a near optimal way. The principle comes from acknowledging that the Kriging prediction uncertainties do not have homogeneous consequences on the uncertainty of the quantity of interest. Some areas of the input space then have a greater influence and require a finer requirement and vise versa. This idea has been developed into many adaptive calibration algorithms dedicated to specific problems with a strong popularity among uncertainty quantification problems [20, 21, 22].

In practice the iterative calibration is driven by a so-called learning function which aims at estimating the individual contribution of any input sample to the uncertainty of the goal quantity. The candidate associated with the highest learning function value is then evaluated using the physical model and added to the calibration data. Such an algorithm provides an elegant and comprehensible compromise between exploration and exploitation and have been proven quite successfully at solving a wide variety of problems [22]. A flowchart of the general adaptive Kriging scheme is provided in Figure 2 for illustration proposes.

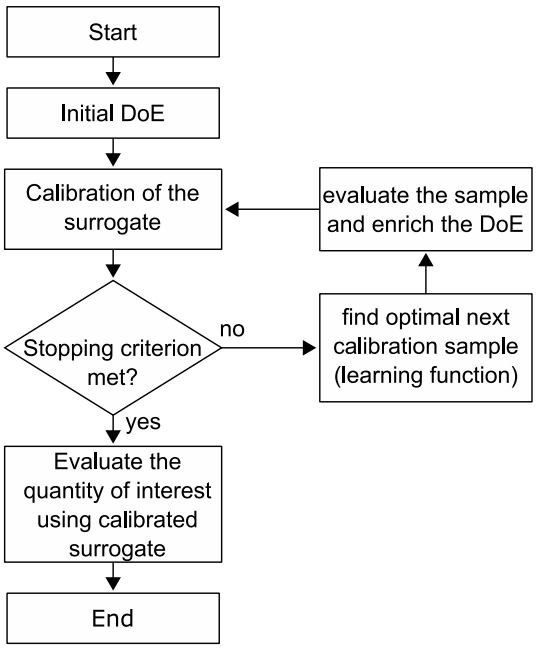


Figure 2: Flowchart of the general adaptive Kriging scheme.

In the context of this work, adaptive Kriging is used to solve a global optimization problem as described in Equation 1, more precisely the outer loop. EGO is usually considered the first adaptive Kriging algorithm and is dedicated to such a task. The associated learning function is called EI for expected improvement and is related to the expected decrease compared to the current optimal value Δy_{min}. I order to explicit the expression of the expected improvement in this context, let us first explicit the quantity of interest $\Delta y(D)$. The problem is to find the optimal robust design as described in section 2.1, as such the quantity of interest is, for any design D, the difference between the maximum $\overline{y}(D)$ and minimum $\underline{y}(D)$ values over the interval of uncertain variables I_u:

$$\begin{aligned} \overline{y}(D) &= \max_{U \in I_u} \left(M_{D,U} \right), \\ \underline{y}(D) &= \min_{U \in I_u} \left(M_{D,U} \right), \\ \Delta y(D) &= \overline{y}(D) - \underline{y}(D). \end{aligned} \tag{3}$$

Let us then consider that a Kriging surrogate $\hat{M}$ is available, and let us call $\hat{\overline{y}}$ and $\hat{\underline{y}}$ the Kriging counterparts of $\overline{y}$ and $\underline{y}$:

$$\begin{aligned}\hat{\overline{y}}(D) &= \max_{U \in I_u} \left(\hat{M}_{D,U} \right), \\ \hat{\underline{y}}(D) &= \min_{U \in I_u} \left(\hat{M}_{D,U} \right), \\ \hat{\Delta} y(D) &= \hat{\overline{y}}(D) - \hat{\underline{y}}(D).\end{aligned} \tag{4}$$

The three quantities described above are random variables since the surrogate $\hat{M}$ is a Gaussian process. $\hat{\overline{y}}(D)$, respectively $\hat{\underline{y}}(D)$ are maximum (resp. minimum) values of a stationary Gaussian process. The expected improvement can then be described directly as:

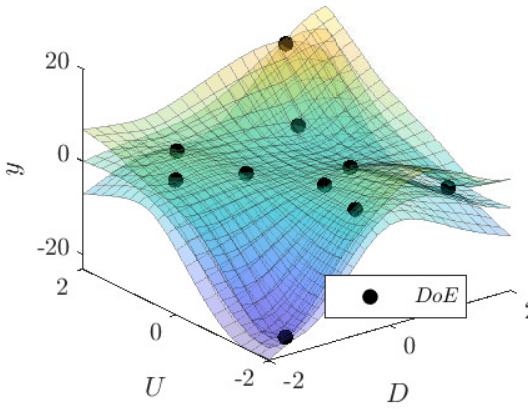
$$\begin{aligned}EI(\hat{M}_{D,U}, y_{min}) &= E\left[max\left(\Delta y_{min} - \hat{\Delta} y(D), 0 \right) \right] \\ EI(\hat{M}_{D,U}, y_{min}) &= F_{\hat{\Delta}_y}\left(\hat{\Delta} y_{min} \right) E\left(\Delta y_{min} - \hat{\Delta} y \middle| \hat{\Delta} y \leq \Delta y_{min} \right)\end{aligned} \tag{5}$$

with $F_{\hat{\Delta}_y}$ the cumulative distribution function of $\hat{\Delta} y$. Considering that if $\hat{\Delta} y$ takes a value higher than Δy_{min} the improvement is null, and in the other case the improvement is equal to the decrease in Δy_{min} value. The expected improvement is then related the expected value of $\hat{\Delta} y$ truncated above Δy_{min} as described in equation 5. In the case of Gaussian variables as it is the case in the standard EGO algorithm, this expression has a closed form allowing for a very fast estimation. However, in the problem considered in this work the distribution of $\hat{\Delta} y$ is unknown.

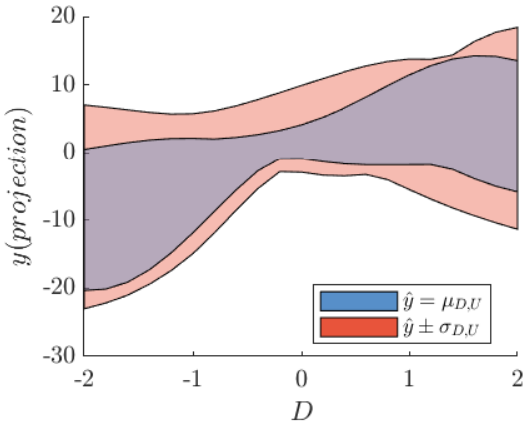
The extrema of Gaussian processes are very difficult to estimate, close form expressions of their distribution are only available for particular cases of covariance functions and against assumptions of the process shape. [23]. Of course the distributions could be estimated empirically by sampling the Gaussian process but such an approach would induce a discretization error, a sampling error and would scale very poorly with increasing dimension. The RDO problem associated with a Kriging surrogate is illustrated in figure 3 on the same example as in figure 1 and considering a Kriging model roughly calibrated with ten DoE points. All the quantities to be represented are then random variables and are illustrated with their expected values and standard deviation. The statistical moments of the robustness metric Δ_y ahve been estimated empirically by sampling the Gaussian process.

3 LEARNING FUNCTION

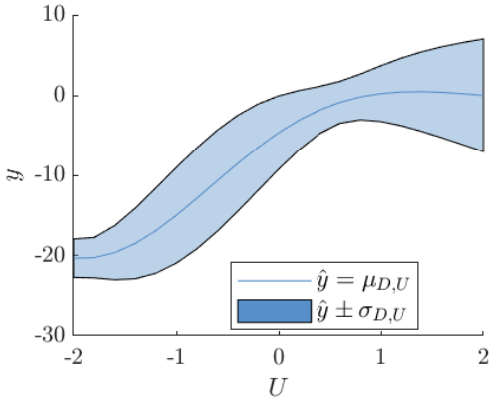
The proposed approach is to use the well known extreme value distributions associated with Gaussian random variables [24] as a rough indicator of the extrema distribution of the Gaussian process. For any set of input parameters, consider a population of $n_e \in \mathbb{N}$ samples is drawn from the associated Gaussian surrogate prediction $\hat{M}_{D,U}$. The maximum $\hat{\overline{M}}_{D,U}$ and minimum $\hat{\underline{M}}_{D,U}$ of this population follow known Gumbel distributions as:



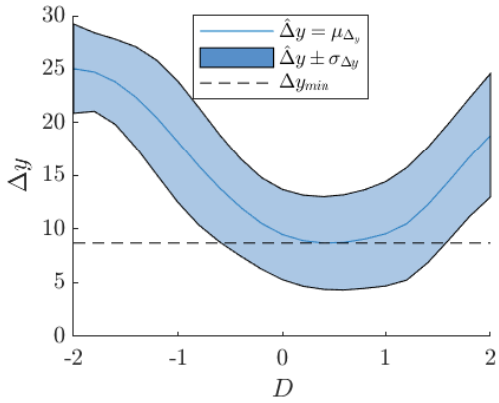
(a) Kriging prediction surface $\mu_{D,U}$ and the confidence interval surfaces $\mu_{D,U} \pm \sigma_{D,U}$



(b) Projection of Kriging estimated performances on the design axis



(c) Kriging estimated performances as a function of U for an arbitrary design



(d) Stochastic robustness metric $\hat{\Delta}y$ and its confidence interval as a function of the design

Figure 3: Illustration of a dummy robust design optimization problem associated with a Kriging surrogate of the system

$$
\begin{aligned}
&\overline{\hat{M}}\,(D,U,n_e) \sim \text{Gumbel}(\overline{\nu},\overline{\beta}) \\
&\overline{\nu}\,(D,U,n_e) = \Phi^{-1}\left(1-\frac{1}{n_e}\right)\sigma_{D,U} + \mu_{D,U} \\
&\overline{\beta}\,(D,U,n_e) = \Phi^{-1}\left(1-\frac{1}{n_e e^1}\right) - \overline{\nu}
\end{aligned}
\tag{6}
$$

$$
\begin{aligned}
&\underline{\hat{M}}\,(D,U,n_e) \sim Gumbel(\underline{\nu},\underline{\beta}) \\
&\underline{\nu}\,(D,U,n_e) = \Phi^{-1}\left(\frac{1}{n_e}\right)\sigma_{D,U} + \mu_{D,U} \\
&\underline{\beta}\,(D,U,n_e) = -\Phi^{-1}\left(\frac{1}{n_e e^1}\right) + \overline{\nu}
\end{aligned}
$$

with ν and β respectively the location and scale parameters of the Gumbel distributions, Φ^{-1} the inverse of the standard Gaussian cumulative distribution function, and μ, σ respectively

the mean and standard deviations of the Gaussian prediction. The main simplification of the approach is then, for a given design, to use the extreme value distribution associated with the highest expected value (resp. lowest) as an indicative distribution of the the maximum (resp. minimum) of the Gaussian process. This approach can be seen as reducing the maximum of the global stochastic process to the maximum of the local Gaussian variable that is the most likely to be responsible for the global maximum and vise versa. This is a very strong simplification assumption but it represents a convenient case to develop this first approach. More sophisticated approaches are possible and are planned to be investigated, for example considering multiple local maxima and their combined mixture distribution or including the local influence of covariance. The simplification can be expressed as:

$$\begin{aligned}\hat{\overline{y}}(D, n_e) &= \max_{U \in I_u} \left(\hat{M}_{D,U} \right) \approx \overline{\hat{M}}_{D,\overline{U}(D),n_e} \\ \hat{\underline{y}}(D, n_e) &= \min_{U \in I_u} \left(\hat{M}_{D,U} \right) \approx \underline{\hat{M}}_{D,\underline{U}(D),n_e}\end{aligned} \tag{7}$$

with $\overline{U}(D)$, respectively $\underline{U}(D)$, the vector of uncertainty values associated with the upper (resp. lower) extreme value distribution having the highest (resp. lowest) expected value:

$$\begin{aligned}\overline{U}(D, n_e) &= \underset{U \in I_U}{\operatorname{argmax}} \left(E \left(\overline{\hat{M}}_{D,U,n_e} \right) \right) \\ \underline{U}(D, n_e) &= \underset{U \in I_U}{\operatorname{argmin}} \left(E \left(\underline{\hat{M}}_{D,U,n_e} \right) \right)\end{aligned} \tag{8}$$

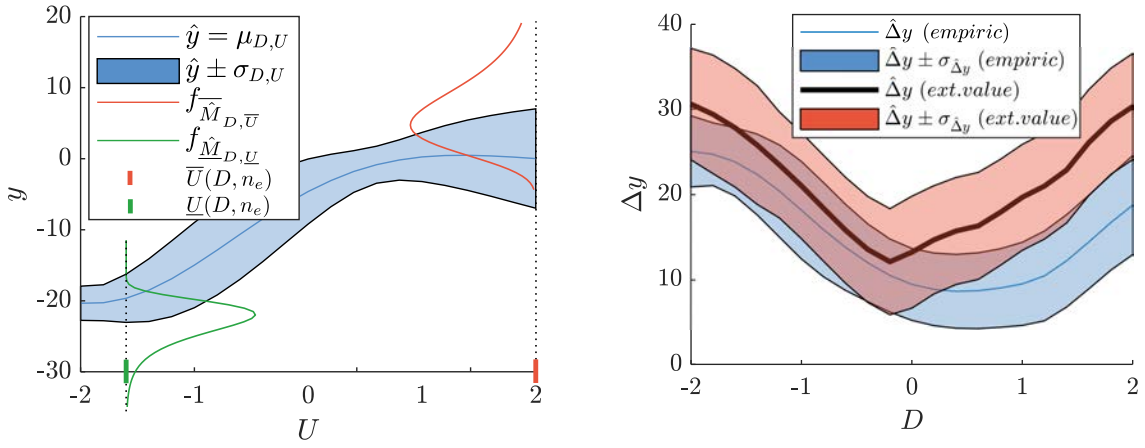
These two coordinates as well as their corresponding extreme value distributions are illustrated in Figure 4a. This figure is drawn on the same case and same Kriging model as Figure 3. Following this simplification, the unknown random variables defined in equation 4 are now simple Gumbel distributions. Their parameters can be obtained by solving the two optimization problems in the sub-space of uncertain parameters described in equation 8. The distribution of the robustness metric $\hat{\Delta}y(D, n_e)$ can finally be directly obtained by simple convolution of the two Gumbel distributions:

$$\begin{aligned}\hat{\Delta}y\,(D, n_e) &\approx \overline{\hat{M}}_{D,\overline{U}_D,n_e} - \underline{\hat{M}}_{D,\underline{U}_D,n_e} \\ f_{\hat{\Delta}y}(y) &= \int_{\tau=-\infty}^{+\infty} f_{\overline{\hat{M}}_{D,\overline{U}_D,n_e}}(\tau) f_{\underline{\hat{M}}_{D,\underline{U}_D,n_e}}(\tau - y) d\tau\end{aligned} \tag{9}$$

with $f_{\hat{\Delta}y}$, $f_{\overline{\hat{M}}}$ and $f_{\underline{\hat{M}}}$ respectively the approximated distribution of $\hat{\Delta}y_{D,n_e}$ and the two Gumbel distributions associated with $\overline{\hat{M}}_{D,\overline{U}(D)}$ and $\underline{\hat{M}}_{D,\underline{U}(D)}$ as described in Equation 6. The idea is also illustrated in Figure 4b where the expected value of $\hat{\Delta}y$ and the associated $\pm\sigma$ confidence interval is drawn both empirically from sampling the Gaussian process and considering the difference of the two extreme Gumbel distributions as detailed in Equation 9. The expected improvement metric described in Equation 5 can finally be estimated by numerical integration (after simplification of the constants) as:

$$EI \left(\hat{M}_{D,n_e,y_{min}} \right) \approx \int_{\tau=-\infty}^{y_{min}} (\Delta y_{min} - \tau)\, f_{\hat{\Delta}y}(\tau) d\tau \tag{10}$$

with Δy_{min} being the current best estimation of the minimal robustness metric in the design space.



(a) Kriging estimated performances for an arbitrary design as a function of U and associated extreme value distributions

(b) Projection of empirical stochastic performances on the design axis, comparison with the simplified difference of extreme value distributions

Figure 4: Illustration of the extreme value distributions and associated stochastic robustness metric

4 ALGORITHM

The goal of this section is to describe the robust optimization algorithm developed on the performance metric described in the previous section. Details of the implementation such as the initial DoE, the choice of the n_e hyper-parameter or the global optimization algorithm chosen for EGO are briefly discussed at the end of the section.

Consider the state of the algorithm at an arbitrary step $k \in \mathbb{N}$ (neither initial or final). A current best estimate of the minimal robustness metric y_{min} is available, as well as a design of experiment in the shape of a $k + k_{ini}$ by d vector $X = [D^i, U^i], i \in (1, .., k + k_{ini})$ with k_{ini} being the size of the initial design of experiment. The goal of the refinement step is to find the sample X^*_{EI} providing the maximum expected improvement regarding the estimation of D^* as defined in equation 10. This sample will then be evaluated using the numerical model M and added to the DoE for next step.

The search for the optimal coordinates is split in two parts as, form the point of view of the adaptive-Kriging scheme, the RDO problem is only related to the design sub-space D. The optimal design coordinates D^*_{EI} of X^*_{EI} are identified first using a standard global optimization algorithm as in the EGO algorithm and as formulated in equation 14. Near-optimal uncertain coordinates U^*_{EI} are then obtained as a by-product of the identification of D^*_{EI}.

$$D^*_{EI} = \underset{D \in I_D}{\operatorname{argmax}} \left(EI \left(\hat{M}_{D,n_e,y_{min}} \right) \right) \tag{11}$$

The evaluation of the expected improvement driving the optimization is done using numerical integration as described in Equation 10. This evaluation requires first the identification of the two samples $\overline{U}(D, n_e)$ and $\underline{U}(D, n_e)$ associated with the two "worst-case" extreme value

distributions in the sense of their expected values as defined in Equation 8. As introduced in Section 3 they are currently unidentified in a discrete way by sampling the uncertainty subspace $U \in I_U$. As this sampling-based approach is used as a temporary first approach no strong efforts have been put in defining a sophisticated sampling strategy. It is simply purposed to any space-filling design such as Latin-Hypercube sampling to generate a population vector of dimension N_s by d_U. On the analytical examples tested in Section 5 N_s has been chosen equal to $d_U * 10^3$.

Once the optimal D^*_{EI} coordinates have been found, it is proposed to only consider the coordinates of the worst-case Gumbel distributions previously identified $\overline{U}(D^*, n_e)$ and $\underline{U}(D^*, n_e)$ as candidates for the U coordinates of X^*_{EI}. Doing so avoids solving additional optimization problems while providing near-optimal coordinates. In order to detail the reasoning let us first clarify what is the contribution of U^*_{EI} in the overall performances of the algorithm. As a comparison the goal of selecting D^*_{EI} is to improve the accuracy of surrogate around possible robustness optima. For a given D^*_{EI} the goal of selecting U^*_{EI} is then to improve the accuracy of the surrogate at estimating the corresponding expected improvement. Since the estimation of the expected improvement only rely on $\overline{U}(D, n_e)$ and $\underline{U}(D, n_e)$ they are naturally preferred candidates. U^*_{EI} is finally chosen as the candidate associated with the highest prediction uncertainty, i.e. the Gumbel distribution with the highest variance:

$$U^*_{EI} = \underset{U \in \{\overline{U}, \underline{U}\}}{\operatorname{argmax}} \left(var \left(\overline{\hat{M}}_{D^*, \overline{U}, n_e}, \underline{\hat{M}}_{D^*, \underline{U}, n_e} \right) \right) \tag{12}$$

Other implementation details are provided bellow:

- The sensitivity of the performances to the extent of the initial DoE k_{ini} has not been thoroughly investigated. Scarce initial designs of experience have the potential of providing better performances as these samples are the only one not guided by the learning function. However this approach also relies very strongly on the ability of the learning function ability to efficiently explore the input space. For the application examples presented in Section 5 a value of $k_{ini} = 10$ has been chosen and samples are generated using the latin-hypercube sampling scheme.

- The global optimization of the expected improvement is performed using the interior-point algorithm. The algorithm was choice for its wide range of applicability but no significant performance difference was observed when compared to the sequential quadratic programming or active set algorithms.

- At the current state of development no satisfying stopping criterion has been found. The one proposed with the EGO algorithm based on values of the expected improvement has been tested and found to be very prone to premature termination. This issue is probably related to the expected improvement values being quite noisy between iterations due to the extra layer of complexity of the RDO problem. For the applications presented in Section 5.2 the method have been ran for as many iterations as the reference method required to converge.

- The influence of the n_e hyper-parameter on the performances of the method has not been thoroughly studied yet. From the currently limited empirical evidence available to the authors increasing n_e have a tendency to increase the mean of the robustness metrics and

decrease its variance. Smaller values of n_e have been found to be more suited to the convergence of the algorithm and for the applications illustrated in section 5.2 a values of $n_e = 4$ have been chosen.

- Ideally the value of Δy_{min} used in the definition of the expected improvement should be updated at every iteration of the outer loop of the proposed algorithm. However doing so would require the resolution of an additional nested optimization problem. Even though it would be based only on the surrogate this extra step would induce significant computation burden. Since the accuracy of the Δy_{min} value is not of critical importance for the convergence of the algorithm, it is proposed to formally evaluate it only once at the initialization stage. Then, at every iteration, after having identified D^*_{EI} the performance amplitude associated with the two samples $\overline{U}(D^*, n_e)$ and $\underline{U}(D^*, n_e)$ is evaluated, compared with Δy_{min} and the smaller of the two is kept.

Finally, once the calibration loop is finished, the nested RDO problem as defined in equation 1 is solved using the same general nested scheme. The inner-loop quantity of interest is the amplitude of variation $\Delta y(D)$ for a given design. This quantity is estimated in a discrete manner using the same sampling scheme as for U^*_{EI}. The outer loop problem of finding the design responsible for the minimal value of Δy is solved using the same global optimization algorithm as the one used to find D^*_{EI}. No extra evaluation of the numerical model are required in this step. The overall algorithm is illustrated by a flowchart in figure 5.

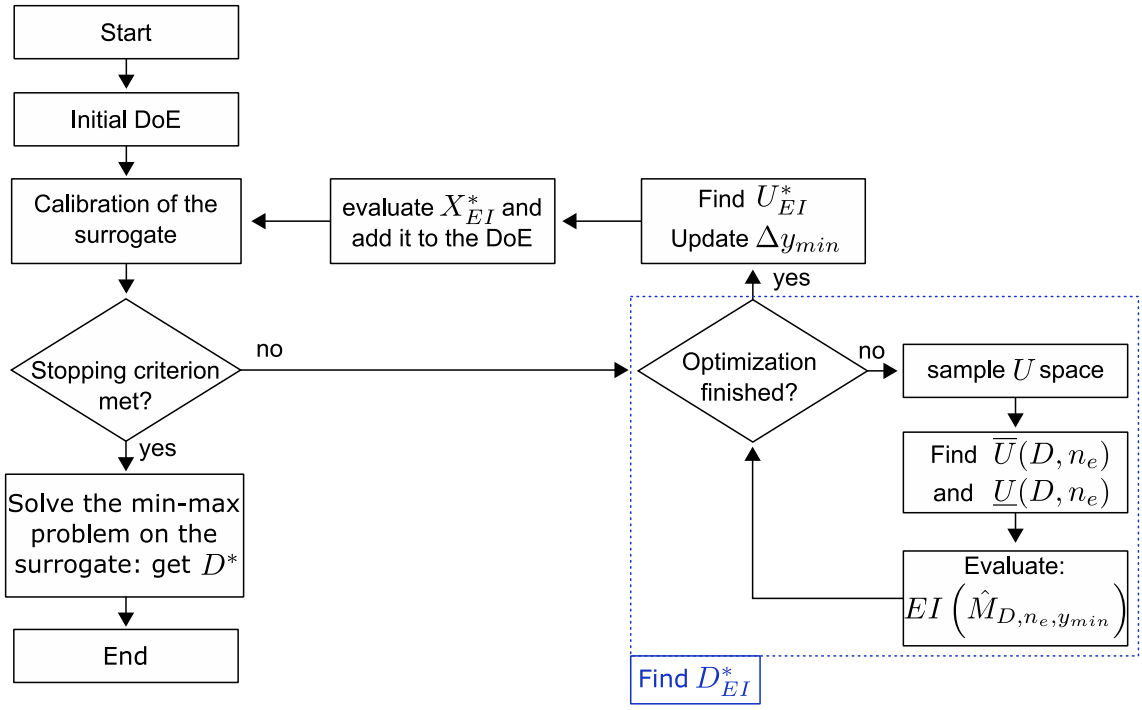


Figure 5: Flowchart of the proposed robust optimization algorithm.

5 BENCHMARK

The performances of the proposed method is illustrated on two analytical problems of varying complexity. The functions have been chosen from the benchmark of the previously pub-

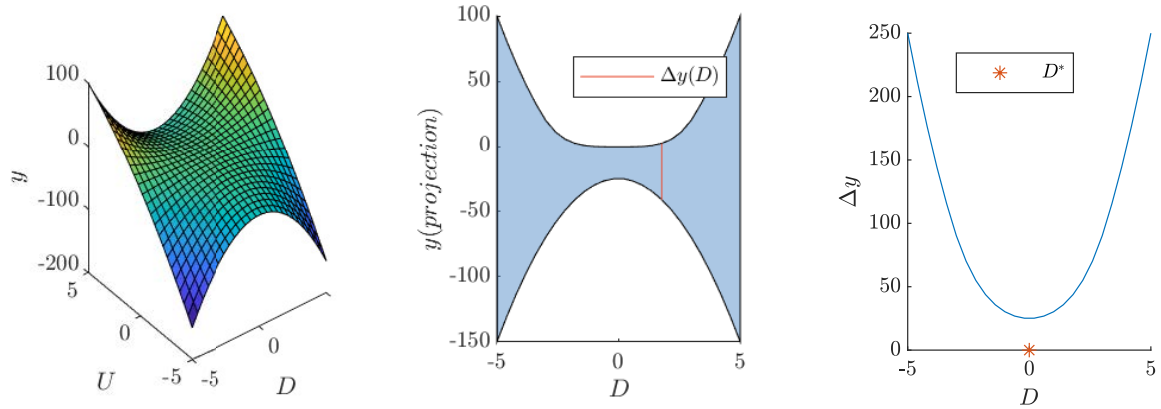
lished RULOK method [13] from the same authors which uses a learning function based on confidence intervals. Since the proposed method currently lacks a stopping criterion the number of required iterations is only studied though a convergence figure.

5.1 Example 1

The first example is a simple polynomial function in dimension two with a unique global optimal design $D^* = 0$ and no local ones. The design space I_D and interval uncertainty I_U are both equal to $[-5, 5]$. The performance function reads as follows:

$$M(D, U) = U^2 D - D^2 \tag{13}$$

The problem is further illustrated in figure 6 with the system's behavior and optimal design point. in the same manner as the RDO problem was introduced in Figure 2.1 .



(a) Performance surface as a function ofd D and U

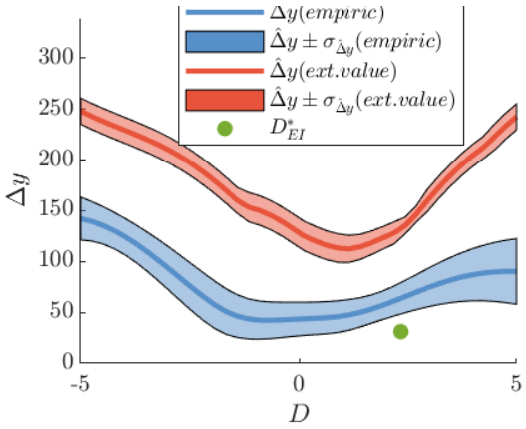
(b) Projected performances as a function of D

(c) Robustness metric as a function of D and optimal design

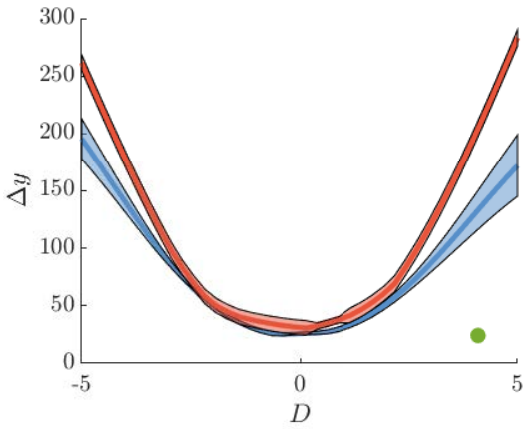
Figure 6: Illustration of the first benchmark problem behaviour and solution

The performances of the method are illustrated in figure 7 where figure a,b and c represent the state of the problem after iterations number 1, 6 and 13 respectively. They consist in a comparison of the empirical surrogate estimation of the robustness metric Δy with the related metric approximated as the difference of the two "worst-case" extreme value distributions. It can be seen that the bias introduced by the approximations of the method is quite significant initially but is quickly reduced as the calibration of the surrogate progresses. It can also be seen that while the bias is significant at the beginning, the general trend is captured, allowing for an efficient convergence of the algorithm.

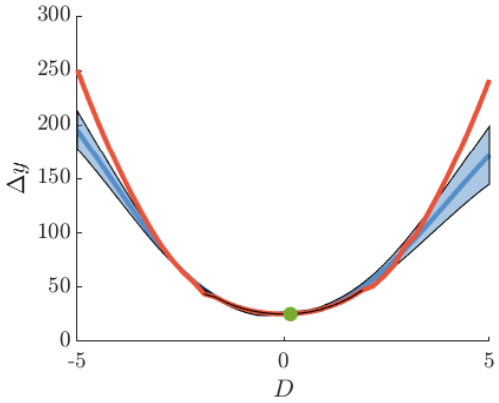
Figure 7d illustrates the convergence of the algorithm over fifteen iterations. It can been seen that the correct solution was identified after iteration ten, corresponding to twenty total evaluations of the model. As a reference the AK-RULOK method required between 38 and 44 evaluations of the model to converge. The performances on this example are very promising with a potential reduction of the computation cost by a factor of two, provided a satisfying stopping criterion can be developed.



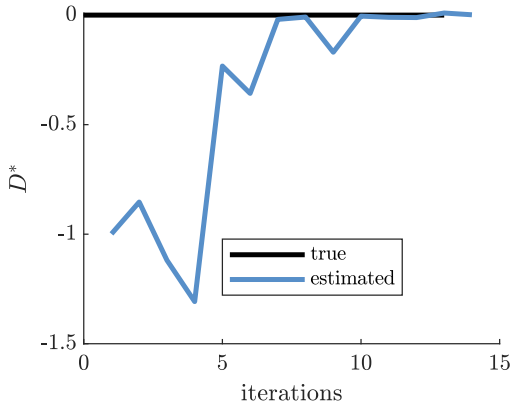
(a) Empiric and simplified robustness metric Δy as a function of D, average and $\pm\sigma$ confidence interval, iteration 1



(b) Empiric and simplified robustness metric Δy as a function of D, average and $\pm\sigma$ confidence interval, iteration 6



(c) Empiric and simplified robustness metric Δy as a function of D, average and $\pm\sigma$ confidence interval, iteration 13



(d) Convergence plot of the estimated location of the optimal design D^*

Figure 7: Results for the first benchmark problem

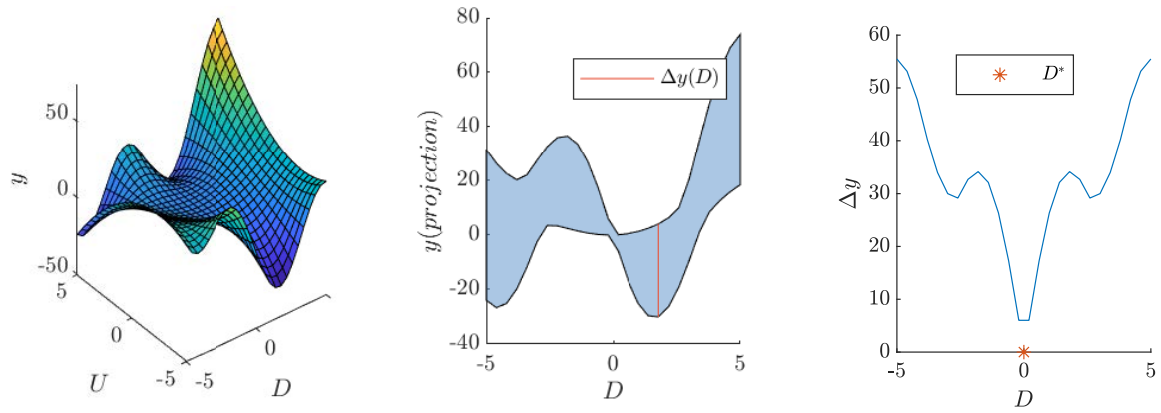
5.2 Example 2

The second example is a more challenging function in dimension two with a unique global optimal design $D^* = 0$ and two local ones. The design space I_D and interval uncertainty I_U are both equal to $[-5, 5]$. The performance function reads as follows:

$$M(D, U) = U.D - sin(D)U^2 + D^2 \tag{14}$$

The problem is further illustrated in figure 8 with the system's behavior and optimal design point.

The performances of the method on the second example are illustrated in figure 9 where figure a,b and c represent the state of the problem after iterations number 1, 14 and 49 respectively in the same fashion as for example one. The conclusions with respect to the accuracy of the approximated robustness metric are similar as for the first example: the bias is initial significant but the trend is capture allowing an efficient calibration of the surrogate. While the calibration progresses, the bias is quickly reduced and the used metric end up being a very accurate



(a) Performance surface as a function ofd D and U

(b) Projected performances as a function of D

(c) Robustness metric as a function of D and optimal design

Figure 8: Illustration of the second benchmark problem behaviour and solution

representation of the difference of the surrogate extrema.

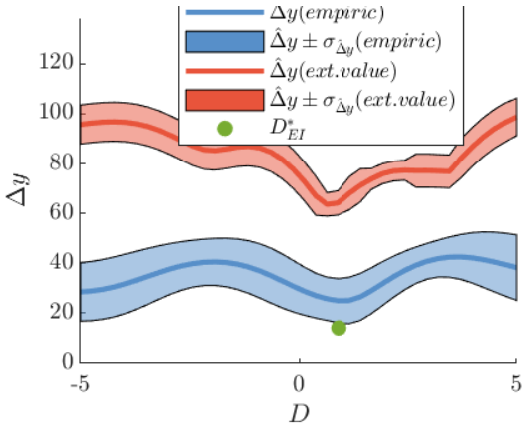
Figure 7d illustrates the convergence of the algorithm over 49 iterations. It can been seen that the correct solution was approximately identified after iteration 11 and exactly identified after iteration 28, corresponding to 38 total evaluations of the model. As a reference the AK-RULOK method required between 30 and 51 evaluations of the model to converge (depending on the confidence desired as stopping criterion). The performances are also very promising of this example with a significant potential to provide reduced computation time compared to the AK-RULOK method.

6 conclusions

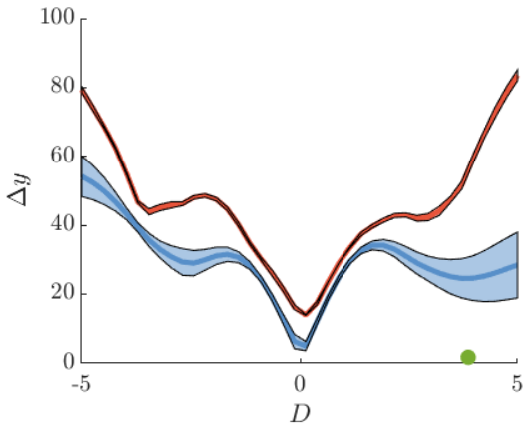
This paper presents a new learning function for solving robust design optimisation problems using the adaptive Kriging framework. The learning function is based on a simplification of the problem of estimating the extrema of a Gaussian process. The proposed simplified distribution is obtained as the difference of the extreme value Gumbel distributions correspond to the ”worst-case” values of uncertain variables for any given design. This metric allows us to solve the RDO problem using a very straightforward implementation of classical adaptive Kriging methods. The proposed method also reduces the necessity of sampling the input space compared to other methods from the literature which could improve of the performances evolve with problems of increasing dimension.

The performances of the method are illustrated on two analytical benchmark problems and compared to the results of a related method based on confidence intervals. The results are very promising with a very fast convergence toward the optimal design. The results also show that the approximation used to compute the robustness metric do not seem to prevent the algorithm to efficiently guide the calibration of the surrogate. In addition the biases due to the approximations are quickly reduce while the calibration progresses.

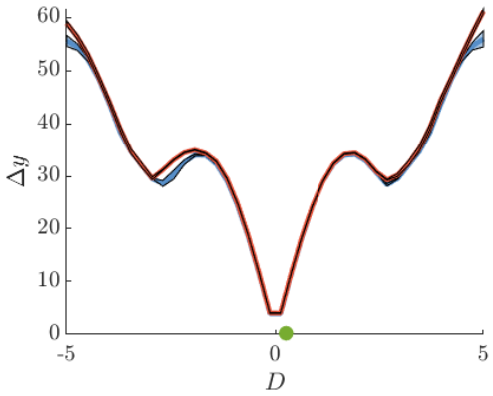
While the results are promising a number of aspects of the proposed method are still to be finished. An efficient stopping criterion is needed and the one proposed with the EGO algorithm was not satisfying as far as the authors tested. Other approaches are possible and will be investigated such as a criterion based on the convergence of the optimal design location. In addition the sensitivity of the algorithm to several hyper-parameters needs to be thoroughly studied such as the extent of the initial design of experiment of the artificial number of samples



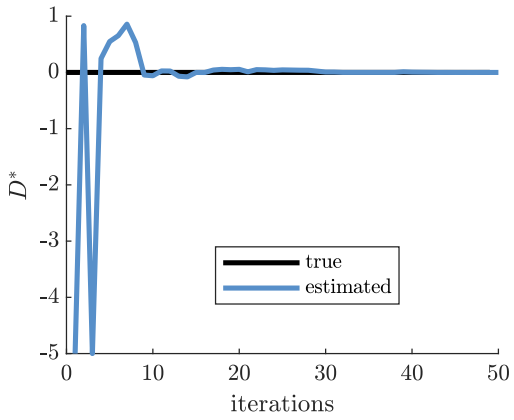
(a) Empiric and simplified robustness metric Δy as a function of D, average and $\pm\sigma$ confidence interval, iteration 1



(b) Empiric and simplified robustness metric Δy as a function of D, average and $\pm\sigma$ confidence interval, iteration 14



(c) Empiric and simplified robustness metric Δy as a function of D, average and $\pm\sigma$ confidence interval, iteration 49



(d) Convergence plot of the estimated location of the optimal design D^*

Figure 9: Results for the second benchmark problem

required in the definition of the extreme value distributions. Finally, the currently algorithm still relies on sampling the space of uncertain variables in the inner loop of the min-max problem. The ambition with this approach is to entirely avoid fine sampling for the input space to increase performances on higher dimension but, from early tests, simply implementing a two-level global optimization problem induces significant computation time and can result in inaccuracies in the selection of the optimal sample to add to the DoE. The source of these limitations is still unclear and needs to be investigated.

Acknowledgements

The authors gratefully acknowledge the support of the Research Foundation Flanders (FWO) under grant 1SA3919N (C. van Mierlo) and GOC2218N (A. Persoons). In addition, we acknowledge the European Union's Horizon 2020 Research and Innovation program GREYDIENT under grant agreement n° 955393.

REFERENCES

[1] G. Taguchi. Quality engineering (taguchi methods) for the development of electronic circuit technology. *IEEE Transactions on Reliability*, 44(2):225–229, 1995.

[2] G. Taguchi. Quality engineering (Taguchi methods) for the development of electronic circuit technology. *IEEE Transactions on Reliability*, 44(2):225–229, June 1995.

[3] Gyung-Jin Park, Tae Hee Lee, Kwon Hee Lee, and Kwang-Hyeon Hwang. Robust Design: An Overview. *AIAA Journal*, 44(1):181–191, January 2006. MAG ID: 2076184585.

[4] Hans-Georg Beyer and Bernhard Sendhoff. Robust optimization – A comprehensive survey. *Computer Methods in Applied Mechanics and Engineering*, 196(33-34):3190–3218, July 2007.

[5] C. Zang, M.I. Friswell, and J.E. Mottershead. A review of robust optimal design and its application in dynamics. *Computers & Structures*, 83(4-5):315–326, January 2005.

[6] Dimitris Bertsimas, David B. Brown, and Constantine Caramanis. Theory and Applications of Robust Optimization. *SIAM Review*, 53(3):464–501, January 2011.

[7] Berc Rustem and Melendres Howe. *Algorithms for worst-case design and applications to risk management*. Princeton University Press, Princeton, N.J. ; Oxford, 2002. OCLC: ocm48931090.

[8] Ruichen Jin, Xiaoping Du, Wei Chen, and Wei Chen. The use of metamodeling techniques for optimization under uncertainty. *Structural and Multidisciplinary Optimization*, 25(2):99–116, June 2003. MAG ID: 2047626293.

[9] Janis Janusevskis and Rodolphe Le Riche. Simultaneous kriging-based estimation and optimization of mean response. *Journal of Global Optimization*, 55(2):313–336, February 2013.

[10] Brian J Williams, Thomas J Santner, and William I Notz. Sequential design of computer experiments to minimize integrated response functions. *Statistica Sinica*, 10:1133–1152, 2000.

[11] Donald R Jones and Matthias Schonlau. Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global Optimization*, 13:455–492, 1998.

[12] Samee ur Rehman, Matthijs Langelaar, and Fred van Keulen. Efficient Kriging-based robust optimization of unconstrained problems. *Journal of Computational Science*, 5(6):872–881, November 2014.

[13] Conradus van Mierlo, Augustin Persoons, Matthias G. R. Faes, and David Moens. Robust design optimisation under lack-of-knowledge uncertainty. *Computers & Structures*, 275:106910, January 2023.

[14] Julien Marzat, Eric Walter, and Hélène Piet-Lahanier. A new expected-improvement algorithm for continuous minimax optimization. *Journal of Global Optimization*, 64(4):785–802, April 2016.

[15] J.-M. Bourinet, F. Deheeger, and M. Lemaire. Assessing small failure probabilities by combined subset simulation and Support Vector Machines. *Structural Safety*, 33(6):343–353, September 2011.

[16] J.-M. Bourinet. Rare-event probability estimation with adaptive support vector regression surrogates. *Reliability Engineering & System Safety*, 150:210–221, June 2016.

[17] Bruno Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety*, 93(7):964–979, July 2008.

[18] Youbao Jiang, Linjie Zhao, Michael Beer, Edoardo Patelli, Matteo Broggi, Jun Luo, Yihua He, and Jianren Zhang. Multiple response surfaces method with advanced classification of samples for structural failure function fitting. *Structural Safety*, 64:87–97, January 2017. Reporter: Structural Safety.

[19] C Lataniotis, D Wicaksono, S Marelli, and B Sudret. UQLab user manual – Kriging (Gaussian process modeling).

[20] B. Echard, N. Gayton, and M. Lemaire. AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation. *Structural Safety*, 33(2):145–154, March 2011. Number: 2 Reporter: Structural Safety.

[21] B. Echard, N. Gayton, M. Lemaire, and N. Relun. A combined Importance Sampling and Kriging reliability method for small failure probabilities with time-demanding numerical models. *Reliability Engineering & System Safety*, 111:232–240, March 2013. Reporter: Reliability Engineering & System Safety.

[22] M. Moustapha, S. Marelli, and B. Sudret. A generalized framework for active learning reliability: survey and benchmark. *arXiv:2106.01713 [stat]*, June 2021. arXiv: 2106.01713.

[23] Jean-Marc Azaïs and Li-Vang Lozada-Chang. A toolbox on the distribution of the maximum of Gaussian process.

[24] Samuel Kotz and Saralees Nadarajah. *Extreme value distributions: theory and applications*. Imperial College Press ; Distributed by World Scientific, London : River Edge, NJ, 2000.