



HAL
open science

Pareto optimal fidelity level selection for multi-fidelity Bayesian optimization applied to drone design

Remy Charayron, Thierry Lefebvre, Nathalie Bartoli, Joseph Morlier

► To cite this version:

Remy Charayron, Thierry Lefebvre, Nathalie Bartoli, Joseph Morlier. Pareto optimal fidelity level selection for multi-fidelity Bayesian optimization applied to drone design. 12 th EASN International conference on "Innovation in Aviation & Space for opening New Horizons", Oct 2022, Barcelone, Spain. hal-03919052

HAL Id: hal-03919052

<https://hal.science/hal-03919052>

Submitted on 2 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pareto optimal fidelity level selection for multi-fidelity Bayesian optimization applied to drone design

R. Charayron^{1,2}, T. Lefebvre¹, N. Bartoli¹ and J. Morlier^{2,3}

¹ ONERA/DTIS, Université de Toulouse, Toulouse, France

² ISAE-SUPAERO, Toulouse, France

³ ICA, Université de Toulouse, INSA, CNRS, MINES ALBI, UPS, Toulouse, France

E-mail: remy.charayron@onera.fr

E-mail: thierry.lefebvre@onera.fr

E-mail: nathalie.bartoli@onera.fr

E-mail: joseph.morlier@isae-supaero.fr

Abstract. Optimizing expensive models is a challenging task in the aeronautical design community. Multi-fidelity optimization proposes to use different levels of fidelity to alleviate the overall computational cost of the optimization. How to know which level of fidelity will be useful for the optimization process ? In this work, we propose a methodology to decide which levels of fidelity are useful. We construct a Pareto front for the fidelity levels in terms of cost and accuracy, and then select only the Pareto optimal fidelity levels. All the models considered in this work were considered as black-boxes, so a gradient-free optimizer is used to solve the optimization problems. We chose to use a multi-fidelity Bayesian algorithm in order to reduce the number of function evaluations. We showed the advantage of selecting the optimal Pareto fidelity levels in the multi-fidelity process by comparing the cost to achieve convergence using different scenarios for level selection. To this end, we applied our methodology to several analytical test cases and to a simple drone design problem for which we obtained an improvement of 38 percent improvement over the mono-fidelity algorithm and a 24 percent improvement over the multi-fidelity algorithm with random selection.

1. Introduction

Multi-fidelity Bayesian optimization proved to overtake mono-fidelity Bayesian optimization when solving problems where the involved models are very expensive [29, 3, 25]. Indeed, multi-fidelity Bayesian optimizer can query low-fidelity models in order to explore the design space but also high expensive fidelity models when exploitation of the surrogate is needed. Sometimes a lot of fidelity levels are available to approach the quantities of interest. In this paper, we address the problem of choosing how many and which levels of fidelity to use in the multi-fidelity process. We propose an efficient approach to select the right combination of fidelity levels among all the possible combinations. To make this selection, we first evaluate each fidelity level according to two objectives which are their cost and accuracy, and then keep only the non-dominant fidelity levels. The paper is divided as follows. Section 2 gives a state of the art on the multi-fidelity Kriging and multi-fidelity Bayesian optimization. Next, Section 3 explains the proposed Pareto optimal fidelity level (POFL) methodology. Some analytical test cases and the associated results are presented in Section 4. The methodology is then experienced on a drone design test case in Section 5. Finally, some conclusions and perspectives are given in Section 6.

2. State of the art

This section presents the multi-fidelity framework that is used in this paper. We want to solve an optimization problem with inequality or/and equality constraints:

$$\begin{aligned} & \min_{x \in \Omega} f(x) \\ \text{such that } & \begin{cases} g(x) \leq 0 \\ h(x) = 0 \end{cases} \end{aligned} \quad (1)$$

where the objective function f and the constraints g, h are evaluated with a Multidisciplinary Design Analysis (MDA) costly to evaluate. In a multi-fidelity context, the f, g and h can be obtained by different MDAs associated to different fidelity levels. To solve this multidisciplinary optimization (MDO) problem with expensive black-box functions, Bayesian optimizers [30, 31] are used to minimize the number of function evaluations. First Section 2.1 shows how to construct a multi-fidelity Gaussian process surrogate model. Then, in Section 2.2, the current Bayesian optimization approach is adapted to multi-fidelity by defining a criterion to choose the fidelity level to query. We assume here that the reader is familiar both with Gaussian processes (GP) interpolation also denoted Kriging models [19, 10, 20] and with the classical mono-fidelity Kriging based Bayesian optimization methods like the Efficient Global Optimization (EGO) [5] algorithm for unconstrained problems and the Super Efficient Global Optimization (SEGO) algorithm [9, 6, 7] for constrained problems.

2.1. Multi-fidelity Kriging

The first extension of the Kriging model for multivariate functions, called co-Kriging, was first developed in geostatistics (see [26, 27]). Next, [1] figured that making assumptions to relate different levels of fidelity is a way to simplify multi-fidelity problems. In this section, focus will be made on Le Gratiet's [2] recursive formulation that we used to construct multi-fidelity Kriging models. Let's suppose that $L + 1$ fidelity levels f_0, \dots, f_L sorted from the lowest to the highest are available to approach the function f (the same philosophy could be applied to the constraint functions) and that we dispose of $L + 1$ design of experiments (DoE), one for each fidelity level, denoted $(D_i)_{i=0 \dots L}$ with $D_i = ((x_0, f_i(x_0)), \dots, (x_j, f_i(x_j)), \dots, (x_{N_i}, f_i(x_{N_i})))^T$, N_i being the number of points in D_i . We make the nested DoE assumption which states that each point evaluated at a fidelity level is also evaluated at all the lowest fidelity levels: $D_L^x \subseteq D_{L-1}^x \subseteq \dots \subseteq D_0^x$ where $D_i^x = (x_0, \dots, x_j, \dots, x_{N_i})^T$ corresponds to the inputs part of D_i . The nested DOE provides some properties useful to express the surrogate model variance in closed form. Let's make the following assumption first introduced by [1] in the bi-fidelity case to link the fidelity levels

$$f_{l+1}(x) = \rho_l f_l(x) + \delta_{l+1}(x) \quad \text{such that} \quad f_l \perp \delta_{l+1} \quad \forall l = 0, \dots, L - 1 \quad (2)$$

where $\delta_{l+1}(x)$ is a discrepancy function that captures the difference between the $l + 1$ -th and the l -th fidelity levels while ρ_l is a scaling factor which is applied to f_l .

Le Gratiet [2] proposed to add the lowest fidelity function to the basis function set $(h_i)_{i=1 \dots p}$ used in the universal Kriging regression term to get:

$$\mu(x) = \sum_{i=1, \dots, p} \left(\beta_i h_i(x) \right) + \beta_{\rho_0} f_0(x) \quad (3)$$

where β_{ρ_0} is an estimation of ρ_0 and $(\beta_i)_{i=1 \dots p}$ is a set of unknown coefficients to be multiplied by the basis functions. Estimations of these coefficients are done at the hyperparameter estimation step performed by maximizing the likelihood [11, 18]. Since we used a nested DoE structure, the independence between fidelity levels of the surrogate model is assumed. Then the following recursive formulation for the mean ($\hat{\mu}_{l+1}$) and variance ($\hat{\sigma}_{l+1}^2$) of each fidelity level GP surrogate model can be written:

$$\forall l = 0, \dots, L - 1 \quad \begin{cases} \hat{\mu}_{l+1} = \rho_l \hat{\mu}_l + \hat{\mu}_{\delta_{l+1}} \\ \hat{\sigma}_{l+1}^2 = \rho_l^2 \hat{\sigma}_l^2 + \hat{\sigma}_{\delta_{l+1}}^2 \end{cases} \quad (4)$$

In this case, ρ_l is considered as a constant but it can depend on x and we have $\rho_l : x \mapsto \rho_l(x)$. The learning process is the following, first the lowest fidelity level is learnt, then the relationships (scaling factor ρ_l and discrepancy function δ_{l+1} with $l = 0, \dots, L - 1$) between every successive fidelity level are consecutively learnt. Since the variances can be expressed in closed form, the contribution of each fidelity level to the total variance of the multi-fidelity model can be expressed too. It constitutes the more remarkable advantage of Le Gratiet's recursive formulation. Denoting $\sigma_{cont}^2(l, x)$ the variance contribution of the l^{th} fidelity level at the point x , with the notation $\sigma_{\delta_0}^2 = \sigma_0^2$ we have:

$$\forall l = 0, \dots, L - 1 \quad \sigma_{cont}^2(l, x) = \sigma_{\delta_l}^2(x) \prod_{j=l}^{L-1} \rho_j^2 \quad \text{and} \quad \sigma_{cont}^2(L, x) = \sigma_{\delta_L}^2(x) \quad (5)$$

2.2. Multi-fidelity Bayesian optimization

With a multi-fidelity Bayesian optimization process, one must decide which point is most promising and the level of fidelity at which to evaluate it. It has been proposed in [2] to solve the problem of finding the point and level of fidelity in two successive steps. First the point is found using a classical acquisition function as in the mono-fidelity Bayesian optimization. Then the variance contribution knowledge at each fidelity level gives some information to smartly decide the fidelity level to choose. Let c_0, \dots, c_L be respectively the querying costs of all the fidelity levels f_0, \dots, f_L . Let us denote $\sigma_{red}^2(l, x^*)$ the variance reduction of the high fidelity model when the point x^* is evaluated with all the fidelity levels $\leq l$ to ensure the nested DoE structure

$$\sigma_{red}^2(l, x^*) = \sum_{i=0}^l \sigma_{\delta_i}^2(x^*) \prod_{j=i}^{L-1} \rho_j^2 \quad (6)$$

A criterion to choose the level of enrichment can then be written as:

$$l^* = \arg \max_{l \in \{0, \dots, L\}} \frac{\sigma_{red}^2(l, x^*)}{(\sum_{i=0}^l c_i)^2} \quad (7)$$

One can remark that Eq (7) is a trade off between variance reduction and computational cost. Then the cost ratios between the different fidelity levels are important parameters to the method. This two step approach from [3] is described in Fig. 3 of Appendix 7.2. It is referred in the following as the Multi-Fidelity Super Efficient Global Optimization (MFSEGO). Now that the multi-fidelity model and optimization technique have been described, the method for selecting the Pareto optimal fidelity level that we experimented with in this paper is explained.

3. Pareto optimal fidelity levels methodology

In this section, we present a preprocessing method whose goal is to select the Pareto optimal fidelity levels (POFL) of a function to be used in a multi-fidelity context. We consider two objectives which are the accuracy and the computational cost associated with each fidelity level. In a multi-objective context, the notion of optimality must be revised, and we speak of Pareto optimality [4]. A point A dominates a point B if it is better in all objectives. For example, in our case, a fidelity level that is more accurate and less expensive than another fidelity level will dominate it. A point is said non-dominated and is called Pareto optimal if no other evaluated point dominates it. The Pareto set is the set of Pareto optimal points. The Pareto front is the set of images of the Pareto set (the objective values at the points of the Pareto set).

Suppose we have $L + 1$ fidelity levels $f_0, \dots, f_1, \dots, f_L$ that evaluate a function with different accuracies. We consider here that the fidelity levels are not hierarchically ordered, except for the last fidelity level L which is the highest fidelity level. Typically, it is possible that some of these fidelity levels are not very interesting to use. For example, if one fidelity level is less accurate than another but still more computationally expensive, it is not interesting to use it in the multi-fidelity process we are building. The strategy we propose here is to evaluate each

level of fidelity at points Q_i of a DoE Q and estimate their accuracy and cost. The number of points is $\text{card}(Q) = n_{pts}$. To approximate the computational cost c_l associated to the l -th level of fidelity, we compute the mean value \hat{c}_l defined by

$$\hat{c}_l = \frac{1}{n_{pts}} \sum_{i=1}^{n_{pts}} c_{i,l} \quad (8)$$

where $c_{i,l}$ is the computational cost at point Q_i with the level l . Assuming that the higher fidelity level f_L is the reference, $\forall l \in \mathbb{L} = \{0, \dots, L\}$ we approximate e_l the accuracy of the l -th fidelity level by

$$\hat{e}_l = \frac{1}{\frac{1}{n_{pts}} \sum_{p=1 \dots P} \frac{\|(f_L(Q))_p - (f_l(Q))_p\|_2^2}{\frac{1}{n_{pts}} \sum_{i=1}^{n_{pts}} (f_L(Q_i))_p}} \quad (9)$$

$f_l(Q)$ is the output of the evaluation of f_l at all the points Q . Denoting P equals to the sum of the number of objectives and the number of constraints, $\forall l, \forall i, f_l(Q_i)$ is a vector of size P . Thus the estimated accuracy of the fidelity levels depends on the accuracy of the objective value but also on the constraints. Next, we construct the Pareto front in terms of cost and accuracy to select only the optimal Pareto fidelity levels. The dominated fidelity levels will not be considered in the process [28].

4. Analytical test cases

The POFL method is validated on a selection of analytical test cases of increasing complexity (unconstrained and constrained Rosenbrock in 2-dimension and the Borehole function in dimension 8) to show its utility.

4.1. Four fidelities for the 2D-Rosenbrock function

For the purpose of this test case, we extend the two-dimensional two-fidelity Rosenbrock function to a four-fidelity function using the method proposed by [24]. The high fidelity (HF) Rosenbrock function is defined by:

$$f_{HF} : \begin{cases} [-2, 2]^2 & \longrightarrow \mathbb{R} \\ (x_0, x_1) & \longmapsto 100(x_1 - x_0^2)^2 + (x_0 - 1)^2 \end{cases} \quad (10)$$

The low fidelity (LF) Rosenbrock functions are of the following form

$$f_{LF, x_s} : \begin{cases} [-2, 2]^2 & \longrightarrow \mathbb{R} \\ (x_0, x_1) & \longmapsto 100(x_s x_1 - x_0^2)^2 - 2((x_1 - x_s)^2 + (x_0 - x_s)^2) \end{cases} \quad (11)$$

The three LF functions are built using the same equation, but are parameterized by x_s . Table 1 summarizes the chosen parameter x_s and the chosen computational cost ratio for each fidelity level. Knowing x_s and the cost ratio for each fidelity level, the POFL method defined in Section 3

Table 1: LF and HF Rosenbrock functions: x_s parameter and associated cost ratio to define the four fidelity levels.

	LF1	LF2	LF3	HF
x_s	0.4	0.65	0.9	1
cost ratio	1000	10	100	1

is applied to build the associated Pareto front in terms of cost and accuracy using the data set Q with $n_{pts} = 100$ points. The Pareto front is depicted in Fig 1a, we can clearly see that with the chosen cost ratios, the second low fidelity function LF2 is dominated and could be removed from the list of available fidelities.

4.2. Four fidelities for the constrained 2D-Rosenbrock function

This test case uses the same functions as the ones defined in Section 4.1 with some different cost ratios. We also add the following constraint

$$c_a(x_0, x_1) = -x_0^2 - (x_1 - 1)^{\frac{1}{2}} + 1.1 + a \sin(10x_0 + 5x_1) \geq 0 \quad (12)$$

where a is an amplitude parameter for the sine. Table 2 sums up the value of this parameter a and the cost ratio for each fidelity level.

Table 2: Constrained Rosenbrock test case: amplitude parameter and cost ratio for each fidelity level

	LF1	LF2	LF3	HF
a	0.15	0.1	0.05	0
cost ratio	100	1000	10	1

The Pareto front obtained with $n_{pts} = 100$ points is shown in Fig 4 in Appendix 7.3. One can observe that the first low fidelity function LF1 does not belong to the Pareto front and could be removed

On these first 2D examples, the POFL method proposes to eliminate one level of fidelity and thus to decrease the number of levels of fidelity from 4 to 3.

4.3. Three fidelities for the 8D-Borehole function

The Borehole function is an unconstrained 8-dimensional test case, derived here to have three fidelity levels denoted by LF, MF and HF. The form of the Borehole function is expressed as:

$$f_{b_0, b_1} : \begin{cases} \Omega & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{b_0 x_2 (x_3 - x_5)}{\ln \frac{x_1}{x_0} (b_1 + \frac{2x_6 x_2}{\ln(x_1/x_0) x_0^2 x_7} + \frac{x_2}{x_4})} \end{cases} \quad (13)$$

where b_0 and b_1 are two parameters and the input space $\Omega \subset \mathbb{R}^8$ is defined by

$$\Omega = [0.05, 0.15] \times [100, 50000] \times [63070, 115600] \times [990, 1110] \times [63.1, 116] \times [700, 820] \times [1120, 1680] \times [9855, 12045]$$

Table 3 gives the values of the parameters b_0 and b_1 and the chosen cost ratio for each fidelity level.

Table 3: Borehole parameters and cost ratio for each fidelity level

	LF	MF	HF
b_0	5	7	2π
b_1	1.5	0.5	1
cost ratio	10	100	1

The Pareto front obtained with $n_{pts} = 100$ points is shown in Fig. 5 in Appendix 7.3. The LF function does not belong to the Pareto front and could be eliminated.

4.4. Optimization results of analytical test cases

In this section the optimization results obtained on the analytical test cases are described. We compared a mono-fidelity optimization with several multi-fidelity optimizations: one using all the fidelity levels available, one using only the Pareto optimal fidelity levels and one using fidelity levels randomly chosen. Thus four scenarios are evaluated and compared.

The multi-fidelity Kriging models are computed with SMT[21, 23, 14]. SMT is an open source toolbox developed by ONERA, ISAE-SUPAERO, NASA Glenn and University of

Michigan specialized on the fast development of tailored surrogate models. The optimizations are performed with SEGOMOE-toolbox [13] a Bayesian optimization toolbox developed at ONERA and ISAE-SUPAERO including the proposed Multi-Fidelity strategy (MFSEGO). For the random fidelity selection case, the HF level is always selected and the number of levels is equal to the one chosen by the POFL strategy. For each scenario, 10 runs are performed using a squared exponential kernel and the WB2 acquisition function [8]. For each run, first we built an initial DoE for the optimization that uses all of the fidelity levels. Next, to construct the initial DoE for each other scenario, we examine the selected fidelity levels with the POFL method, and then select the parts corresponding to those fidelity levels from the full initial DoE. For all analytical test cases, the initial DoE size for all multi-fidelity optimizations was chosen as follows. First, we ranked all fidelity levels by accuracy, from $k = 0$ the least accurate to $k = L$ the most accurate. Next, we chose the initial DoE size for the HF level with $ndoe_{HF}$ equal to twice the problem dimension, i.e., $ndoe_{HF} = 4$ for the two Rosenbrock problems and $ndoe_{HF} = 16$ for the Borehole problem. Finally, the number of points evaluated with the k -th level of fidelity is determined by the formula

$$ndoe_{fid_k} = (ndoe_{HF})2^{L-k}$$

As MFSEGO needs a nested DoE structure as mentioned in Section 2.1, the initial DoEs have then been created respecting this requirement. Reference solutions are known for the analytical problems. Given this reference solution ref_{sol} , a convergence criterion for the mean (over the 10 runs) of the objective value at the best feasible point $\bar{y}_{best}^{feasible}$ is defined by:

$$|\bar{y}_{best}^{feasible} - f_{HF}(ref_{sol})| \leq \epsilon_0 |f_{HF}(ref_{sol})| + \epsilon_1 \quad (14)$$

where ϵ_0 and ϵ_1 are respectively the relative and absolute tolerances. Typically we chose $\epsilon_0 = \epsilon_1 = 1e^{-3}$ for the analytical test cases. Figure 1b shows the evolution of the mean of the best feasible objective value against the budget for the unconstrained Rosenbrock test case. Figures 7 and 8 in Appendix 7.4 show the mean value evolutions respectively for the constrained Rosenbrock and Borehole test cases. Overall results are summed up in Table 4. The term "not reached" in this table means that the maximum allocated budget for the optimization has been reached before the convergence criterion Eq. (14) was satisfied. For all the analytical test cases we can see that the POFL method overtakes all other methods. Moreover these results highlight that using all of the fidelity levels available is not efficient, even compared to the mono-fidelity strategy. In Fig 1b and Fig 7 for unconstrained and constrained Rosenbrock cases, the curve for the random scenario quickly approaches the reference solution, which means that for some of the 10 runs the random strategy was effective in finding the optimum. Nevertheless, the convergence criterion (see Eq. (14)) on the average over the 10 runs is not even reached using the whole allowed budget. We can therefore deduce that the random strategy is not at all robust.

Table 4: Summary of the needed budget to reach the convergence criterion for the three analytical test cases with $\epsilon_0 = \epsilon_1 = 1e^{-3}$ ("not reached" means that the maximum allocated budget for the optimization has been reached before the convergence criterion Eq. (14) was satisfied).

	Unconstrained Rosenbrock	Constrained Rosenbrock	Borehole
All fidelity levels	not reached	not reached	37.0
Mono-fidelity	89	67.0	34.4
POFL	75.2	30.686	19.84
Random	not reached	not reached	43.95

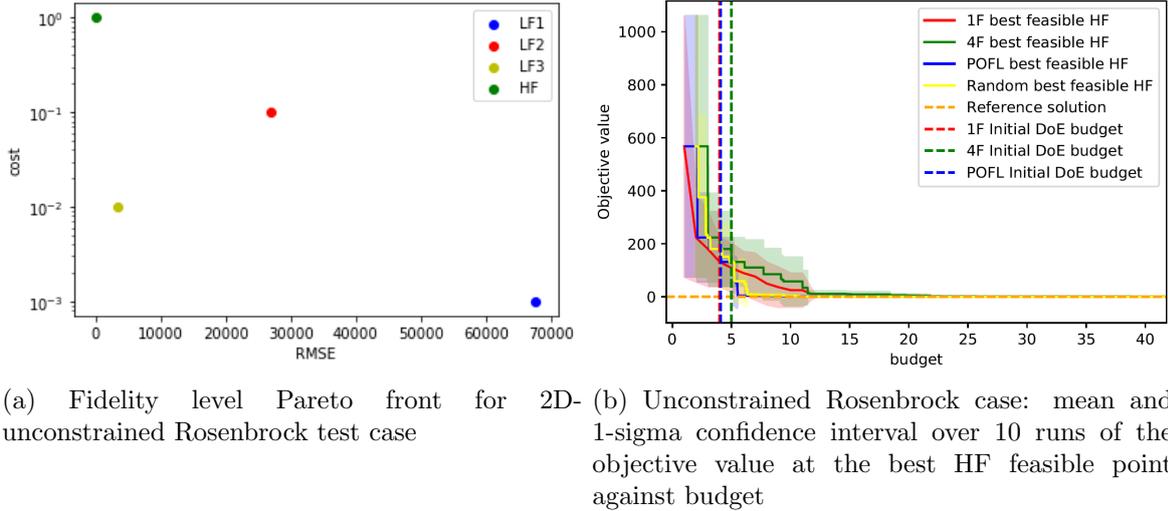


Figure 1: Unconstrained Rosenbrock test cases fidelity level Pareto front and convergence curves.

5. Drone design test case

5.1. The drone model

For the purpose of this test case, we developed a parametric drone model with the OpenMDAO framework [17]. The core of the model is the aerostructural part which relies on an aerostructural simulation code called OpenAeroStruct [15] [12] [16] (OAS) developed by University of Michigan within OpenMDAO. The model has a propulsion group handled using the electrical extension of the Breguet range equation. This equation is valid in the steady level flight case and its expression is given in Appendix 7.6. In order for the drone to ensure longitudinal stability, the model also has a group that balances the root chord tail such that the tail surface over wing surface ratio is equal to 0.15. This model has been derived in 5 fidelity levels by combining different mesh dimensions and drone speed parameterizations. The different combinations are given in Table 5. Using different speeds allows us to simulate several codes of a priori similar levels. This is a way to emulate more levels of fidelity for our problem without using more tools and codes.

Table 5: Parameterization of the OpenAeroStruct models

	wing mesh dim (chordwise, spanwise)	tail mesh dim (chordwise, spanwise)	Drone speed
HF	(19,41)	(17,37)	20 $m.s^{-1}$
MF1	(9,15)	(7,11)	20 $m.s^{-1}$
MF0	(9,15)	(7,11)	30 $m.s^{-1}$
LF1	(3,5)	(3,5)	20 $m.s^{-1}$
LF0	(3,5)	(3,5)	30 $m.s^{-1}$

The MDO problem to be solved can be defined as:

$$\max_{x \in \mathbb{R}^{19}} SOC(x)$$

$$\text{such that } \begin{cases} CL \geq 0.4 \\ \text{wing}_{\text{failure}} \leq 0 \\ \text{tail}_{\text{failure}} \leq 0 \end{cases}$$

where the SOC is the state of charge of the battery at the end of the mission and x is the vector of 19 design variables relative to wing and tail geometrical parameters. They are detailed in

Table 8 in Appendix 7.5. The Pareto front obtained with $n_{pts} = 100$ points is shown in Fig 6 in Appendix 7.3. We can see that the two models MF0 and LFO that consider a 30 ms^{-1} speed are not Pareto optimal.

5.2. Optimization results for the drone design

This section presents the optimization results obtained with the drone design test case. The size of the initial DoEs and the cost ratios are given in Table 6.

Table 6: Size of initial DoEs and cost ratios for the drone design test case.

	LF0	LF1	MF0	MF1	HF
Initial DoE size	20	20	10	10	5
cost ratio	100	100	10	10	1

The same convergence criterion, see Eq. (14), as for the analytical cases is now parameterized by $\epsilon_0 = 0.005$ and $\epsilon_1 = 0.0025$ and the reference solution has been found using the SNOPT algorithm [22] which stands for Sparse Nonlinear OPTimizer. For this test case we also compared the results with an other type of fidelity selection strategy that we named dominated selection. Indeed, since the low and medium fidelity levels are mainly intended to explore the design space, we decided to check the impact on the optimization performance of selecting the dominated low fidelity models instead of the non-dominated low fidelity models. Then, the same methodology as for the analytical cases is used to obtain the results presented in Fig. 2. The representation is slightly different to show the budget. For each of the 5 strategies, the number of runs that found the optimum is given by the blue line and the number of executions that found a feasible point with respect to the budget by the green line. Table 7 sums up the required budget to reach the convergence criterion. In the end, the POFL method is the most efficient. The dominated selection method is less efficient than the POFL method but still better than all other strategies. The mono-fidelity strategy is not the most effective but it is quite robust. On the opposite the random strategy is very effective for most of the runs but is slow for one of them. It emphasizes that this strategy is not robust to the initial DoE as we started to see with the analytical test cases. Finally, the all fidelity level strategy seems to be the worst. Using a large number of fidelity levels can lead to redundancy and complicate surrogate modeling, which could explain this inefficiency.

Table 7: Summary of the needed budgets to reach the convergence criterion for the drone design test case with $\epsilon_0 = 5 \times 1e^{-3}$ and $\epsilon_1 = 2.5 \times 1e^{-3}$.

All fidelity levels	Mono-fidelity	POFL	dominated selection	Random
24.87	14	8.71	10.59	11.41

6. Conclusion

In this work, we proposed a method to select the most useful fidelity levels to use in a multi-fidelity process among all available levels. We experimented with this POFL method on several analytical optimization problems and on a UAV design problem that we solved using single and multi-fidelity Bayesian optimizations. Overall, the POFL selection proves to be the most computationally efficient for solving these problems. The budget required to achieve convergence was divided by a factor between 1.2 and 2.2 using the POFL method compared to the mono-fidelity method. However, the POFL has some limitations. For example, if the cost and accuracy of each model is not known or has not yet been estimated, it must be done to obtain the Pareto front and this can be expensive. Nevertheless, once this has been done once, the cost and accuracy estimates can be reused.

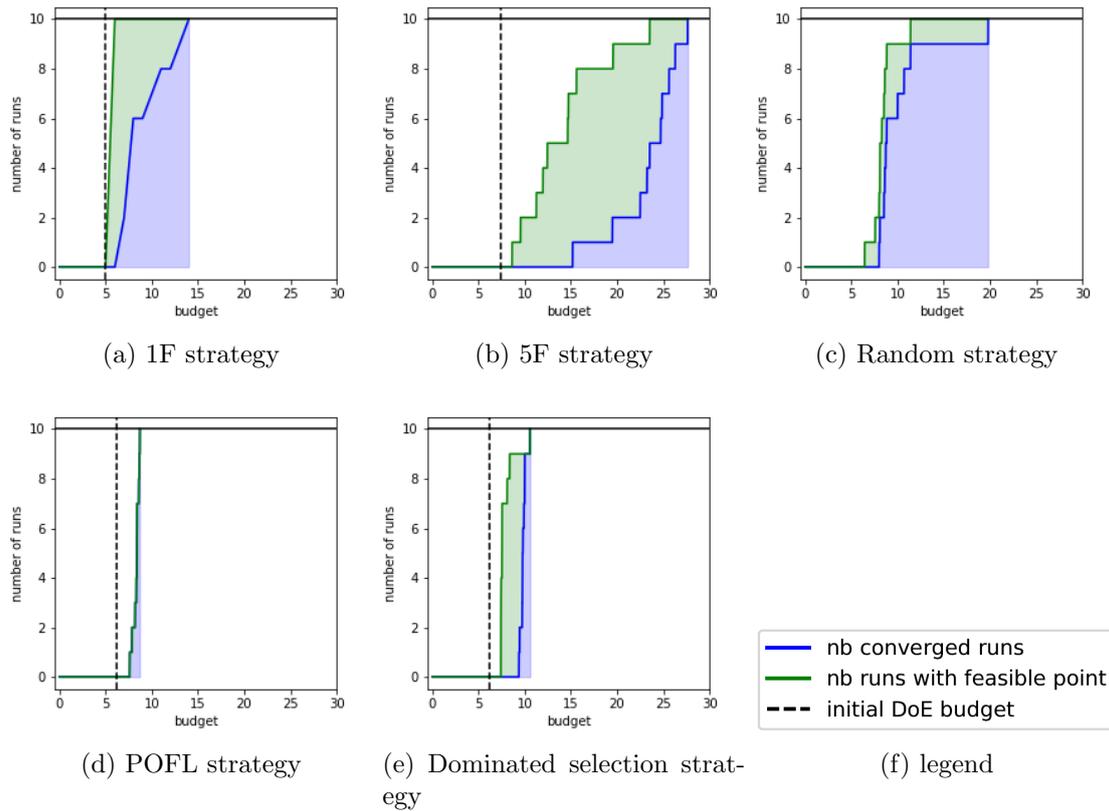


Figure 2: Evolution of the number of runs that have found a feasible point and the number that have found the optimum.

Acknowledgments

The PhD is funded by the defense innovation agency (AID) and by the french Directorate General of Armaments (DGA) as part of the CONCORDE project. This work is part of the activities of ONERA - ISAE - ENAC joint research group.

References

- [1] Bayesian calibration of computer models, Kennedy, Marc C and O’Hagan, Anthony, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, volume 63, number 3, pages 425–464, 2001, publisher Wiley Online Library.
- [2] Multi-fidelity Gaussian process regression for computer experiments, Le Gratiet, Loic, 2013, Université Paris-Diderot-Paris VII.
- [3] Multi-fidelity efficient global optimization: Methodology and application to airfoil shape design, Meliani, Mostafa and Bartoli, Nathalie and Lefebvre, Thierry and Bouhleh, Mohamed-Amine and Martins, Joaquim RRA and Morlier, Joseph, *AIAA aviation 2019 forum*, pages 3236, 2019.
- [4] Martins, Joaquim R. R. A. and Ning, Andrew, Cambridge University Press, *Engineering Design Optimization*, 2020.
- [5] Efficient global optimization of expensive black-box functions, Jones, Donald R and Schonlau, Matthias and Welch, William J, *Journal of Global optimization*, volume 13, number 4, pages 455–492, 1998, Springer.
- [6] An adaptive optimization strategy based on mixture of experts for wing aerodynamic design optimization, Bartoli, Nathalie and Lefebvre, Thierry and Dubreuil, Sylvain and Olivanti, Romain and Bons, Nicolas and Martins, Joaquim RRA and Bouhleh, Mohamed-Amine

- and Morlier, Joseph, 18th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, pages 4433, 2017.
- [7] Improvement of efficient global optimization with application to aircraft wing design, Bartoli, Nathalie and Bouhleb, Mohamed-Amine and Kurek, Igor and Lafage, Rémi and Lefebvre, Thierry and Morlier, Joseph and Priem, Rémy and Stilz, Vivien and Regis, Rommel, 17th AIAA/ISSMO Multidisciplinary analysis and optimization conference, pages 4001, 2016.
 - [8] Infill sampling criteria to locate extremes, Watson, Alan G and Barnes, Randal J, *Mathematical Geology*, volume 27, number 5, pages 589–608, 1995, Springer.
 - [9] Flexibility and efficiency enhancements for constrained global design optimization with Kriging approximations, Sasena, Michael James, 2002, Citeseer.
 - [10] A statistical approach to some basic mine valuation problems on the Witwatersrand, Krige, Daniel G, *Journal of the Southern African Institute of Mining and Metallurgy*, volume 52, number 6, pages 119–139, 1951, Southern African Institute of Mining and Metallurgy.
 - [11] Cross validation and maximum likelihood estimations of hyper-parameters of Gaussian processes with model misspecification, Bachoc, François, *Computational Statistics & Data Analysis*, volume 66, pages 55–69, 2013, Elsevier.
 - [12] Low-fidelity aerostructural optimization of aircraft wings with a simplified wingbox model using OpenAeroStruct, Chauhan, Shamsheer S and Martins, Joaquim RRA, *International Conference on Engineering Optimization*, pages 418–431, year 2018, Springer.
 - [13] SEGOMOE: Super Efficient Global Optimization with Mixture of Experts, Priem, Rémy and Bartoli, Nathalie and Diouane, Youssef and Lefebvre, Thierry and Dubreuil, Sylvain and Salaün, Michel and Morlier, Joseph, *Workshop CIMI Optimization & Learning*, 2018.
 - [14] A Python surrogate modeling framework with derivatives, Bouhleb, Mohamed Amine and Hwang, John T and Bartoli, Nathalie and Lafage, Rémi and Morlier, Joseph and Martins, Joaquim RRA, *Advances in Engineering Software*, volume 135, pages 102662, 2019, Elsevier.
 - [15] Open-source coupled aerostructural optimization using Python, Jasa, John P and Hwang, John T and Martins, Joaquim RRA, *Structural and Multidisciplinary Optimization*, volume 57, number 4, pages 1815–1827, 2018, Springer.
 - [16] Multifidelity optimization under uncertainty for a tailless aircraft, Chaudhuri, Anirban and Jasa, John and Martins, Joaquim RRA and Willcox, Karen E, 2018 AIAA Non-Deterministic Approaches Conference, pages 1658, 2018.
 - [17] OpenMDAO: An open-source framework for multidisciplinary design, analysis, and optimization, Gray, Justin S and Hwang, John T and Martins, Joaquim RRA and Moore, Kenneth T and Naylor, Bret A, *Structural and Multidisciplinary Optimization*, volume 59, number 4, pages 1075–1104, 2019, Springer.
 - [18] Computing the maximum likelihood estimates: concentrated likelihood, EM-algorithm, Pavlyuk, Dmitry.
 - [19] *Gaussian processes for machine learning*, vol. 1, Rasmussen, Carl Edward and Williams, CK, 2006, MIT press Cambridge MA.
 - [20] *Mémoires du Bureau de Recherche Géologiques et Minières*, n. 14, Matheron, G and de Géostatistique Appliquée, *Traité and Tome, I*, Ed. Technip, Paris, 1962.
 - [21] Mohamed Amine Bouhleb and John T. Hwang and Nathalie Bartoli and Rémi Lafage and Joseph Morlier and Joaquim R. R. A. Martins, *Advances in Engineering Software*, A Python surrogate modeling framework with derivatives, pages 102662, 2019, issn 0965-9978, <https://doi.org/10.1016/j.advengsoft.2019.03.005>, 2019.
 - [22] *User’s Guide for SNOPT Version 7: Software for Large-Scale Nonlinear Programming*, Philip, E and Murray, Walter and Saunders, Michael A, 2015.
 - [23] Mohamed Amine Bouhleb and John T. Hwang and Nathalie Bartoli and Rémi Lafage and Joseph Morlier and Joaquim R. R. A. Martins, *Surrogate Modelling Toolbox*, 2022, GitHub repository, <https://github.com/SMTorg/smt>

- [24] Comparison of generic multi-fidelity approaches for bound-constrained nonlinear optimization applied to adjoint-based CFD applications, Olivanti, Romain and Gallard, François and Brézillon, Joël and Gourdain, Nicolas, AIAA Aviation 2019 Forum, 3102, 2019.
- [25] Non-Myopic Multifidelity Method for Multi-regime Constrained Aerodynamic Optimization, Di Fiore, Francesco and Mainini, Laura, AIAA AVIATION 2022 Forum, pages 3716, 2022.
- [26] Geostatistics: modeling spatial uncertainty, Chiles, Jean-Paul and Delfiner, Pierre, volume 497, 2009, John Wiley & Sons.
- [27] Multivariate geostatistics: an introduction with applications, Wackernagel, Hans, 2003, Springer Science & Business Media.
- [28] A Gradient-based Sequential Multifidelity Approach to Multidisciplinary Design Optimization, Wu, Neil and Mader, Charles A and Martins, Joaquim RRA, Structural and Multidisciplinary Optimization, volume 65, number 4, pages 1–20, 2022, Springer.
- [29] Multi-fidelity constrained Bayesian optimization, application to drone design, Charayron, Remy and Lefebvre, Thierry and Bartoli, Nathalie and Morlier, Joseph, CAID 2021 Forum, 2021.
- [30] Bayesian optimization and data science, Archetti, Francesco and Candelieri, Antonio, 2019.
- [31] A tutorial on Bayesian optimization, Frazier, Peter I, 2018.

7. Appendix

In this appendix, we first find a pseudo-code of the MFSEGO methodology in Section 7.1. Then, Section 7.2 presents this MFSEGO methodology in a diagrammatic form. Section 7.3 shows the fidelity level Pareto fronts for the constrained Rosenbrock, Borehole and drone design test cases. Section 7.4 shows the optimization convergence curves for the constrained Rosenbrock and Borehole test cases. A table summarizing the lower and upper bounds, the unit and the dimension of all of the design variables for the drone test case is provided in Section 8. Finally, Section 7.6 details how to find the expression of the electric Breguet range equation.

7.1. Pseudo-code of the multi-fidelity Bayesian optimization methodology based on Le Gratiet’s recursive formulation.

The following pseudo-code explains the MFSEGO methodology.

Algorithm 1 Multi-fidelity Bayesian optimization algorithm, based on Le Gratiet’s recursive formulation.

```

Compute initial DoE using LHS
while (maximum budget is not reached) and (convergence criterion not reached) do
  Learn LF Kriging surrogate model ( $\hat{\mu}_0$  and  $\hat{\sigma}_0^2$ )
  for  $k = 1 \dots L$  do
    Learn  $\rho_{k-1}$  and  $\hat{\mu}_{\delta_k}$ 
    Deduce  $\hat{\mu}_k$  and  $\hat{\sigma}_k^2$  and so the  $k$ -th fidelity level Kriging surrogate model
  end for
  Choose  $x_{next}$  that maximizes the acquisition function
  Select the level of enrichment  $t$  (see Eq. (7))
  for  $l = 0 \dots t$  do
    Add  $(x_{next}; f_l(x_{next}))$  to the DoE
  end for
   $y_{best}^{feasible} = \min Y^{feasible}$ 
end while
return  $y_{best}^{feasible}$ 

```

7.2. Diagram of the multi-fidelity Bayesian optimization methodology based on Le Gratiet's recursive formulation.

Figure 3 describes the MFEGO methodology schematically.

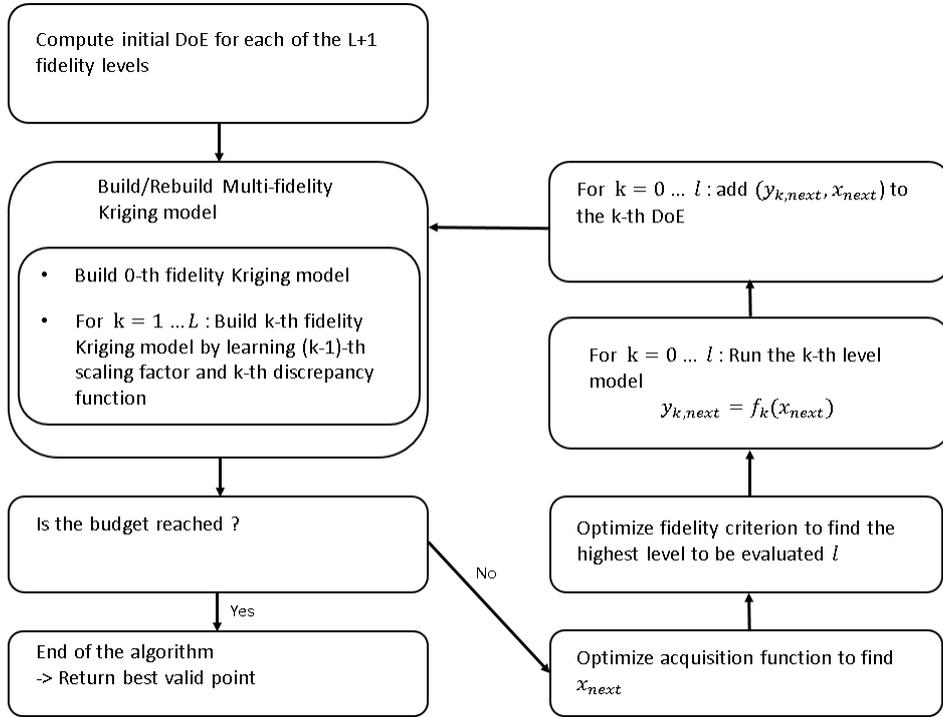


Figure 3: MFSEGO methodology diagram.

7.3. Others Pareto front figures

Figure 4 depicts the fidelity level Pareto front for the constrained 2D-Rosenbrock test case.

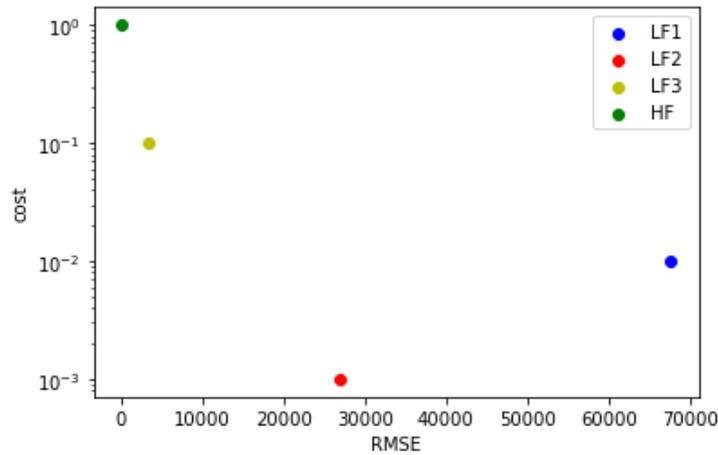


Figure 4: Fidelity level Pareto front for the constrained 2D-Rosenbrock test case

Figure 5 depicts the fidelity level Pareto front for the 8D-Borehole test case.

Figure 6 depicts the fidelity level Pareto front for the drone design test case.

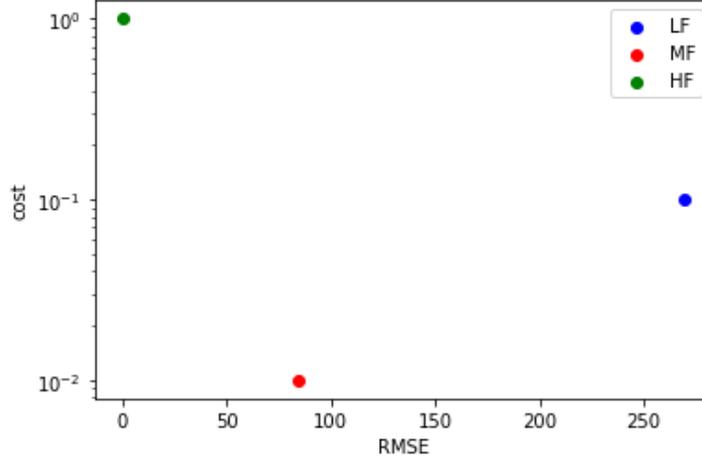


Figure 5: Fidelity level Pareto front for the Borehole test case

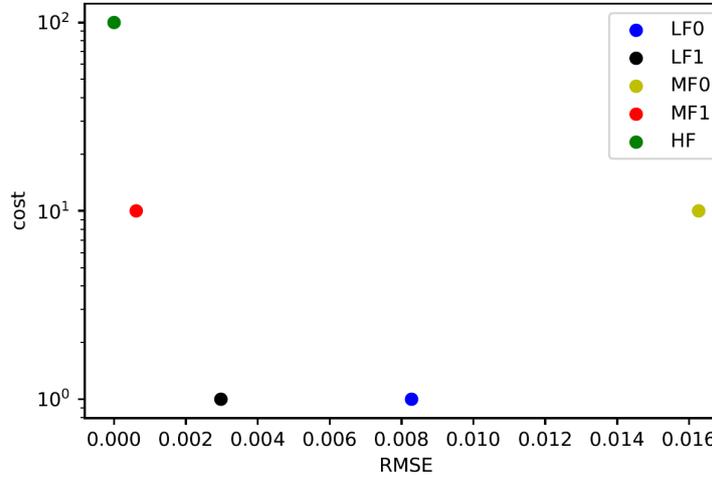


Figure 6: Fidelity level Pareto front for the drone design test case

7.4. Others optimization results figures

Figure 7 shows the evolution of the mean over the 10 runs of objective value at the best HF feasible point along the optimization for the constrained 2D-Rosenbrock test case. Figure 8 shows the evolution of the mean over the 10 runs of objective value at the best HF feasible point along the optimization for the constrained 8D-Borehole test case.

7.5. Design variables of the drone design problem

Table 8 sums up the design variables of the drone design optimization problem, their lower and upper bound, their unit and their dimension.

7.6. Electric Breguet range equation demonstration

First lets notice that the range R can be expressed with the speed and the time flying at that speed: $R = V \times t$. Under ideal conditions, the time needed to empty a battery is given by:

$$t = \frac{m_{\text{batt}} e_b}{P_{\text{batt}}} \quad (15)$$

where e_b is the specific energy density of the battery (unit: J.kg^{-1}), m_{batt} is the battery mass and P_{batt} is the battery power. Lets define the overall propulsion system efficiency $\eta_{\text{tot}} = \frac{P_{\text{prop}}}{P_{\text{batt}}}$ with P_{prop} is the required propulsive power to reach the equilibrium. η_{tot} can be expressed as

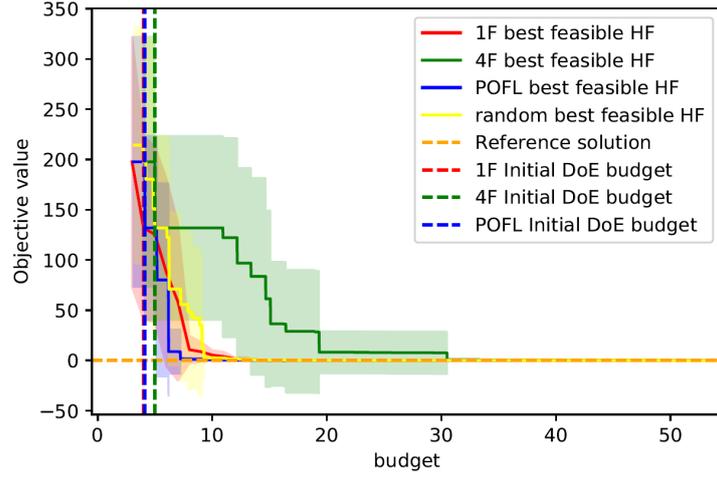


Figure 7: Constrained Rosenbrock case: mean and 1-sigma confidence interval over 10 runs of the objective value at the best HF feasible point against budget

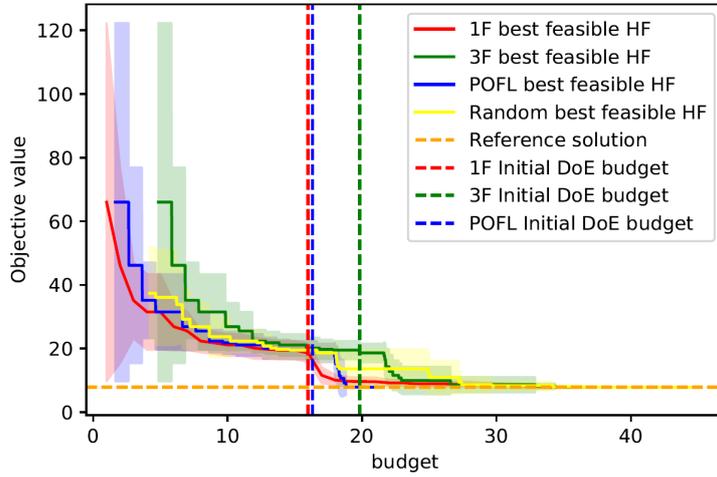


Figure 8: Borehole case: mean and 1-sigma confidence interval over 10 runs of the objective value at the best HF feasible point against budget

Table 8: Design variables of the drone design problem

	lower bound	upper bound	unit	dimension
$\text{span}_{\text{wing}}$	5	6	m	1
$\text{root}_{\text{chord wing}}$	0.9	1.2	m	1
$\text{taper}_{\text{wing}}$	0.6	1	\emptyset	1
$\text{sweep}_{\text{wing}}$	0	5	deg	1
$\text{sweep}_{\text{tail}}$	0	5	deg	1
$\text{thickness}_{\text{cp wing}}$	[0.001; 0.001; 0.001]	[0.01; 0.01; 0.01]	m	3
$\text{thickness}_{\text{cp tail}}$	[0.001; 0.001; 0.001]	[0.01; 0.01; 0.01]	m	3
$\text{dihedral}_{\text{wing}}$	-3	3	deg	1
$\text{dihedral}_{\text{tail}}$	-3	3	deg	1
$\text{twist}_{\text{cp wing}}$	[0; 0.5; 1]	[0.5; 1; 1.5]	deg	3
$\text{twist}_{\text{cp tail}}$	[0; 0.5; 1]	[0.5; 1; 1.5]	deg	3
Total number of variables	19			

the product of 3 individual efficiencies: η_p the propeller efficiency, η_m the motor efficiency and η_{ESC} the electric speed controller efficiency. $P_{\text{prop}} = TV$ with T the required thrust. Assuming the steady level flight hypothesis, we have

$$\begin{cases} T = D \\ L = W \end{cases} \quad (16)$$

where L is the lift force, D is the drag force and W is the weight. Equation (16) implies

$$T = \frac{L}{\frac{L}{D}} = \frac{W}{\frac{L}{D}} = \frac{mg}{\frac{L}{D}}$$

with g being the gravity acceleration. A relation is now available for the speed:

$$V = \frac{P_{\text{prop}}}{T} = \frac{P_{\text{batt}} \eta_{\text{tot}}}{\frac{mg}{\frac{L}{D}}} \quad (17)$$

Then the final expression for the range is given by:

$$R = V \times t = \frac{m_{\text{batt}} e_b}{P_{\text{batt}}} \times \frac{P_{\text{batt}} \eta_{\text{tot}}}{\frac{mg}{\frac{L}{D}}} = \frac{L}{D} \eta_p \eta_m \eta_{ESC} \frac{e_b}{g} \frac{m_{\text{batt}}}{m} \quad (18)$$

Note that the $\frac{L}{D}$ ratio can be replaced by the ratio of lift and drag coefficient $\frac{C_L}{C_D}$. Indeed, $L = qSC_L$ and $D = qSC_D$ with S being the lifting surface area and q being the dynamic pressure.