IX International Conference on Computational Methods in Marine Engineering
MARINE 2021
I.M. Viola and F. Brennan (Eds)

TOWARDS AUTOMATIC PARAMETER SELECTION FOR MULTI-FIDELITY SURROGATE-BASED OPTIMIZATION

RICCARDO PELLEGRINI*, JEROEN WACKERS†,⋆, ANDREA SERANI*, MICHEL VISONNEAU†, AND MATTEO DIEZ*

*National Research Council - Institute of Marine Engineering (CNR-INM)
Via di Vallerano 139, 00128 Rome, Italy

† LHEEA Lab, Ecole Centrale de Nantes, CNRS UMR 6598,
1 rue de la Noe, 44321 Nantes cedex 3, France
e-mail: jeroen.wackers@ec-nantes.fr

Key words: Multi-fidelity, Surrogate-based Optimization, Stochastic Radial Basis Functions

Abstract. The performance of surrogate-based optimization are highly affected by how the surrogate training set is defined. This is especially true for multi-fidelity surrogate models, where different training sets exist for each fidelity. Adaptive sampling methods have been developed to improve the fitting capabilities of surrogate models, avoiding to define an a priori design of experiments, adding training points only where necessary or most useful (i.e., providing the highest knowledge gain) to the optimization process. Nevertheless, the efficiency of the adaptive sampling is highly affected by its initialization. The paper presents and discusses a novel initialization strategy with a limited training set for adaptive sampling. The proposed strategy aims to reduce the computational cost of evaluating the initial training set. Furthermore, it allows the surrogate model to adapt more freely to the data. In this work, the proposed approach is applied to single- and multi-fidelity stochastic radial basis functions for an analytical test problem and the shape optimization of a NACA hydrofoil. Numerical results show that the results of the surrogate-based optimization are improved, thanks to a more effective and efficient domain space exploration and a significant reduction of high-fidelity evaluations.

1 INTRODUCTION

Automatic shape optimization and uncertainty quantification offer rigorous and effective mathematical approaches to the design and performance assessment of modern ships and ships' subsystems. These methods generally require a large number of evaluations of one or more merit factors. If these performance metrics are evaluated via high-fidelity computations, the computational cost can become prohibitively expensive and unaffordable for most users. Surrogate models give a solution to this problem: computations are only performed in a few design points or conditions and an interpolatory or regressive model is build based on these computations. Shape optimization and uncertainty quantification are then performed using the surrogate model,

which is inexpensive to evaluate. An even greater gain in efficiency is obtained through multi-fidelity surrogate models, which combine computationally inexpensive low-fidelity simulations with high-fidelity computations [1]. Multiple fidelities can be defined combining computations performed with different physical models (*e.g.* Reynolds averaged Navier-Stokes equations and potential flow) or with different accuracy (*e.g.* computations performed varying the grid size).

The performance of multi-fidelity surrogate models depends on several factors [2], such as the presence of nonlinearities, the problem dimensionality, the noisy or smooth behavior of the function, and the approach used for the definition of the training set. Numerical experiments show that there is no unique optimal multi-fidelity approach: the best choice of the surrogate model depends on the data being modeled. For example, a significant local variation of the merit factors may require a high density of training data, the presence of noise may require heavy filtering, and finally low-fidelity corrections are only useful when the low- and high-fidelity data are sufficiently correlated (similar). The first two issues led the authors to adopt solution-adaptive sampling for the training points [3] and automatically tuned noise filters [4]. Obviously, these methods achieve better performance if a larger computational budget is available.

Here, the multi-fidelity model is based on a surrogate model of the low-fidelity solutions and a discrepancy surrogate model based on the difference between high- and low-fidelity solutions. Specifically, the surrogate model is based on stochastic radial-basis functions (SRBF) with power kernel. In the authors previous work, the initial training set for starting the surrogate-based optimization is defined using $2D + 1$ samples where D is the design space dimension. The training points are placed in the center of the domain and the center of the boundaries, for all the fidelities. This initialization approach requires a considerable computational cost only for the initialization.

This is the motivation to develop an approach to minimize the number of initial training points.

The objective of the present work is to introduce a new approach for the definition of the initial training set for surrogate-based optimization. The proposed approach reduces the computational cost of the initialization making available a larger budget for the adaptive sampling of the training points. Furthermore, less information is initially provided to the surrogate model in order to have it freely adapt to the data. The new approach uses only one initial point for all the fidelities except for the lowest. When a single training point is available, the surrogate model prediction is an extrapolation based on that single training point. The power kernel used in the SRBF lacks a compact support, as a consequence the extrapolated prediction may not be well correlated with the desired function behavior, negatively affecting the adaptive sampling. Although other kernels exist with compact support, the SRBF with power kernels is robust, showing good results for several applications [4] and are here preferred. Therefore, a constraint is imposed on both the surrogate model prediction and the associated uncertainty to improve the the adaptive sampling.

The proposed initialization approach is assessed for an analytical test problem and a simulation-based optimization problem: the drag-coefficient minimization of a NACA hydrofoil. For both problems one and three fidelities are used and the results are compared with the previous initialization approach [4]. The simulations are performed with the unsteady RANS solver ISIS-CFD [7], developed at Ecole Centrale de Nantes/CNRS and integrated in the FINE/Marine simulation suite from NUMECA Int. Mesh deformation and adaptive grid refinement are adopted,

different fidelity levels are defined by increasing the grid refinement.

2 MULTI-FIDELITY APPROACH

Consider $\mathbf{x} \in \mathbb{R}^D$ as a design variables vector of dimension D. Let the true function $f(\mathbf{x})$ be assessed by N fidelity levels: the highest-fidelity level is $f_1(\mathbf{x})$, the lowest-fidelity is $f_N(\mathbf{x})$, and the intermediate fidelity levels are $\{f_i\}_{i=2}^{N-1}(\mathbf{x})$. Using $\tilde{\cdot}$ to denote surrogate model prediction, the multi-fidelity (MF) approximation $\hat{f}_i(\mathbf{x})$ of $f_i(\mathbf{x})$ $(i = 1, \dots, N-1)$ is the sum of the lowest-fidelity surrogate and surrogates of the errors (inter-level errors or bridge-functions, $\tilde{\varepsilon}$) between subsequent levels

$$\hat{f}_i(\mathbf{x}) = \tilde{f}_N(\mathbf{x}) + \sum_{k=i}^{N-1} \tilde{\varepsilon}_k(\mathbf{x}). \tag{1}$$

For each i-th fidelity level the training set is $\mathcal{T}_i = \{\mathbf{y}_j, f_i(\mathbf{y}_j)\}_{j=1}^{J_i}$, with J_i the training set size. The resulting inter-level error training set is defined as $\mathcal{E}_i = \{\mathbf{y}_j, \varepsilon_i(\mathbf{y}_j)\}_{j=1}^{J_i}$, where

$$\varepsilon_i(\mathbf{y}_j) = f_i(\mathbf{y}_j) - \hat{f}_{i+1}(\mathbf{y}_j). \tag{2}$$

Consider the surrogate model providing the prediction and the associated uncertainty. The uncertainty $U_{\tilde{f}_N}$ of the lowest-fidelity prediction is considered as uncorrelated with the uncertainty $U_{\tilde{\varepsilon}_k}$ of the inter-level error predictions. Therefore, the uncertainty $U_{\hat{f}_i}$ of the MF prediction can be evaluated as $(i = 1, \dots, N-1)$

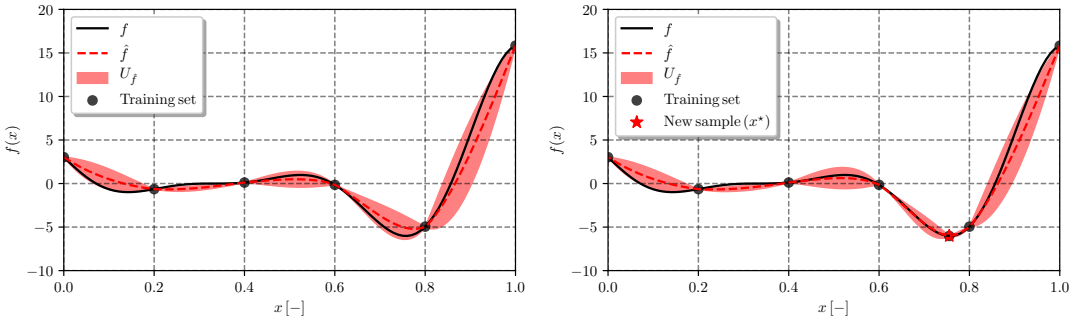
$$U_{\hat{f}_i}(\mathbf{x}) = \sqrt{U^2_{\tilde{f}_N}(\mathbf{x}) + \sum_{k=i}^{N-1} U^2_{\tilde{\varepsilon}_k}(\mathbf{x})}. \tag{3}$$

2.1 Adaptive Sampling Method

The multi-fidelity surrogate model is dynamically updated by adding new training points. First, a new training point $\mathbf{x}^\star$ is identified based on the aggregate-criteria adaptive sampling (ACAS, see Fig. 1) presented in [3]. It aims to find points with large prediction uncertainty and small objective function value. Accordingly, ACAS identifies a new training point by solving the single-objective minimization $\mathbf{x}^\star = \underset{\mathbf{x}}{\operatorname{argmin}} \left[\hat{f}(\mathbf{x}) - U_{\hat{f}}(\mathbf{x})\right]$. It may be noted that the ACAS method is a special case of the lower-confidence bounding [16], with equally weighted contributions of $\hat{f}$ and $U_{\hat{f}}$. Once $\mathbf{x}^\star$ is identified, the fidelity for its evaluation needs to be selected. To this extent, the surrogate model prediction uncertainty vector is defined as

$$\mathbf{U} \equiv \{U_{\tilde{\varepsilon}_1}/\beta_1, ..., U_{\tilde{\varepsilon}_{N-1}}/\beta_{N-1}, U_{\tilde{f}_N}/\beta_N\}, \tag{4}$$

where $\beta_i = c_i/c_1$ with c_i the computational cost associated to the i-th level and c_1 the computational cost of the highest-fidelity. Then, the k-th fidelity level to sample is $k = \text{maxloc}(\mathbf{U})$, and the new training point is added to the k-th training set $\mathcal{T}_k$ and to the lower-fidelity sets from $k+1$ up to N.



(a) (b)

Figure 1: Example of the adaptive sampling method using one fidelity: (a) shows the initial surrogate model with the associated prediction uncertainty and training set; (b) shows the position of the new training point and the new surrogate model prediction and its uncertainty.

2.2 Stochastic Radial Basis Functions

Given a (single-fidelity) training set $\mathcal{T} = \{\mathbf{y}_i, f(\mathbf{y}_i)\}_{i=1}^{J}$, the RBF surrogate model prediction $\tilde{f}(\mathbf{x})$ is computed as the expected value (EV) over a stochastic tuning parameter of the surrogate model [8], $\tau \sim \text{unif}[1,3]$

$$\widetilde{f}(\mathbf{x}) = \text{EV}\left[g(\mathbf{x},\tau)\right]_\tau, \quad \text{with} \quad g(\mathbf{x},\tau) = \text{EV}\left[\mathbf{f}\right] + \sum_{j=1}^{M} w_j ||\mathbf{x} - \mathbf{c}_j||^\tau, \tag{5}$$

where w_j are unknown coefficients, $||\cdot||$ is the Euclidean norm and $\mathbf{c}_j$ are the RBF centers, with $j = 1, \ldots, M$ and $M \leq J$. If the training set is not affected by numerical noise then exact interpolation of the training set is imposed and the coefficients w_j are computed by solving $\mathbf{A}\mathbf{w} = (\mathbf{f} - \text{EV}[\mathbf{f}])$, with $\mathbf{c}_j = \mathbf{y}_j$ (yielding $M = J$) and $\mathbf{f} = \{f(\mathbf{y}_i)\}_{i=1}^{J}$. If numerical noise affects the training set then noise reduction is achieved by choosing a number of RBF centers M smaller than the number of training points J, and $\mathbf{c}_j$ coordinates are defined via k-means clustering [15] of the training points. Hence, w_j are determined with least squares regression by solving $\mathbf{w} = (\mathbf{A}^\mathsf{T}\mathbf{A})^{-1}\mathbf{A}^\mathsf{T}(\mathbf{f} - \text{EV}[\mathbf{f}])$. The optimal number of stochastic RBF centers ($M^\star$) is defined by minimizing a leave-one-out cross-validation (LOOCV) metrics [4].

The uncertainty $U_{\widetilde{f}}(\mathbf{x})$ associated with the SRBF prediction is quantified by the 95%-confidence band using the cumulative density function of $g(\mathbf{x},\tau)$ [8].

3 INITIAL TRAINING SET AND BOUNDED SURROGATE MODEL

In this section a new approach for defining the initial training set is proposed. The new method uses a reduced training set (RS), as opposed to the authors previous work where a full training set was used (FS). Both the RS and FS approaches can be used with single- or multi-fidelity methods. Table 1 summarizes the RS and FS approaches for the single- ($N = 1$) and multi-fidelity ($N > 1$) cases.

The challenge for the RS approach is to create a SRBF surrogate model which can handle extrapolation. Using the RS approach, at the first iteration of the adaptive sampling the surro-

Table 1: Comparison between the new reduced training set (RS) and the full training set (FS) for the single- ($N = 1$) and multi-fidelity ($N > 1$) cases

Approach	N	Fidelity level	n. of training points	Training points placement
RS	1	1	1	Center of domain
	> 1	from 1 to $N-1$	1	Center of domain
		N	2D+1	Center of domain and of the boundaries
FS	1	1	2D+1	Center of domain and of the boundaries
	> 1	from 1 to $N-1$	2D+1	Center of domain and of the boundaries
		N	2D+1	Center of domain and of the boundaries

gate model prediction is an extrapolation based on the single training point available. Since the SRBF with power kernel has a low accuracy when extrapolating, a bounded surrogate model prediction and the associated uncertainty (both identified with the B subscript) are defined as described in Algorithm 1. When one training point is available, the surrogate model prediction and the associated uncertainty are set equal to the function value in the training point. When more training points are available, the surrogate model prediction and the associated uncertainty are bounded only in regions of the domain far from these training points. Specifically, the definition of $U_{\tilde{\varepsilon}_{Bi}}(\mathbf{x})$, as defined in Algorithm 1, stems from the consideration that the error surrogates represent errors in the multi-fidelity approximation $\hat{f}$. Therefore the average error can be used as reference for the surrogate model prediction uncertainty when an extrapolation is performed.

In Algorithm 1 a sigmoid-like function $s(r)$ is used to provide a smooth transition between the SRBF prediction and the bounded prediction

$$s(r) = \frac{1}{1 + e^{\alpha(r-\gamma)}}, \tag{6}$$

where, for the present work, $\alpha = -75$ and $\gamma = 0.2$. The argument r is the Euclidean distance of $\mathbf{x}$ from the existing k-th training point when one sample is available (J or $J_i = 1$). When more training points becomes available, the smallest hyperrectangle (whose edges are parallel to the Cartesian coordinated axis) containing the training points is defined and r becomes the distance of $\mathbf{x}$ from the hyperrectangle boundaries.

Figure 2 shows the first iteration of the RS and FS approaches for a monodimensional example. In the single fidelity case the RS approach yields a constant surrogate prediction and associated uncertainty, see Fig. 2a. In the multi-fidelity case, the RS approach produces a multi-fidelity prediction that only relies on the low-fidelity surrogate model to approximate the trend of the desired function. Differently, the FS approach takes advantage also of the medium- and high-fidelity evaluations. The RS approach produces an obviously less accurate prediction in the first iteration but allows to preserve a reasonable trend while using only one evaluation of the medium- and high-fidelity.

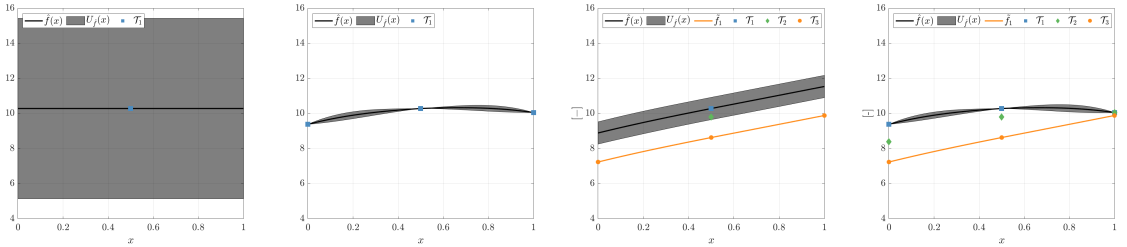
Algorithm 1: Bounding of the SRBF prediction and associated uncertainty

if $N = 1$ **then** // Single-fidelity case
  if $J = 1$ **then** // One training point available
    $\tilde{f}_B(\mathbf{x}) = f(\mathbf{x}')$;
    $U_{\tilde{f}_B}(\mathbf{x}) = f(\mathbf{x}')$;
  else if $J > 1$ **then** // J training points available
    $\tilde{f}_B(\mathbf{x}) = \tilde{f}(\mathbf{x})\,[1 - s(r)] + \mathrm{EV}[\mathbf{f}]s(r)$;
    $U_{\tilde{f}_B}(\mathbf{x}) = \min(U_{\tilde{f}}, 2\mathrm{EV}[\mathbf{f}])$;
  end
else if $N > 1$ **then** // Multi-fidelity case
  if $J_i = 1,\ i = 1, \ldots, N-1$ **then** // One training point available
    $\tilde{\varepsilon}_{Bi}(\mathbf{x}) = \varepsilon(\mathbf{x}'_i)$;
    $U_{\tilde{\varepsilon}_{Bi}}(\mathbf{x}) = \varepsilon(\mathbf{x}'_i)$;
  else if $J_i > 1,\ i = 1, \ldots, N-1$ **then** // J_i training points available
    $\tilde{\varepsilon}_{Bi}(\mathbf{x}) = \tilde{\varepsilon}_i(\mathbf{x})\,[1 - s_i(r)] + \mathrm{EV}[\varepsilon_i]s_i(r)$;
    $U_{\tilde{\varepsilon}_{Bi}}(\mathbf{x}) = \min(U_{\tilde{\varepsilon}_i}, 2\mathrm{EV}[\varepsilon_i])$;
  end
end

4 TEST PROBLEMS

4.1 Analytical Test Problem

The analytical test problem is based on the Rosenbrock function with two variables, three fidelities are considered ($N = 3$). The Rosenbrock function is a function defined using two parameters a and b. The high-fidelity function f_1 is a Rosenbrock function with $a = 1$ and $b = 100$. The medium-fidelity function f_2 is the additive combination of a Rosenbrock function, with $a = -2$ and $b = 50$, with a second order function, see Eq. 7. Finally, the low-fidelity



(a) RS, single fidelity (b) FS, single fidelity (c) RS, multiple fidelities (d) FS, multiple fidelities

Figure 2: Example of the two initialization strategies, monodimensional function from [9].

function f_3 is a transformation of f_1 by addition and multiplication of first order functions [10]

$$\begin{aligned} f_1(\mathbf{x}) &= 100(x_2 - x_1^2)^2 + (1 - x_1)^2, \\ f_2(\mathbf{x}) &= (-2 - x_1)^2 + 50(x_2 - x_1^2)^2 - 80 - 0.5x_1 0.5x_2, \\ f_3(\mathbf{x}) &= (f_1(x_1, x_2) - 4 - 0.5x_1 - 0.5x_2)/(10 + 0.25x_1 + 0.25x_2). \end{aligned} \tag{7}$$

where $\mathbf{x} \in [-2.048, 2.048] \otimes [-2.048, 2.048]$. The functions are represented in Fig. 3, they are challenging from the optimization viewpoint since the minimum is located in a narrow and flat valley. Specifically, the minimum of f_1 is in $\mathbf{x} = \{1, 1\}$ and equal to 0.

A synthetic computational cost is associated to the evaluation of each fidelity, the resulting computational cost ratios are equal to $\beta = \{1, 0.1, 0.05\}$.

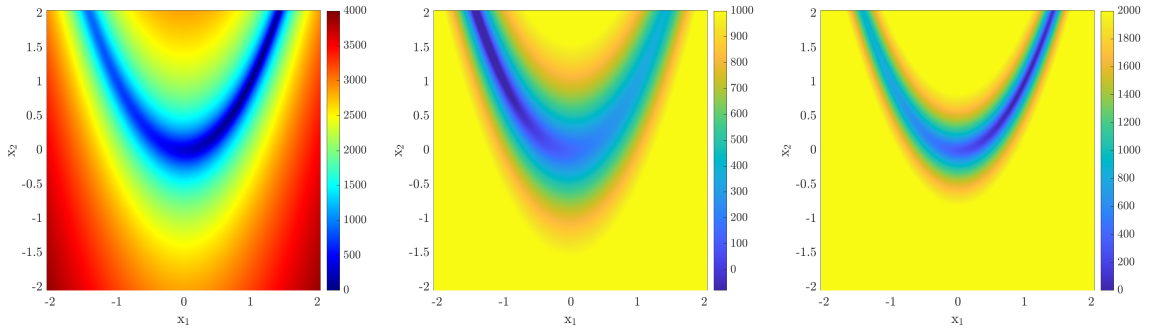
4.2 NACA Hydrofoil Optimization

This problem addresses the drag coefficient minimization of a NACA four-digit airfoil. The following minimization problem is solved

$$\text{minimize } f(\mathbf{x}) = C_D(\mathbf{x}), \text{ subject to } C_L(\mathbf{x}) = 0.6, \text{ and to } \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}. \tag{8}$$

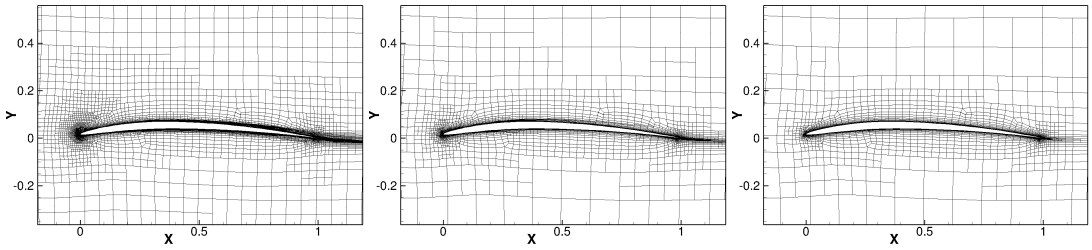
where $\mathbf{x}$ is the design variable vector; $\mathbf{l}$ and $\mathbf{u}$ are respectively the lower and upper bounds of the design space; C_D and C_L are respectively the drag and lift coefficient. The equality constraint on the lift coefficient is necessary in order to compare different geometries at the same lift force (equal to the weight of the object), since the drag depends strongly on the lift. The simulation conditions are: velocity $U = 10$ m/s, chord $c = 1$ m, fluid density $\rho = 1,026$ kg/m^3, with a chord based Reynolds number $\mathrm{Re} = 8.41 \cdot 10^6$. The hydrofoil shape is defined by the general equation for four-digit NACA foils [11]. In this work the design variables vector is defined as $\mathbf{x} = \{t, m\}$ with $t \in [0.030, 0.120]$ is the maximum thickness and $m \in [0.025, 0.065]$ is the maximum camber value with the maximum camber position fixed at $p = 0.4$. Numerical simulations are performed with the RANS solver ISIS-CFD. Tests are run with one and three fidelity levels ($N = 1, 3$), see Fig. 4.

ISIS-CFD is an incompressible unstructured finite-volume solver for multifluid flow. The velocity field is obtained from the momentum conservation equations and the pressure field is



(a) f_1 (b) f_2 (c) f_3

Figure 3: Test problem, the Rosenbrock function.



(a) (b) (c)

Figure 4: NACA hydrofoil computational grids for ISIS-CFD: (a) Fine grid, 12.8k cells, (b) Medium grid, 5.7k cells, and (c) Coarse grid, 3.6k cells.

extracted from the mass conservation constraint transformed into a pressure equation. These equations are similar to the Rhie and Chow SIMPLE method [12], but have been adapted for flows with discontinuous density fields. The unstructured discretization is face-based. While all unknown state variables are cell-centered, the systems of equations used in the implicit time stepping procedure are constructed face by face. Therefore, cells with an arbitrary number of arbitrarily-shaped constitutive faces are accepted. The code is fully parallel using the message passing interface protocol. A detailed description of the solver is given in [7].

Computational grids are created through adaptive grid refinement (AGR) [13, 14], to optimize the efficiency of the solver and to simplify the automatic creation of suitable grids. The AGR method adjusts the computational grid locally, during the computation, by dividing the cells of an original coarse grid. The decision where to refine comes from a refinement criterion, a tensor field $\mathcal{C}(x, y, z)$ computed from the flow. The tensor is based on second derivatives of pressure and velocity, which gives a crude indication of the local truncation errors. The grid is refined until the dimensions $\mathbf{d}_{p,j}$ $(j = 1, 2, 3)$ of each hexahedral cell p satisfy $\|\mathcal{C}_p \mathbf{d}_{p,j}\| = T_r$. The refinement criterion based on the second derivatives of the flow is not very sensitive to grid refinement [14], so the cell sizes everywhere are proportional to the constant threshold T_r.

For the MF optimization, grid adaptation is used to take into account the need for several fidelities. The interest of this procedure is that different fidelity results can be obtained by running the same simulations and simply changing the threshold T_r. Thus, it is straightforward to automate the MF simulations. Highest- to lowest-fidelity simulations require about 17, 9, and 5 minutes, respectively, of wall-clock time to converge. The resulting computational cost ratios are equal to $\beta = \{1, 0.5, 0.3\}$.

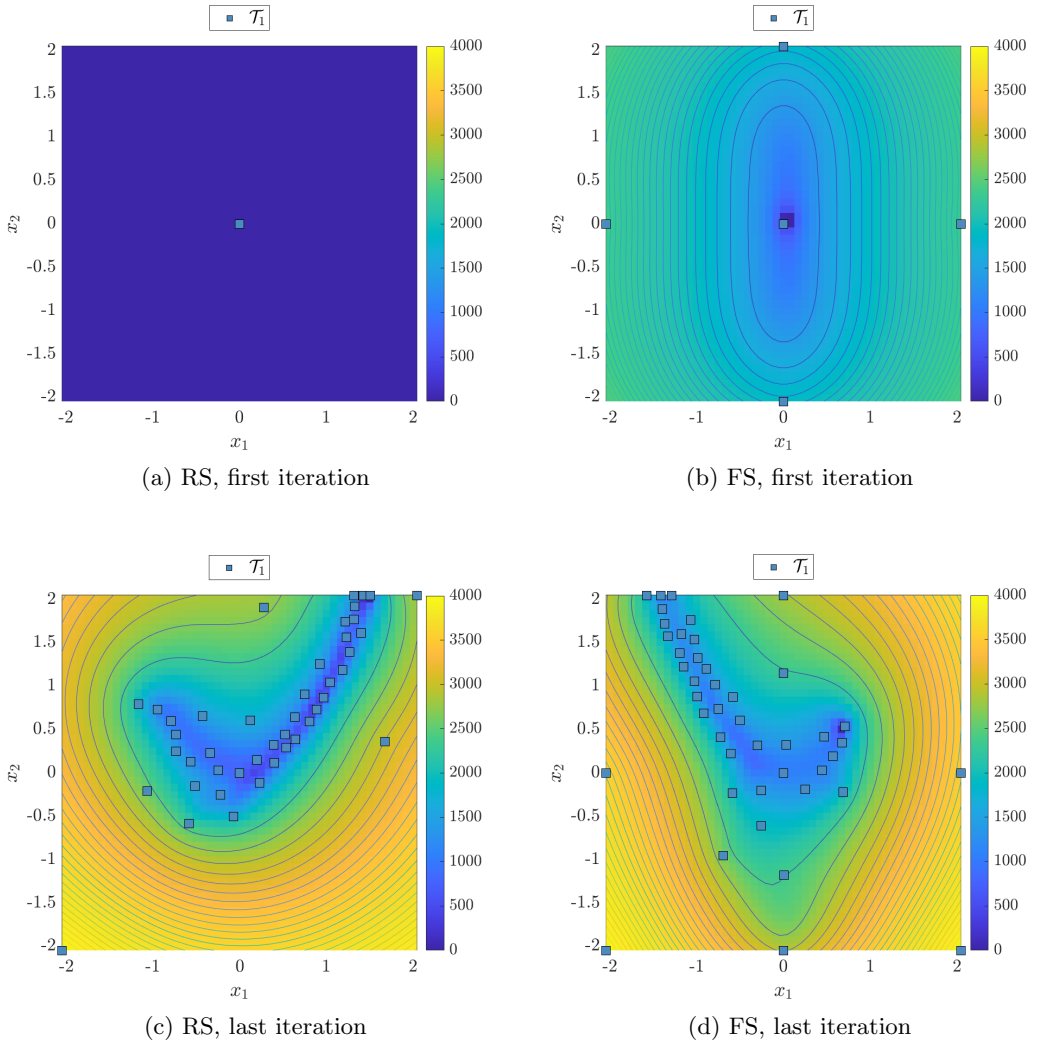
5 NUMERICAL RESULTS

The performance of the method is assessed using three metrics. Specifically, the accuracy of the surrogate model in the identification of the predicted minimum $\hat{f}(\mathbf{x}_{\min})$ is quantified by the prediction error (E_p); the error in the identification of the reference minimum $f(\check{\mathbf{x}})$ in the function space is quantified by the validation error (E_v); and the effectiveness of the surrogate model in identifying the location of the minimum in the variable space is quantified by the

location error (E_x)

$$E_p = \left| \frac{\hat{f}(\mathbf{x}_{\min}) - f_1(\mathbf{x}_{\min})}{R_1} \right|, \quad E_v = \left| \frac{f_1(\mathbf{x}_{\min}) - f(\check{\mathbf{x}})}{R_1} \right|, \quad E_x = \sqrt{\sum_{j=1}^{D} \left(\frac{x_{\min,j} - \check{x}_j}{u_j - l_j} \right)^2}, \tag{9}$$

where $f_1(\mathbf{x}_{\min})$ is the verified minimum by a high-fidelity evaluation, R_1 is the reference high-fidelity function range, $\check{\mathbf{x}}$ if the position of the reference minimum, and l_j and u_j (for $j = 1, \ldots, D$) are the lower and the upper bounds of the variables domain, respectively. The performance of the RS approach are compared with the performance of the FS approach at fixed computational budget equivalent to 45 high-fidelity simulations, for both problems.



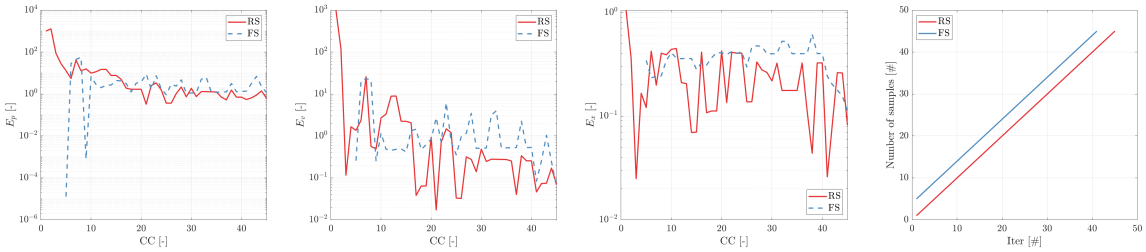
(a) RS, first iteration

(b) FS, first iteration

(c) RS, last iteration

(d) FS, last iteration

Figure 5: Analytical problem. Results for the single fidelity case, response surfaces and training sets for the two initialization approaches at the first and last iterations of the adaptive sampling.



(a) E_p convergence (b) E_v convergence (c) E_x convergence (d) Training set size

Figure 6: Analytical problem. Results for the single fidelity case.

5.1 Analytical Test Problem

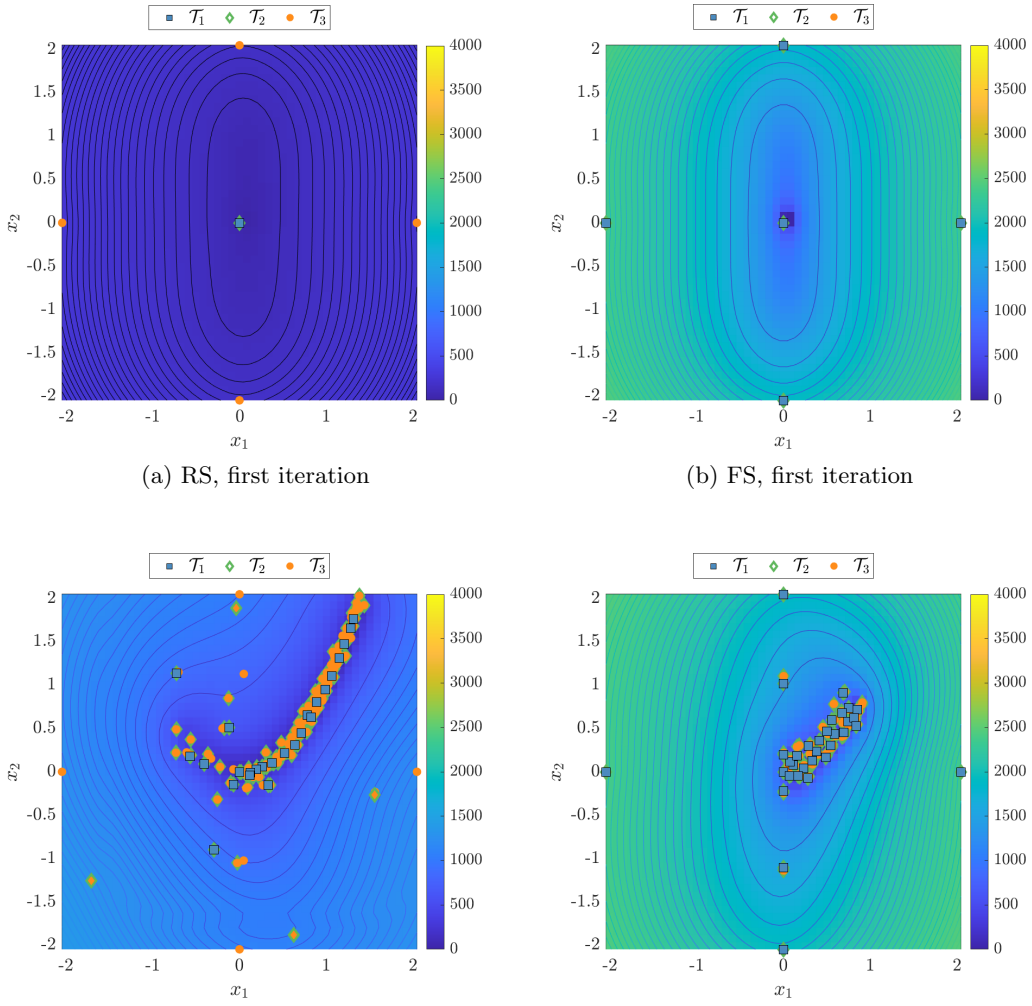
The evaluation of the analytical test is not affected from numerical noise, therefore an interpolative formulation of the SRBF is used. The computational costs associated with the initial RS training sets are 1 for $N = 1$ and 1.35 for $N = 3$, while for FS the costs are 5 and 5.75, respectively. The R_1 value is the high-fidelity function range of the FS initial training set and is equal to 3906.

Figure 5 shows the response surfaces with the training sets for the two initialization approaches using one fidelity. The response surface of the RS method is a constant flat surface whereas for the FS approach a concavity is already evident, see Figs. 5a and 5b. At the last iteration it is interesting to notice that the ACAS method with RS has explored the entire valley where the minimum is located whereas with FS the ACAS method has partially explored the valley and has then focused on the opposite side, see Figs. 5c and 5d.

Figure 6 shows the convergence of E_p, E_v, and E_x and the training set size for the two initialization approaches, using one fidelity. Although several oscillations are present, the RS approach achieves slightly better results than FS. Figure 6d shows that the ACAS method with RS performs more sampling iterations than with FS.

Figure 7 shows the response surfaces with the training sets for the two initialization approaches using three fidelities. At the first iteration the surrogate model prediction with the RS approach shows the same trend as with the FS approach. At the last iteration the ACAS method with RS has explored the entire valley where the minimum is located. Differently, the ACAS method with FS has partially explored the valley and did not perform any exploration of the domain, see Figs. 7c and 7d.

Figure 8 shows the convergence of E_p, E_v, and E_x and the training set size for the two initialization approaches, using three fidelities. While several oscillations are present, the RS approach achieves better results than FS. Figure 8d shows that the ACAS method with RS performs a significantly higher number of iterations than with the FS approach. This is due to the smaller computational cost spent for the evaluation of the initial training set with the RS approach. Furthermore, the ACAS method with RS uses a significantly lower number of high-fidelity evaluations. It is interesting to notice that the FS approach uses the same number of medium- and low-fidelity evaluations. This does not occur with the RS approach that uses less medium-fidelity evaluations than FS. Finally, the ACAS method with RS performs a high-fidelity



(a) RS, first iteration

(b) FS, first iteration

(c) RS, last iteration

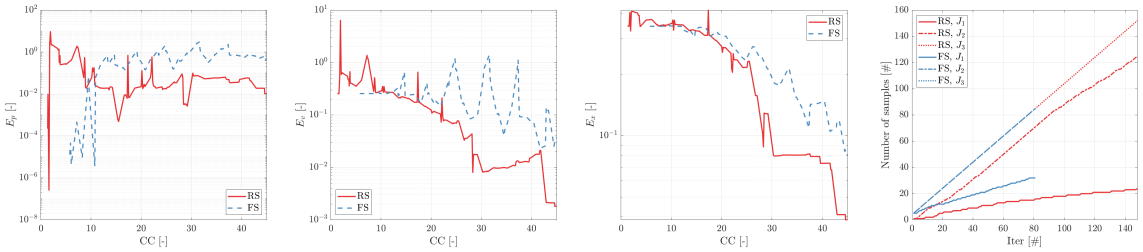
(d) FS, last iteration

Figure 7: Analytical problem. Results for the three fidelity case, response surfaces and training sets for the two initialization approaches at the first and last iterations of the adaptive sampling.

evaluation after the 8-th iteration, showing that a partial exploration is initially performed using only low- and medium-fidelity evaluations.

5.2 NACA Hydrofoil Optimization

The data pertaining to the NACA hydrofoil optimization with the ACAS method and the FS approach are take from [4]. The computational costs associated with the initial RS training sets are 1 for $N = 1$ and 3.03 for $N = 3$, while for FS the costs are 5 and 9.2, respectively. The R_1 value is the high-fidelity function range of the FS initial training set and is equal to $1.523E-3$. It is known from [4] that this problem is affected by numerical noise, especially the low-fidelity



(a) E_p convergence (b) E_v convergence (c) E_x convergence (d) Training sets size

Figure 8: Analytical problem. Results for the three fidelity case.

evaluations, therefore a least-squares regressive approach, as described in Section 2.2, is used.

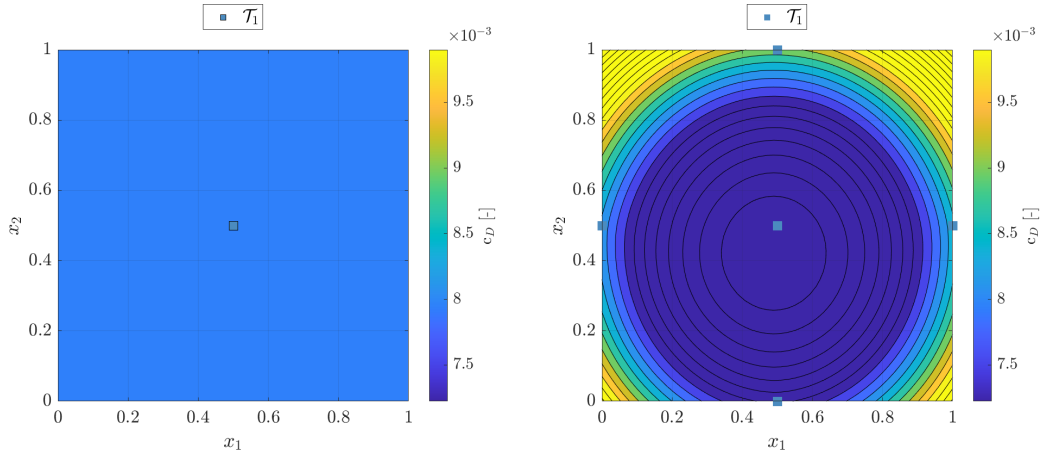
Figure 9 shows the response surfaces with the training sets for the two initialization approaches, using one fidelity. The response surface of the RS method is a constant flat surface whereas for the FS approach a concavity is already evident, see Figs. 9a and 9b. Figure 9c and 9d show the last iteration of the adaptive sampling, the RS approach has clustered the training points in the region of the minimum with only one training point in the opposite side of the domain (in $\{1, 1\}$), whereas the FS approach has performed also a considerable exploration of the design space.

Figure 10 shows the response surfaces with the training sets for the two initialization approaches, using three fidelities. The two response surfaces at the first iteration are significantly different, see Figs. 10a and 10b. Figures 10c and 10d show the last iteration of the adaptive sampling method. The ACAS method with the RS approach did not sampled the upper region of the domain, where the drag coefficient is higher. Differently, the ACAS method with the FS approach sampled also the upper corners of the domain. The response surface is mostly different but for the region of the minimum, correctly identified from both approaches. It is worth noting that the RS approach only adds one high-fidelity simulation to the initial training set, evaluating it in the center of the region of the minimum. Differently, the FS approach adds two high-fidelity simulations to the initial training set, evaluating them in the two corners of the lower region of the domain. Table 2 summarizes the minimum position, value, and associated uncertainty and the E_p and E_v metrics along with the training sets size. The RS approach identifies different minima. The predicted minimum is significantly different among the approaches whereas all the verified minima show an error, with respect to the reference, below 1%. The RS approach shows with three fidelities shows the largest prediction error due to overfitting of the low-fidelity training data, see Fig. 10c, which is the most affected from the numerical noise. Finally, the RS approach uses the lowest number of high-fidelity evaluations.

6 CONCLUSIONS

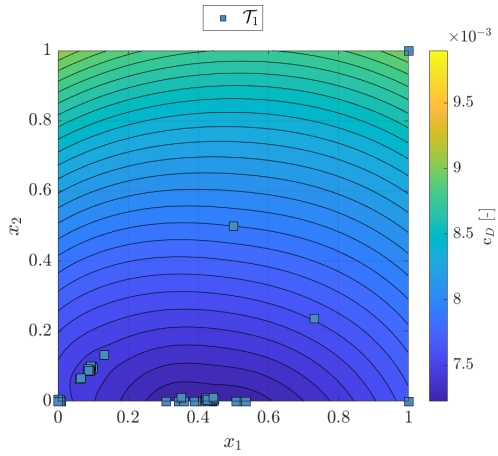
A new approach is defined to reduce the computational cost of the training set initialization of a surrogate-based optimization method, for both single- and multi-fidelity (MF). The proposed approach is specifically developed for stochastic radial basis functions with power kernel (SRBF). Reducing the computational cost of the initialization means that a larger computational budget is available for an adaptive sampling of the design space. Furthermore, less information are

initially provided to the surrogate model in order to have it freely adapt to the data. The computational cost of the initialization is reduced in comparison with the initialization approach used in the authors previous work. The previous initialization approach uses a full training set (FS) of $2D + 1$ samples, where D is the design space dimension, for all the fidelities. In the single-fidelity case, the reduced training set (RS) uses one point as initial training set. In the multi-fidelity case, the RS approach uses $2D + 1$ points as training set only for the lowest (and cheapest to evaluate) fidelity and one single point as initial training set for the other fidelities. When a single training point is available, the surrogate model prediction is an extrapolation

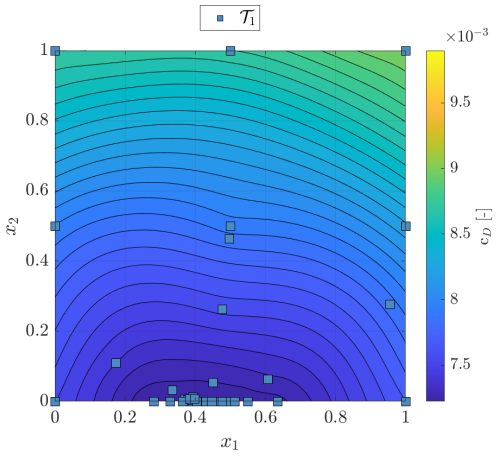


(a) RS, first iteration

(b) FS, first iteration

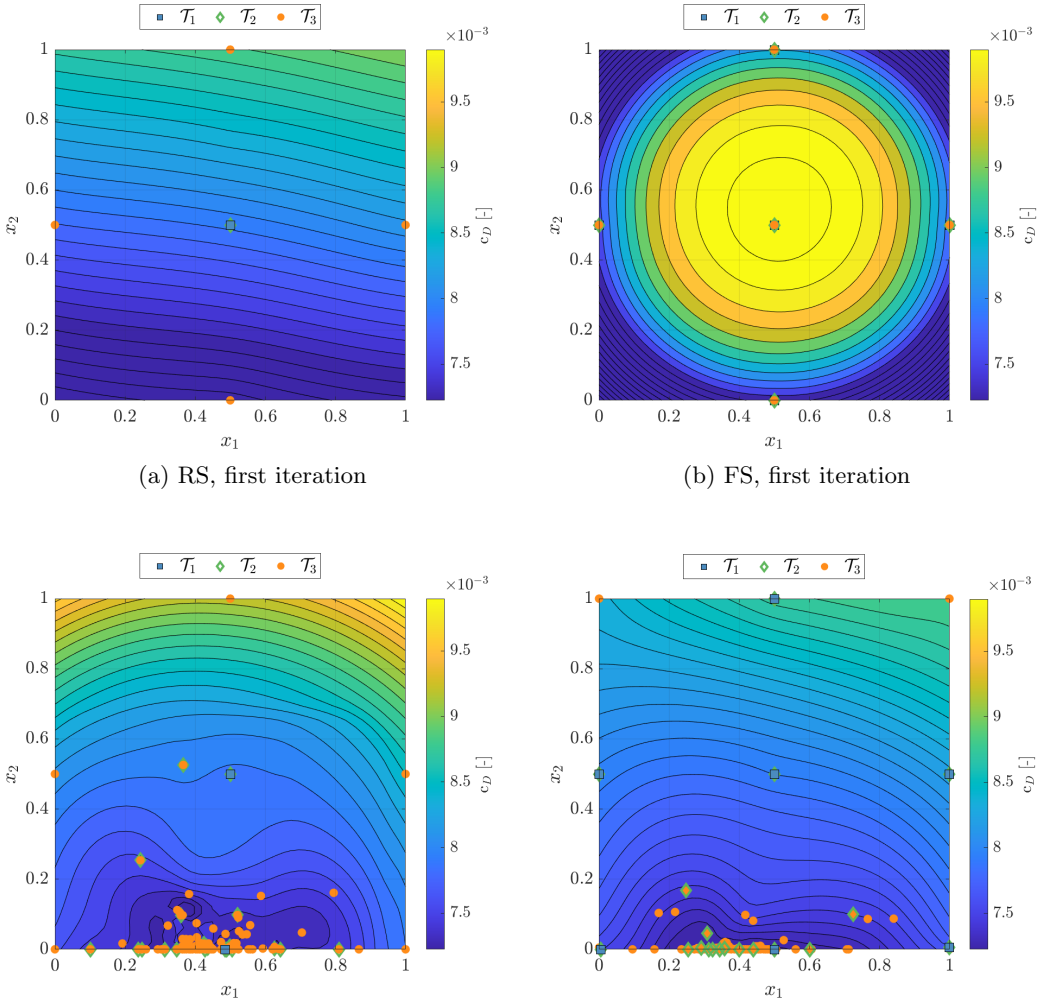


(c) RS, last iteration



(d) FS, last iteration

Figure 9: NACA hydrofoil optimization problem. Results for the single fidelity case, response surfaces and training sets for the two initialization approaches at the first and last iterations of the adaptive sampling.



(a) RS, first iteration

(b) FS, first iteration

(c) RS, last iteration

(d) FS, last iteration

Figure 10: NACA hydrofoil optimization problem. Results for the three fidelity case, response surfaces and training sets for the two initialization approaches at the first and last iterations of the adaptive sampling.

based on that single training point. The power kernel used in the SRBF lacks a compact support, as a consequence the extrapolated prediction may not be well correlated with the desired function behavior, negatively affecting the adaptive sampling. Although other kernels exist with compact support, the SRBF with power kernels is robust, showing good results for several applications and are here preferred. Therefore, a constraint is imposed on both the surrogate model prediction and the associated uncertainty when an extrapolation is performed, to improve the adaptive sampling. New training points are iteratively added by and adaptive sampling method based on the surrogate prediction value, the associated uncertainty, and the computational cost associated to each fidelity.

The proposed initialization approach (RS) is assessed and compared to the previous (FS) for an analytical test problem and a NACA hydrofoil optimization problem, both with one and three fidelities. The comparison is provided for a fixed computational budget. A synthetic computational cost is associated to each evaluation of the analytical test. Finally, the comparison is based on three metrics to quantify the capability of the surrogated-based optimization to converge towards the global optimum. The RS approach reduced the computational cost of evaluating the initial training set by 76% for the analytical test problem and by 67% for the NACA hydrofoil optimization.

The results of the analytical test problem show that the adaptive sampling method with the RS approach uses less high-fidelity evaluations, especially in the early iterations, relying more on low- and medium-fidelity evaluations for the identification of the region of interest. This is due to the reduced data provided to the adaptive sampling method and to the bounding of the uncertainty, that make the method not interested in exploring the extrema of the domain with high-fidelity evaluations. For the NACA hydrofoil optimization problem, the adaptive sampling method adds only one high-fidelity sample in the region of the minimum. Although this means that the method is effective in using high-fidelity data only where they are most useful, this also means that the flat region of minimum of the NACA problem is not completely explored. This is most likely due to the numerical noise in the low-fidelity training set that is affecting the adaptive sampling procedure.

Future work aims to investigate the effects of the RS approach on the adaptive sampling using several analytical test problems. Specifically, the analytical test set defined within the NATO STO Research Task Group AVT-331 on "Goal-driven, multi-fidelity approaches for military vehicle system-level design" will be considered [17]. Furthermore, an improved strategy to perform regression with the SRBF in presence of numerical noise will be investigated in order to avoid overfitting.

ACKNOWLEDGMENTS

CNR-INM is grateful to Drs. Elena McCarthy and Woei-Min Lin of the Office of Naval Research for their support through the Naval International Cooperative Opportunities in Science and Technology Program. The work is conducted in collaboration with NATO STO Research Task Group AVT-331 on "Goal-driven, multi-fidelity approaches for military vehicle system-level design".

Table 2: NACA hydrofoil optimization problem, summary of the results

Initial set	N	x_1	x_2	$\hat{f}_{\min}$	$f_{\min}$	$\lvert E_p\rvert\%$	$\lvert E_v\rvert\%$	$\mathcal{J}_1$	$\mathcal{J}_2$	$\mathcal{J}_3$
Reduced	1	0.4173	0.0000	7.2132E-3	7.2195E-3	0.41	0.52	45	-	-
Full	1	0.3799	0.0000	7.2130E-3	7.2140E-3	0.01	0.03	45	-	-
Reduced	3	0.4354	0.0000	7.0500E-3	7.2245E-3	11.45	0.85	2	22	105
Full	3	0.3615	0.0000	7.1616E-3	7.2182E-3	0.78	0.09	7	19	96
Reference	1	0.3776	0.0000	-	7.2116E-3	-	-	-	-	-

REFERENCES

[1] Peherstorfer, Benjamin, Karen Willcox, and Max Gunzburger. Survey of multifidelity methods in uncertainty propagation, inference, and optimization, *Siam Review* 60, **3**, 2018, 550-591.

[2] Giselle Fernández-Godino, M., Chanyoung Park, Nam H. Kim, and Raphael T. Haftka, Issues in deciding whether to use multifidelity surrogates, *AIAA Journal* 57, **5**, 2019, 2039-2054.

[3] Serani, A., Pellegrini, R., Wackers, J., Jeanson, C.-E., Queutey, P., Visonneau, M., and Diez, M., Adaptive multi-fidelity sampling for CFD-based optimisation via radial basis function metamodels, *International Journal of Computational Fluid Dynamics*, **33**(6-7), 2019, pp. 237–255.

[4] Wackers, J., Visonneau, M., Serani, A., Pellegrini, R., Broglia, R., Diez, M., 2020. Multi-Fidelity Machine Learning from Adaptive- and Multi-Grid RANS Simulations. *33rd Symposium on Naval Hydrodynamics*, Osaka, Japan.

[5] Mackman, T. J., C. B. Allen, Mehdi Ghoreyshi, and K. J. Badcock. Comparison of adaptive sampling methods for generation of surrogate aerodynamic models, *AIAA journal* 51, **4**, 2013, 797-808.

[6] Cai, Xiwen, Haobo Qiu, Liang Gao, Li Wei, and Xinyu Shao. Adaptive radial-basis-function-based multifidelity metamodeling for expensive black-box problems, *AIAA journal* 55, **7**, 2017, 2424-2436.

[7] Queutey, P. and Visonneau, M., An Interface Capturing Method for Free-Surface Hydrodynamic Flows, *Computers & Fluids*, **36**(9), 2007, pp. 1481–1510.

[8] Volpi, S., Diez, M., Gaul, N., Song, H., Iemma, U., Choi, K. K., Campana, E. F., and Stern, F., Development and validation of a dynamic metamodel based on stochastic radial basis functions and uncertainty quantification, *Structural and Multidisciplinary Optimization*,**51**(2), 2015, pp. 347–368.

[9] Wackers, Jeroen, Michel Visonneau, Simone Ficini, Riccardo Pellegrini, Andrea Serani, and Matteo Diez. Adaptive N-Fidelity Metamodels for Noisy CFD Data. *AIAA AVIATION 2020 FORUM*, p. 3161. 2020.

[10] Rumpfkeil, M. P., and Beran, P. S., Multi-Fidelity, Gradient-enhanced, and Locally Optimized Sparse Polynomial Chaos and Kriging Surrogate Models Applied to Test Problems, *AIAA Scitech 2020 Forum*, 2020, p. 0677.

[11] Moran, J. An introduction to theoretical and computational aerodynamics. *Courier Corporation*, 2003.

[12] Rhie, C. M. and Chow, W. L., A Numerical Study of the Turbulent Flow Past an Isolated Airfoil With Trailing Edge Separation, *AIAA Journal*, **17**, 1983, pp. 1525–1532.

[13] Wackers, J., Deng, G., Guilmineau, E., Leroyer, A., Queutey, P., and Visonneau, M., Combined refinement criteria for anisotropic grid refinement in free-surface flow simulation, *Computers and Fluids*, **92**, 2014, pp. 209 – 222.

[14] Wackers, J., Deng, G., Guilmineau, E., Leroyer, A., Queutey, P., Visonneau, M., Palmieri, A., and Liverani, A., Can adaptive grid refinement produce grid-independent solutions for incompressible flows?, *Journal of Computational Physics*, **344**, 2017, pp. 364 – 380.

[15] Lloyd, S., Least squares quantization in PCM, *IEEE transactions on information theory*, **28**(2), 1982, pp. 129–137.

[16] Cox, Dennis D. and John, Susan, A statistical method for global optimization, *In proceedings of 1992 IEEE International Conference on Systems, Man, and Cybernetics*, 1241–1246, 1992

[17] Beran, Philip S and Bryson, Dean and Thelen, Andrew S and Diez, Matteo and Serani, Andrea, Comparison of multi-fidelity approaches for military vehicle design, *In proceedings of AIAA AVIATION 2020 FORUM*, 2020