



**HAL**  
open science

## Hydrofoil Optimization via Automated Multi-Fidelity Surrogate Models

Hayriye Pehlivan Solak, Jeroen Wackers, Riccardo Pellegrini, Andrea Serani,  
Matteo Diez, Paolo Perali, Matthieu Sacher, Jean-Baptiste Leroux, Benoit  
Augier, Frédéric Hauville, et al.

► **To cite this version:**

Hayriye Pehlivan Solak, Jeroen Wackers, Riccardo Pellegrini, Andrea Serani, Matteo Diez, et al..  
Hydrofoil Optimization via Automated Multi-Fidelity Surrogate Models. 10th International Confer-  
ence on Computational Methods in Marine Engineering (MARINE 2023), Jun 2023, Madrid, Spain.  
hal-04089604

**HAL Id: hal-04089604**

**<https://hal.science/hal-04089604v1>**

Submitted on 5 May 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.

# Hydrofoil Optimization via Automated Multi-Fidelity Surrogate Models

Hayriye Pehlivan Solak<sup>1\*</sup>, Jeroen Wackers<sup>1</sup>, Riccardo Pellegrini<sup>†</sup>, Andrea Serani<sup>†</sup>,  
Matteo Diez<sup>†</sup>, Paolo Perali<sup>‡</sup>, Matthieu Sacher<sup>‡</sup>, Jean-Baptiste Leroux<sup>‡</sup>, Benoît  
Augier<sup>§</sup>, Frédéric Hauville<sup>||</sup> and Patrick Bot<sup>||</sup>

<sup>1</sup> LHEEA, CNRS UMR 6598, Centrale Nantes, Nantes, France

<sup>†</sup> CNR-INM, National Research Council-Institute of Marine Engineering, Rome, Italy

<sup>‡</sup> ENSTA Bretagne, CNRS UMR 6027, IRDL, Brest, France

<sup>§</sup> IFREMER, RDT, Plouzané, France

<sup>||</sup> IRENav, Ecole Navale, Brest, France

\* Corresponding author: H. Pehlivan Solak, hayriye.pehlivan@ec-nantes.fr

## ABSTRACT

Lifting hydrofoils are gaining importance, since they drastically reduce the wetted surface area of a ship hull, thus decreasing resistance. To attain efficient hydrofoils, the geometries can be obtained from an automated optimization process, based on simulations. However, hydrofoil high-fidelity simulations are computationally demanding, since fine meshes are needed to accurately capture the pressure field and the boundary layer on the hydrofoil. Simulation-based optimization can therefore be very expensive.

Automated surrogate models, trained by a limited number of simulations, can reduce the required computational demand for the optimization by performing simulations where these are more informative. Furthermore, if an efficient low-fidelity hydrofoil performance prediction tool (with a low computational cost) is available, using surrogates in a multi-fidelity framework can provide a further reduction in required simulations, by combining the accuracy of a few high-fidelity with an exploration process based on a larger number of low-fidelity computations.

In this study, we propose a hydrofoil optimization procedure based on two simulation codes, a dedicated hydrofoil potential flow solver for low-fidelity simulations and a RANS solver for both medium- and high-fidelity simulations. The RANS solver uses adaptive grid refinement to attain high accuracy with a limited computational budget. Two different multi-fidelity frameworks are compared for a parameterized geometric model of a realistic hydrofoil: only RANS based and potential-RANS based. The effect of different combinations of fidelity levels on the efficiency of the optimization and the performance of the kitefoil-type hydrofoil is investigated and discussed.

**Keywords:** SDDO; multi-fidelity; RANS; potential solver; kitefoil.

## 1. INTRODUCTION

While numerical simulation is becoming widely used as a design tool in naval hydrodynamics, automatic Simulation-Driven Design Optimization (SDDO) has so far seen limited practical use. Even with today’s technology and computational cost reduction methods, a recent study showed that under realistic stochastic conditions, hull form optimization may require up to 500k CPU hours (Serani et al., 2022). For the typical time and budget available in ship design, this is prohibitively expensive.

Surrogate models, which replace time-consuming numerical computations for the optimization by a model of the simulation response based on a limited amount of simulation results, may overcome this issue. Thus, surrogate models have become widely accepted in design optimization and several prominent models exist, such as radial basis functions (Volpi et al., 2015) or Gaussian processes (Ortali et al., 2022). Furthermore, multi-fidelity strategies are especially efficient since they combine the exploration ability of numerous low-fidelity numerical simulations with the accuracy of high-fidelity computations (Beran et al., 2020). Low- and high-fidelity simulations may be based for example on different physical models (Nagawkar and Leifsson, 2022), or mesh resolution (Wackers et al., 2022b). Additive and multiplicative correction methods (Kennedy and O’Hagan, 2000) are used to combine these different simulations into multi-fidelity surrogate models.

Nevertheless, especially for complex geometries such as lifting hydrofoils, accurate optimization using high-fidelity solvers remains challenging. Hydrofoil shape optimization is important for naval architects, since foils can significantly increase a vessel’s energy efficiency or performance in yacht races. However, hydrofoils are expensive to simulate since they require fine meshes to capture the flow characteristics, such as the pressure peaks near the leading and trailing edges. Also the free surface needs to be resolved accurately, since it has a great effect on flow field (Pernod et al., 2023). The cost of the optimization further increases because the characteristics of a hydrofoil depend on many geometrical parameters, hence the optimization is conducted in a high-dimensional space. Thus, to be usable in practice, the optimization procedure must be highly efficient.

The goal of this paper is to study if high-fidelity hydrofoil shape optimization can be conducted today in a realistic time for naval engineering, which is about two weeks of computation on a modern workstation or a small cluster. We use two separate approaches for increasing the optimization efficiency.

First, optimal use of computational resources can be achieved by automatic adaptation of a method with respect to the requirements of the task it accomplishes. For optimization, adaptivity is used for i) active learning, i.e. adaptive sampling of the surrogate training data, and ii) producing training data for the surrogate, i.e. adaptivity in the flow solvers. Regarding the surrogate model, adaptive selection of the simulations which form its training data provides more accurate and efficient predictions, by sampling the most promising locations using knowledge available from the surrogate, such as the uncertainty and the possible minimum of the objective value (Serani et al., 2019). In a multi-fidelity framework, active learning also assigns the fidelity level to be sampled in search of the next design candidate (Wackers et al., 2022c).

For simulations, automatic mesh adaptation can lower the cost of the predictions by placing fine grid cells only where they are needed to resolve the flow. In this study, the Reynolds-averaged Navier-Stokes (RANS) flow solver uses adaptive grid refinement and overset meshes for hydrofoils.

Second, the multi-fidelity paradigm provides even larger benefits if: (a) the lowest fidelity simulations are very fast, so they can efficiently explore even large design spaces, while expensive high-fidelity points are placed only around the optimum (as an example, using potential flow solvers as low-fidelity

models can provide a significant reduction of the computational cost of optimization). And (b) the more the different fidelity levels are in correlation, the better the efficiency of the optimization process, which means that specialized potential flow solvers for each application are attractive. In this study, a potential flow solver dedicated to hydrofoils (Perali et al., 2022) featuring a nonlinear free surface condition and a dynamically resolved wake, will be combined with high-fidelity RANS to analyse the cost savings this can bring to a multi-fidelity framework.

To perform these tests, we set up an automated multi-fidelity hydrofoil optimization framework and compare two different optimization approaches:

1. RANS simulation for both low- and high(er)-fidelity levels, using the full design space,
2. Potential flow simulation as the lowest fidelity with RANS for high(er)-fidelity levels,

The optimization is performed for the main wing of a kitefoil, operating at a heel angle in close proximity to the water surface. The flow solvers used are the potential flow code PUFFIn (Perali et al., 2022) developed in the partnership of IRDL, IRENav, and IFREMER, and the RANS flow solver ISIS-CFD developed at LHEEA (Queutey and Visonneau, 2007; Wackers et al., 2017). The multi-fidelity surrogate modeling approach developed at CNR-INM (Pellegrini et al., 2022) is used.

The remainder of this paper is organised as follows: Sect. 2 presents the surrogate modeling with active learning strategy. Section 3 introduces the flow solvers PUFFIn and ISIS-CFD. Then Sect. 4 presents the optimization problem and geometric modeler tool we use for automated shape optimization of hydrofoils. This is followed by the simulation setup and tests for each flow solver (ect. 5) and optimization results (Sect. 6). Finally the conclusions and future directions are presented.

## 2. SURROGATE MODELING

All the optimizations in this paper are performed with the multi-fidelity optimization approach of Pellegrini et al. (2022): it is based on Stochastic Radial Basis Function surrogate models (SRBF) and uses an arbitrary number of low-fidelity surrogates, combined with additive corrections for the difference between fidelity levels, and several active learning methods. Furthermore, noise canceling methods are available. This section describes the approach as used in this paper.

### 2.1 Multi-Fidelity Framework

The response  $f(\mathbf{x})$  of the problem is modeled based on  $N$  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  $\{f_i\}_{i=2}^{N-1}(\mathbf{x})$ , where  $\mathbf{x} \in \mathbb{R}^D$  is the design variables vector of dimension  $D$ . The multi-fidelity (MF) prediction  $\hat{f}_i(\mathbf{x})$  of  $f_i(\mathbf{x})$  ( $i = 1, \dots, N-1$ ) is the sum of surrogates (denoted by  $\tilde{\cdot}$ ) for the lowest level and the inter-level errors

$$\hat{f}_i(\mathbf{x}) = \tilde{f}_N(\mathbf{x}) + \sum_{k=i}^{N-1} \tilde{\varepsilon}_k(\mathbf{x}). \quad (2.1)$$

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

$$\varepsilon_i(\mathbf{x}_j) = f_i(\mathbf{x}_j) - \hat{f}_{i+1}(\mathbf{x}_j). \quad (2.2)$$

A surrogate model is used which provides both a prediction and an associated uncertainty (see below). 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 ( $i = 1, \dots, N - 1$ )

$$U_{\hat{f}_i}(\mathbf{x}) = \sqrt{U_{\tilde{f}_N}^2(\mathbf{x}) + \sum_{k=i}^{N-1} U_{\tilde{\varepsilon}_k}^2(\mathbf{x})}. \quad (2.3)$$

## 2.2 Stochastic Radial Basis Functions with Least Squares Approximation

The individual surrogate models in equation (2.1) are constructed with the Stochastic Radial Basis Functions (SRBF) of Volpi et al. (2015). Given a (single-fidelity) training set  $\mathcal{T} = \{\mathbf{x}_i, f(\mathbf{x}_i)\}_{i=1}^J$ , the surrogate model prediction  $\tilde{f}(\mathbf{x})$  is computed as the expected value (EV) over a stochastic tuning parameter of the surrogate model,  $\tau \sim \text{unif}[1, 3]$

$$\tilde{f}(\mathbf{x}) = \text{EV}[g(\mathbf{x}, \tau)]_{\tau}, \quad \text{with} \quad (2.4)$$

$$g(\mathbf{x}, \tau) = \text{EV}[\mathbf{f}] + \sum_{j=1}^M w_j \|\mathbf{x} - \mathbf{c}_j\|^\tau, \quad (2.5)$$

where  $w_j$  are unknown coefficients,  $\|\cdot\|$  is the Euclidean norm, and  $\mathbf{c}_j$  are the RBF centers, with  $j = 1, \dots, M$  and  $M \leq J$ . The uncertainty  $U_{\tilde{f}}(\mathbf{x})$  associated with the SRBF prediction is quantified by the 95%-confidence band using the cumulative density function of  $g(\mathbf{x}, \tau)$ .

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 (Lloyd, 1982) of the training points. Hence,  $w_j$  are determined with least squares regression by solving  $\mathbf{w} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top (\mathbf{f} - \text{EV}[\mathbf{f}])$ . The optimal number of stochastic RBF centers ( $M^*$ ) is defined by minimizing a leave-one-out cross-validation (LOOCV) metric (Pellegrini et al., 2022).

## 2.3 Initial Training set and Bounded Surrogate Model

The reduced initial training set (RS) of Wackers et al. (2022c) is used: except on the lowest fidelity level, where the domain center and the centers of the boundary faces are sampled, the surrogate models are initialized with only a point in the domain center. For the error surrogate models, this requires a SRBF surrogate which can handle extrapolation. Therefore, a bounded surrogate model prediction and uncertainty (both identified with the  $B$  subscript) are defined as

$$\tilde{\varepsilon}_{Bi}(\mathbf{x}) = \tilde{\varepsilon}_i(\mathbf{x}) [1 - s_i(r)] + \text{EV}[\varepsilon_i] s_i(r), \quad (2.6)$$

$$U_{\tilde{\varepsilon}_{Bi}}(\mathbf{x}) = \min(U_{\tilde{\varepsilon}_i}, \text{EV}[\varepsilon_i]), \quad (2.7)$$

where a sigmoid-like function  $s(r)$  is used to provide a smooth transition between the SRBF prediction and the bounded prediction (Wackers et al., 2022c), with  $r$  the distance from the boundaries of the smallest hypercube containing the training set. Since error surrogates represent numerical errors in training data, the average error  $\text{EV}[\varepsilon_i]$  provides an acceptable reference value away from high-fidelity points when an extrapolation is needed.

## 2.4 Active Learning Method

The MF surrogate model is dynamically updated by adding new training points (Serani et al., 2019). First, a new training point  $\mathbf{x}^*$  is identified based on a modified version of the Lower Confidence Bound

(LCB) infill criterion, which samples points with large prediction uncertainty and small objective function value:

$$\mathbf{x}^* = \operatorname{argmin}_{\mathbf{x}} \left[ \widehat{f}(\mathbf{x}) - U_{\widehat{f}}(\mathbf{x}) + P_x(\mathbf{x}) \right]. \quad (2.8)$$

$P_x(\mathbf{x})$  is a penalization factor based on the distance from the existing training sets (considering all the fidelities) to prevent the sampling of already sampled points. Once  $\mathbf{x}^*$  is identified, the fidelity used for the evaluation of  $f(\mathbf{x}^*)$  is selected. 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  $N$ , where  $k$  is the fidelity level that has the highest uncertainty divided by the computational cost (Pellegrini *et al.*, 2022).

### 3. FLOW SOLVERS AND MULTI-FIDELITY FRAMEWORK

For multi-fidelity optimization, to obtain simulations with various fidelity levels, the present paper uses two flow solvers. The following subsections describe the potential flow solver PUFFIn (Perali *et al.*, 2022) followed by the RANS-based ISIS-CFD and its mesh adaptation method. Finally, the foil sock approach (Robin *et al.*, 2022) for automated kitefoil optimization is introduced.

#### 3.1 PUFFIn

An incompressible potential approach is used for the PUFFIn flow solver. Therefore, viscous effects are neglected and the flow is irrotational. The velocity field  $\mathbf{u}(x, y, z, t)$  is obtained as the superposition of the undisturbed flow  $\mathbf{U}_0$  and the gradient of the velocity potential  $\phi(x, y, z, t)$ :

$$\mathbf{u}(x, y, z, t) = \mathbf{U}_0 + \nabla\phi(x, y, z, t). \quad (3.1)$$

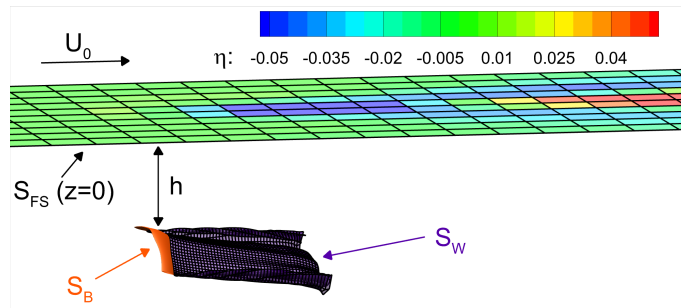
Under these circumstances, mass conservation reduces to a Laplace equation:

$$\nabla^2\phi(x, y, z, t) = 0. \quad (3.2)$$

A Boundary Integral Equation (BIE) is written on the domain boundaries  $S$  by imposing the second Green identity to Eq. (3.2). The potential is obtained by a sum of doublets  $\mu$  and sources  $\sigma$  on the boundary  $S$  (Katz and Plotkin, 2001):

$$\phi = -\frac{1}{4\pi} \int_S \left[ \sigma \frac{1}{r} - \mu \frac{\partial}{\partial n} \left( \frac{1}{r} \right) \right] dS, \quad \text{where } r = \|x - x_P\|, P \in S, \quad (3.3)$$

with  $\mathbf{n}$  the outward pointing normal vector and  $\frac{\partial}{\partial n} = \mathbf{n} \cdot \nabla$ . The typical simulation configuration for PUFFIn is represented in Fig. 1. The distributions of doublets  $\mu$  and sources  $\sigma$  on the boundary  $S$  is:



**Figure 1:** Configuration for PUFFIn computations.

$$\mu = -\phi(x, y, z, t) \quad \text{and} \quad \sigma = -\frac{\partial\phi(x, y, z, t)}{\partial n}. \quad (3.4)$$

On the surface of the geometry, the sources are given by the non-penetration condition:

$$\sigma = -\mathbf{U}_0 \cdot \mathbf{n}. \quad (3.5)$$

The source distribution on the wake surface must be null, so Eq. (3.3) can be written:

$$\phi(x, y, z, t) = \frac{1}{4\pi} \int_{S_B+S_{FS}} \left[ \sigma \frac{1}{r} - \mu \frac{\partial}{\partial n} \left( \frac{1}{r} \right) \right] dS + \frac{1}{4\pi} \int_{S_W} \left[ \mu \frac{\partial}{\partial n} \left( \frac{1}{r} \right) \right] dS. \quad (3.6)$$

In addition, linearized kinematic and dynamic boundary conditions are imposed around the initial free surface position  $z = 0$  to reduce the computational time. Keeping only the first order terms:

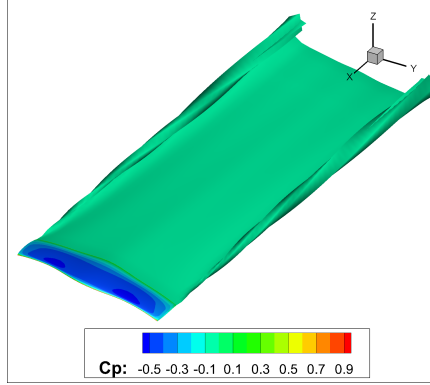
$$\frac{\partial\phi}{\partial t} + \mathbf{U}_0 \cdot \nabla\eta = \frac{\partial\phi}{\partial z}, \quad (3.7)$$

$$\frac{\partial\phi}{\partial t} + \mathbf{U}_0 \cdot \nabla\phi + g\eta = 0. \quad (3.8)$$

The Neumann-Kelvin (NK) formulation is obtained by combining these two equations:

$$\left( \frac{\partial}{\partial t} + \mathbf{U}_0 \cdot \nabla \right)^2 \phi + g \frac{\partial\phi}{\partial z} = 0. \quad (3.9)$$

The NK condition is solved using the finite difference method, with a third order upwind scheme for the spatial derivative and a second order backward scheme for the time derivative.



**Figure 2:** PUFFIn base geometry and wake.

Boundaries are discretized by quadrilateral panels and the source and doublet distributions are supposed to be constant on each panel. The matricial form of the Boundary Element Method (BEM) can be obtained by writing Eq. (3.6) at the geometrical center of each panel. The doublet distribution is obtained by solving:

$$\mathcal{A}\mu = \mathcal{B}\sigma_b + \mathcal{W}\mu_w, \quad (3.10)$$

where  $\sigma_b$  are the source strengths on the hydrofoil panels and  $\mu_w$  are the doublet strengths on the wake panels. While the sources  $\sigma_b$  are given by Eq. (3.5), the position of the wake and the doublet distribution  $\mu_w$  are still unknown. The wake surface is seen in Fig. 2 constructed using a Lagrangian approach. At each time step, new panels are created downstream of the trailing edge. The doublet values on these new panels are obtained with a Kutta condition at the trailing edge, imposing equality

of the pressure on the lower and upper part of the foil. For inviscid flow, the doublet strength on a wake panel must remain constant between two time steps. Thus, the panels are subsequently transported by the flow using an explicit first order time scheme and keeping their initial values of the doublet.

The pressure distribution on the hydrofoil is obtained with the Bernoulli relation based on the velocity potential. Integration of the pressure over the surface  $S_B$  gives the hydrodynamic forces and moments acting on the hydrofoil.

### 3.2 ISIS-CFD

RANS simulations are performed with the Navier-Stokes solver ISIS-CFD, developed at ECN – CNRS and distributed in the FINE/Marine computing suite from Cadence Design Systems. ISIS-CFD is an incompressible unstructured finite-volume solver for multifluid flow (Queutey and Visonneau, 2007). The velocity field is obtained from the momentum conservation equations and the pressure field is extracted from the mass conservation constraint transformed into a pressure equation. Free-surface flow is simulated with a conservation equation for the volume fraction of water, discretized with compressive discretization schemes. Body motion can be solved for up to six degrees of freedom (Leroyer and Visonneau, 2005) and is either accompanied by mesh deformation or, for larger amplitude motions, overset meshing is available, where one or more small domains can move freely through a larger background mesh; the domains are dynamically coupled for the resolution of the flow equations.

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, which enables for example adaptive mesh refinement.

**Quasi-Static Positioning** Since the resistance of hydrofoils depends strongly on the lift it generates, it is useful to compare different foils at the same lift and sideforce. For this, the hydrofoil rake (pitch) and sideslip angles are adjusted in a quasi-static positioning method, to find the equilibrium position that generates the target forces (Ploé, 2018). In this procedure, corrections to the angles are regularly computed, dividing the missing forces by estimated lift slopes for the hydrofoil. The foil is rotated to the new position over several time steps, then remains in position for a few time steps to enable the forces to stabilize, until a new correction is computed. This procedure converges rapidly and it is stable even for large time steps, since no unsteady motion has to be resolved.

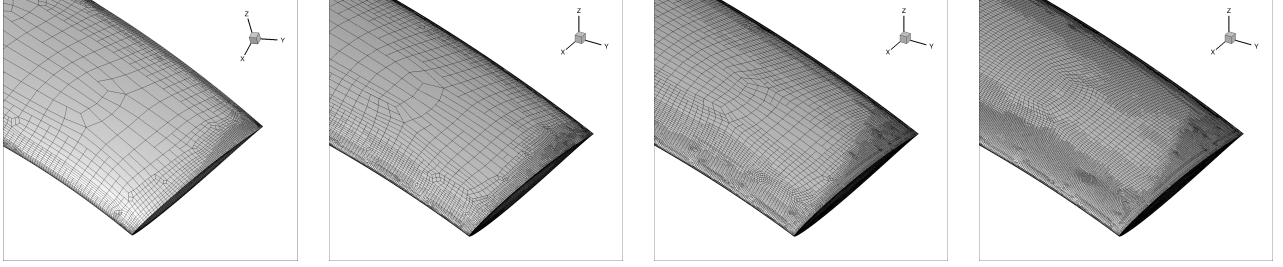
**Adaptive Grid Refinement** The adaptive grid refinement (AGR) method utilizes anisotropic refinement of unstructured hexahedral meshes by dividing the cells into finer meshes (Wackers et al., 2017, 2022a). The refinement procedure is fully parallel and is called at regular intervals during the simulation. Once the grid is adapted to the flow, the refinement procedure stops modifying the mesh.

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 the water surface position and on second derivatives of the flow variables, which give an 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

$$\|\mathcal{C}_p \mathbf{d}_{p,j}\| = T_r. \quad (3.11)$$

The refinement criterion based on the second derivatives of the flow is not very sensitive to grid refinement (Wackers et al., 2017), so the cell sizes everywhere are proportional to the constant  $T_r$ .



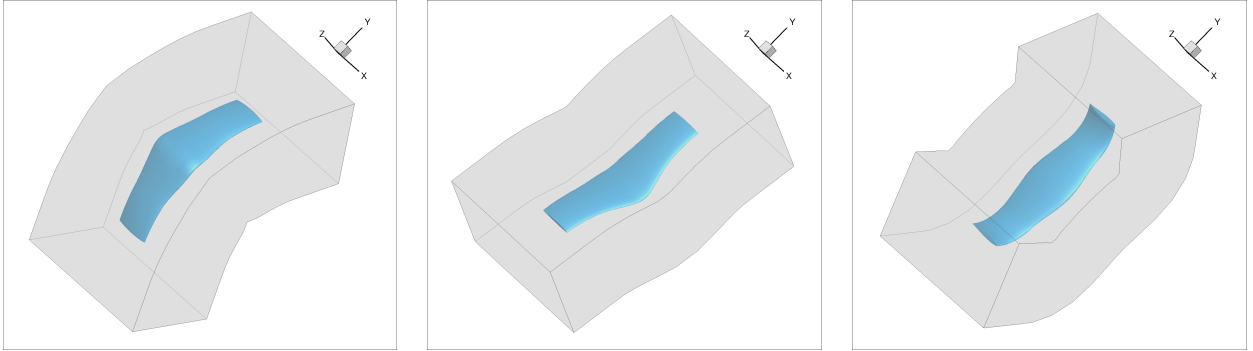


**Figure 3:** Refinement levels from low- to high-fidelity (left to right).

For MF optimization, grid adaptation is used to create several RANS fidelities. The interest of this procedure is that different fidelity results can be obtained by running the same simulations and simply changing the threshold  $T_r$  as seen in Fig. 3, where the threshold  $T_r$  ratio is 2:1 between fidelity levels. Hence, the threshold variation enables to automate the MF simulations. The values for  $T_r$  and other parameters for the RANS simulations follow the protocol of (Richeux, 2022).

**Foil Sock Approach** For hydrofoil calculations, this paper uses a ‘sock mesh’: a curved domain that follows the shape of the foil with a body-aligned mesh, embedded in a larger domain by the overset mesh (or chimera) technique.

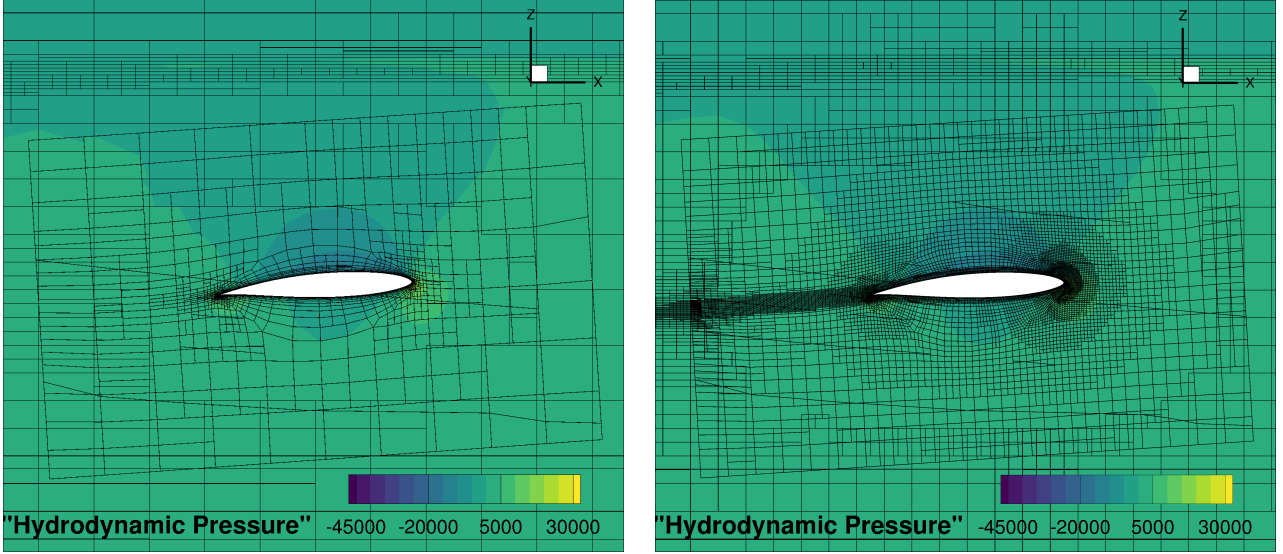
The overset domain is generated automatically as a fixed offset from the spine of each hydrofoil shape (see Sect. 4.2). The foil-shaped domain and an associated structured curved initial mesh are then imported in Hexpress, which transforms the coarse initial mesh into a geometry-adapted Hexpress mesh (the final meshes are created by adaptive refinement of this mesh in ISIS-CFD). In Fig. 4 various kitefoil geometries are shown in their adapted foil-shaped domains. Even for very steep-angled hydrofoil cases or high tip angles, the domain aligns itself generating a foil-shaped overset domain.



**Figure 4:** Foil-shaped overset mesh domain for various kitefoil geometries.

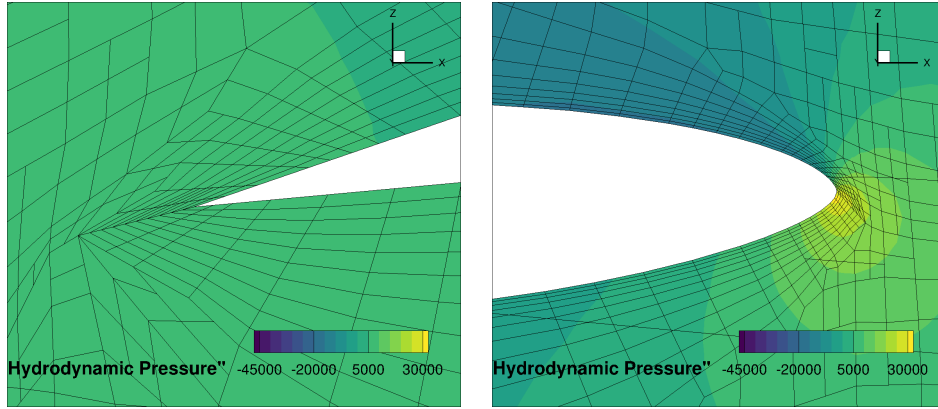
Since the overset mesh is aligned with the foil, it produces surface meshes that locally resemble structured body-aligned grids. Especially for hydrofoils this approach provides better quality cells on the leading and trailing edges. Also, it reduces the total required number of cells. In Fig. 5 the adapted mesh and the rotated overset domain is seen for the low and highest fidelity levels around the foil, respectively.

AGR effectively captures waves, wake flows and boundary layers, so combining this strategy with the foil sock approach enables efficient automated optimization even for challenging hydrofoil cases with very slender geometries. The mesh resulting from the AGR combined with the foil sock approach



**Figure 5:** Adapted mesh and the rotated overset domain around the kitefoil centerplane for low fidelity (left) and highest fidelity (right) levels.

is shown in Fig. 6 for the trailing and leading edges at the lowest fidelity level. This figure shows the refinement layers around the foil and how the mesh resolution accurately captures even the very sharp end of the geometry. The ultimate aim in this procedure is to utilize adaptation without user intervention in any kind of industrial problem.



**Figure 6:** Foil sock approach combined with AGR: mesh for the lowest fidelity at the trailing and leading edge.

#### 4. OPTIMIZATION PROBLEM

In this study, two different approaches for realizing the MF surrogate model are confronted: i) RANS flow solver only, where 4 fidelity levels are obtained varying the mesh refinement, ii) RANS flow solver for the 4 highest fidelities, combined with the potential flow solver as an additional 5-th fidelity level.

These approaches are applied to the optimization of a kitefoil geometry. A typical kitefoil is presented in Fig. 7, where three main parts can be seen: the wings (main/lifting foil & rear/stabilizing foil), the mast, and the fuselage. Especially the main foil may profoundly change the characteristics of the kitefoil and dominates the foiling performance. Therefore, this study focuses on the main/lifting wing.



**Figure 7:** A kitefoil geometry (adopted from: Alpinefoil company kitefoil-RAVE).

#### 4.1 Problem Formulation

The problem addresses the drag coefficient minimization of a kitefoil-type hydrofoil at constant side and lift force, modeled with 11 parameters: 10 for the parametrization of the geometry and one for the heel angle. The following minimization problem is solved

$$\begin{aligned}
 & \text{minimize} && f(\mathbf{x}) = F_X(\mathbf{x}) \\
 & \text{subject to} && F_Y(\mathbf{x}) = 150N, \\
 & && F_Z(\mathbf{x}) = 800N, \\
 & \text{and to} && \mathbf{l} \leq \mathbf{x}_{1-11} \leq \mathbf{u},
 \end{aligned} \tag{4.1}$$

where  $\mathbf{x}$  is the design variable vector,  $F_X$ ,  $F_Y$  and  $F_Z$  are respectively the drag, sideforce, and lift. The two equality constraints are necessary to compare different geometries at the same lift force (equal to the weight of the object), and sideforce (imposed by the traction of the kite) since the drag depends strongly on the foil's loads. The vectors  $\mathbf{l}$  and  $\mathbf{u}$  are the lower and upper bounds of the design space.

The simulation conditions are: velocity  $U = 8$  m/s, fluid density  $\rho = 1,026$  kg/m<sup>3</sup>, reference root chord  $c = 0.175$  m with a chord based Reynolds number  $Re = 1.47 \cdot 10^6$ . A free surface is considered in the simulations and is located at 0.2 m above the hydrofoil center.

#### 4.2 Kitefoil-type Geometry

**Parametrization** The hydrofoil is described with 10 parameters (variables). These parameters, together with default values and the ranges for the design space, are given in Tab. 1. The default valued kitefoil will be called ‘base geometry’ and is used in base and optimum geometry comparisons.

meaning Fig. 8 shows the geometrical meaning of the variables. In the top, front and cross-sectional views, the optimization variables are shown on the main foil. In addition to these, another parameter, the heel angle, is shown which is optimized so that the foil adopts the required lift and sideforce.

In several cases, the design parameters are ratios: the span ratio  $\text{arc2}$  is  $\frac{\text{Span arc2}}{\text{Span}}$ , while the chord ratios are the indicated chords divided by the root chord: the chord ratio  $\text{arc2}$  is  $\frac{\text{Chord arc2}}{\text{Chord root}}$  and chord ratio tip is  $\frac{\text{Chord tip}}{\text{Chord root}}$ . The use of these ratios allows the main size parameters (such as the span or chord) to be adjusted separately from the wing shape, given by the ratios.

**Geometrical Modeller** The geometric modeller is based on the doctorate thesis (Ploé, 2018) Chapter 6. It follows the principle of specifying the physical dimensions of the foil as design parameters (such as the span, chord, dihedral angle, etc.). The principle used in the creation of variant geometries

**Table 1:** Base geometry and design parameters variation in DoE.

Parameter	Default	Range	Units
Span	0.5	[0.2, 0.8]	<i>m</i>
Dihedral angle	-5.0	[-25, 15]	<i>deg</i>
Span ratio arc2	0.55	[0.2, 0.9]	
Tip angle	-5.0	[-40, 30]	<i>deg</i>
Sweep angle	0.0	[-20, 20]	<i>deg</i>
Chord root	0.175	[0.1, 0.25]	<i>m</i>
Chord ratio arc2	0.75	[0.3, 1.2]	
Chord ratio tip	0.7	[0.2, 1.2]	
Camber	0.0	[-0.05, 0.05]	<i>deg</i>
Twist	0.0	[-10, 10]	<i>deg</i>

has 4 steps; i) spine generation, ii) creating sections; base profile mixing and section transformation, iii) positioning along the spine and iv) converting to output.

First, the foil spine is defined, then divided into sections. The spine curve is 3D, allowing the optimization of its shape when the modeler is embedded in an optimization process. The spine is generated from 4 continuous arcs defined by radius and chord lengths. These arcs are joined based on the tangency constraint to get a smooth transition. The tip of the foil is parameterized based on the angle between the chord and the horizontal axis.

The spine gives the outline of the hydrofoil but the envelope is constructed by the foil sections, whose shape is varied based on the position on the spine. Smooth variation of the section parameters is achieved by interpolation between control points, with a B-spline. First, the initial section is modified by conformal mapping to obtain the desired camber. Once the camber is adjusted, the required chord length profile is produced by scaling transformations.

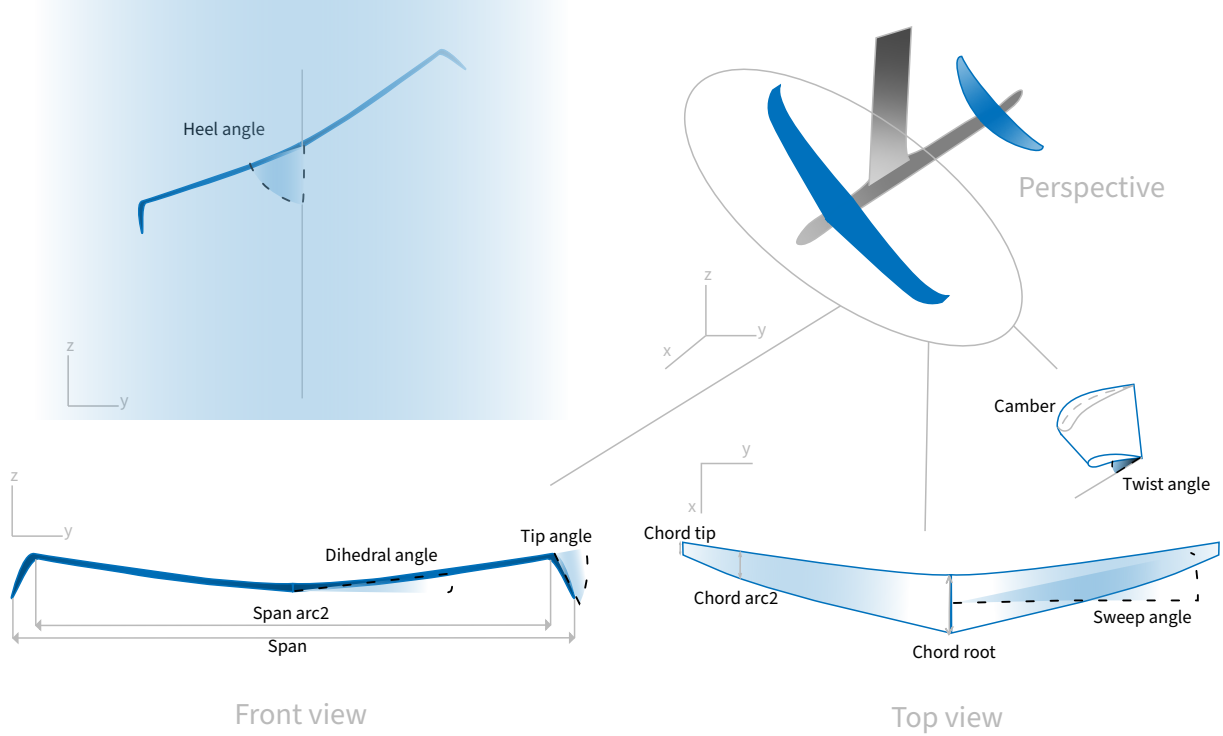
To conclude the creation of variant geometry, the sections are rotated and translated to adopt the required twist, spine alignment and translation. Once the final position assigned for each section, the foil envelope is created as a triangulated point cloud. The heel angle, which is the final optimized parameter, is applied during the flow simulations.

## 5. SIMULATION SETUP AND TESTS

The setups of the flow solvers PUFFIn and ISIS-CFD for the kitefoil optimization are described in the following subsections and the results of these two solvers are compared.

### 5.1 PUFFIn

The size of the free-surface mesh (Fig. 1) is chosen as ( $x_{\min}: -3.5$ ,  $x_{\max}: 1.0$ ,  $y_{\min}: -1.5$ ,  $y_{\max}: 1.0$ ) m. The number of panels used on the free surface in  $x$ -direction is 70 and in  $y$ -direction 50, while it is 34 and 62 for the profile. Adding the potential flow solver as the lowest fidelity to RANS based intermediate and highest fidelities, the resulting computational cost ratio (i.e. the cost divided by the computational cost for the highest fidelity)  $\beta_5 = 0.005$  is used, since one simulation with PUFFIn is completed in one minute (wall-clock time) approximately. On the other hand a complete simulation with the highest fidelity level with ISIS-CFD takes about four hours (wall-clock time).



**Figure 8:** Kitefoil parametric geometry description.

## 5.2 ISIS-CFD

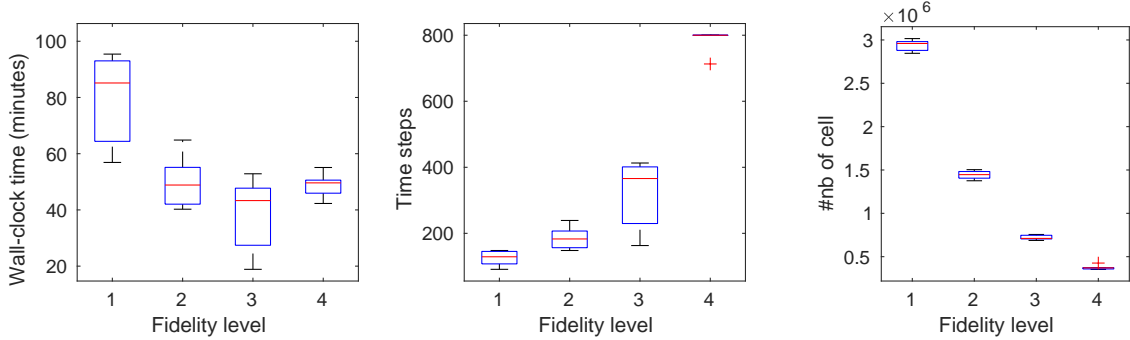
The computational domain runs from  $32c$  in front of the leading edge to  $-80c$  behind the hydrofoil, from  $-48c$  to  $48c$  laterally, and from  $-40c$  to  $24c$  vertically. Dirichlet conditions on the velocity are imposed, except on the top and bottom which have an imposed pressure. The hydrofoil surface is treated with a wall law and  $y^+ = 60$  for the first layer. Turbulence is modeled with the standard  $k-\omega$  SST model (Menter, 1994).

Tests are run with four fidelity levels ( $N = 1, 2, 3, 4$ ). The refinement threshold value  $T_r$  is set to 0.0625, 0.125, 0.25 and 0.5 from highest- to lowest-fidelity. This results in a cell size ratio of 2 : 1 between each grid and the next coarser one. The final grids have about 3M, 1.5M, 700k and 370k cells respectively. In Fig. 3 the resulting mesh refinement is shown on the base geometry for low- to high fidelity levels. Computational cost (CC) ratios of  $\beta_1 = 1$ ,  $\beta_2 = 0.62$ ,  $\beta_3 = 0.39$  and  $\beta_4 = 0.20$  are used.

In Fig. 9 the required wall-clock time for a simulation, time steps to convergence, and corresponding mesh size are plotted against the fidelity level. Figure 9 explains why CC hardly increases between the levels: each high-fidelity simulation is restarted from the preceding lower-fidelity one, so the number of time steps for convergence of the high-fidelity simulations is lower than for the lowest fidelity. Due to the restarts, higher-fidelity CC are the sum of the costs for the level itself and all the coarser ones.

## 5.3 Comparison of ISIS-CFD and PUFFIn

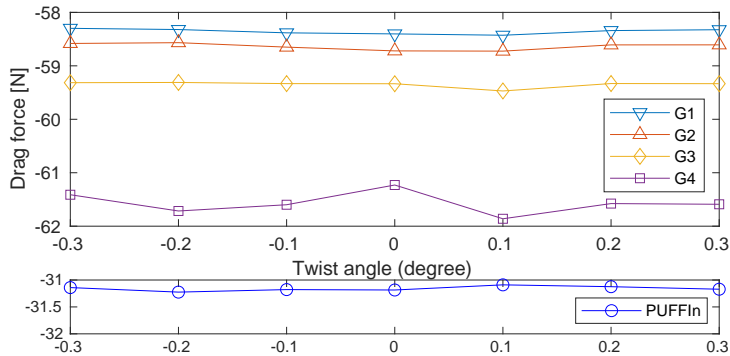
Since the correlation of the solutions provided by the RANS and potential flow solvers have great impact on the performance of multi-fidelity surrogate, this subsection is dedicated to the comparison of the flow solvers. If the discrepancy between the solvers is negligible, then the exploration stage



**Figure 9:** ISIS-CFD: simulation time, required time steps and total number of cell according to fidelity levels.

with the lowest fidelity will be highly useful for higher levels, thus providing efficient multi-fidelity surrogate modeling.

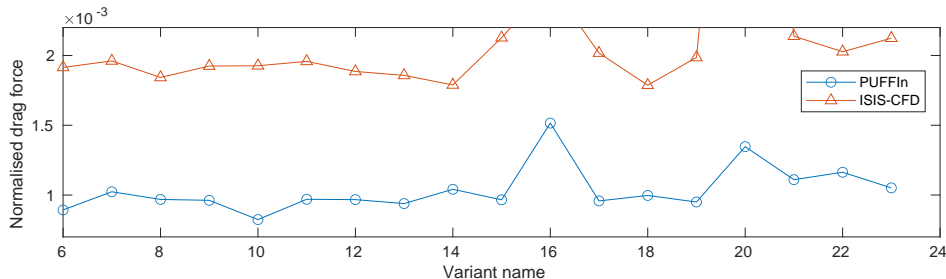
To determine the presence and intensity of numerical noise in the flow solvers, a systematic variation of the twist angle in the range  $\pm 0.3^\circ$  is presented in Fig. 10. The figure shows the drag force value versus the twist angle: for ISIS-CFD, the noise level decreases from lowest to the higher fidelities, since the oscillations diminish for the corresponding drag values. The noise, caused by variations in the meshes (Pellegrini et al., 2022), is proportional to numerical errors as it decays for the finer meshes. Since the potential flow solver always has the same panel distribution, its noise is low.



**Figure 10:** Drag sensitivity to  $0.1^\circ$  twist angle change, top: ISIS-CFD, bottom: PUFFIn.

The compatibility of these flow solvers is further investigated over the startset geometries. In Fig. 11 the variant geometries 6-23 from the startset are evaluated with the two solvers. Since the extreme cases for the heel angle and span caused the simulation to crash, these cases are excluded from the comparison. PUFFIn and ISIS-CFD show similar tendencies over the startset geometries, indicating a good correlation between the two solvers. Therefore, it is expected that PUFFIn has the ability to reliably approximate RANS simulations over a wide range of geometries.

To provide a further quantitative comparison of the two flow solvers, the forces and incidence angles for the baseline geometry are summarized in Tab. 2. The forces are the mean values over the last 50 and 150 time steps for PUFFIn and ISIS-CFD computations, respectively. The  $F_x$  difference stems from the lack of viscous component of the drag for PUFFIn. The forces in  $y$  and  $z$  direction, resulting from the dynamic positioning, are in good agreement, although ISIS-CFD is slightly more precise than PUFFIn. For low-fidelity results, this is probably not an issue.



**Figure 11:** Startset comparison for ISIS-CFD and PUFFIn.

**Table 2:** Comparison of forces and incidence angles for the base geometry evaluated with PUFFIn and ISIS-CFD (highest fidelity).

Solver	$F_x$ [N]	$F_y$ [N]	$F_z$ [N]	Yaw [deg]	Pitch [deg]	Roll (imposed) [deg]
PUFFIn	-30.153	146.852	783.376	0.6841	-2.5844	-10.6
ISIS-CFD	-58.395	150.022	800.066	1.6215	-3.4893	-10.6

Results on the base geometry for PUFFIn, ISIS-CFD lowest, and highest fidelity are compared in Fig. 14. The columns shows the pressure coefficient contour over the hydrofoil surface, the top view, and the free surface elevation affected by the foil geometry. When the accuracy increases the  $C_p$  has a smoother transition which indicates the better resolution of the hydrodynamic characteristics around the foil, see first and second columns of Fig. 14. The free surface visualisations show how much these fidelities are correlated, indicating that the solvers are able to serve as a fidelity level hierarchy.

## 6. OPTIMIZATION RESULTS

The following subsection presents the results for kitefoil optimization in a multi-fidelity framework based on varied combinations of flow solvers. The optimization results are assessed with different metrics inspired by (Pellegrini et al., 2022). These metrics are given in Eq. 6.1 as the position of the design point, the optimum value with respect to the base geometry and the error of the surrogate model in predicting the minimum, respectively.

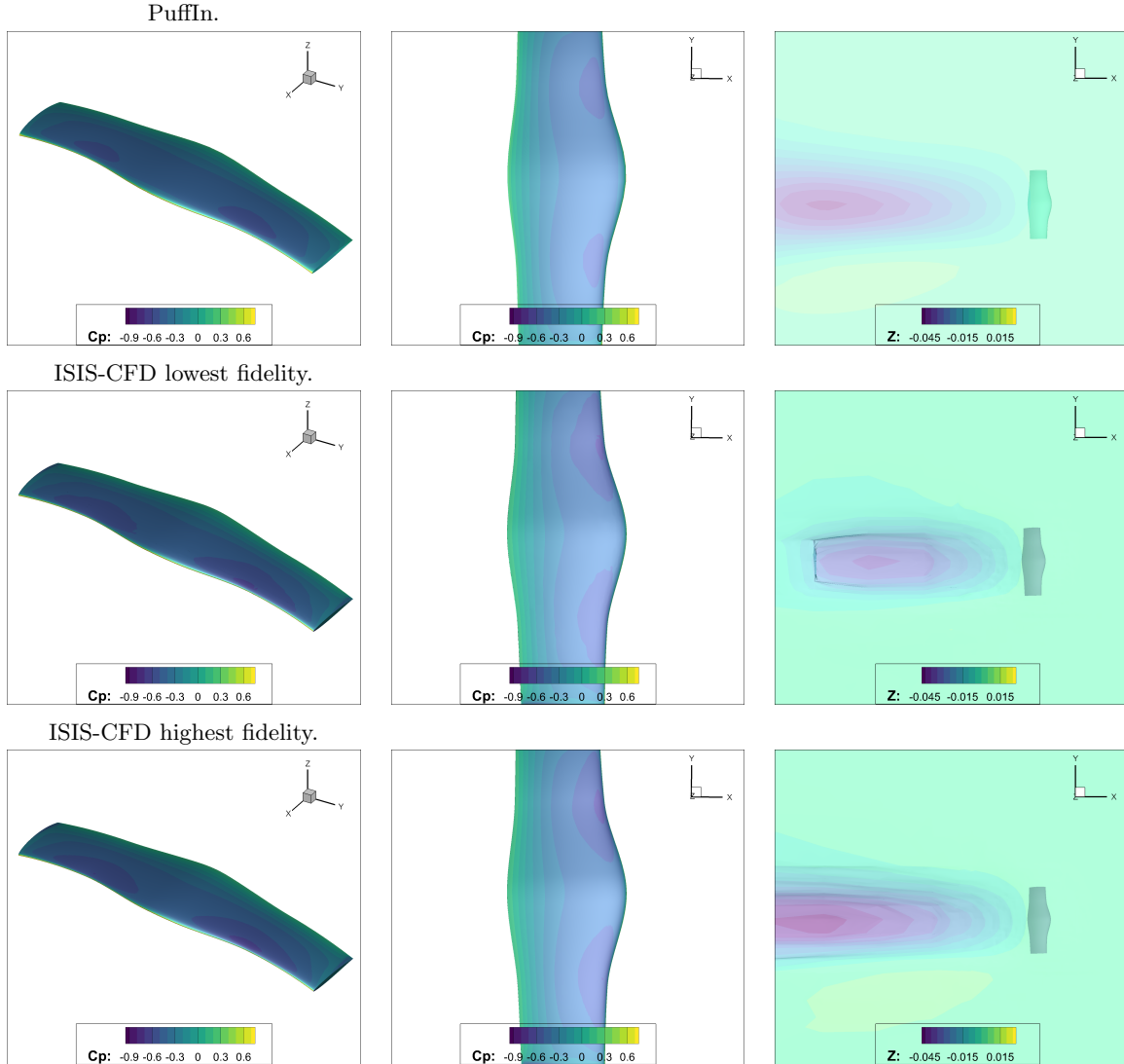
$$\Delta_x \equiv \frac{\|\mathbf{x}^* - \mathbf{x}_0\|}{\sqrt{D}}, \quad \Delta_f \equiv \frac{f(\mathbf{x}^*) - f(\mathbf{x}_0)}{f(\mathbf{x}_0)}, \quad E_p = \frac{\hat{f}(\mathbf{x}^*) - f(\mathbf{x}^*)}{f(\mathbf{x}_0)}. \quad (6.1)$$

In these metrics,  $\mathbf{x}_0$  is the base geometry and  $f(\mathbf{x}_0)$  the original objective function value (drag value of the base geometry),  $\mathbf{x}^*$  is the location of the optimum of the MF surrogate model  $\hat{f}$ . The objective function is the scaled resistance:  $f(\mathbf{x}) = -F_x(\mathbf{x})/\frac{1}{2}\rho V^2$ .

The optimization budget is fixed at  $CC = 50$ , which corresponds to about two weeks of wall clock time. Certain computations crash, since the dynamic positioning may not find a solution if the heel

**Table 3:** Summary of the optimization results.

Framework	$D$	$CC$	$\Delta_x\%$	$\Delta_f\%$	$E_p\%$	$\mathcal{J}_1$	$\mathcal{J}_2$	$\mathcal{J}_3$	$\mathcal{J}_4$	$\mathcal{J}_5$
A	11	50.11	8.9	9.6	-270.3	1	1	33	178	-
B	11	5.93	39.3	87.9	-236.6	1	1	1	1	796



**Figure 12:** Comparison of the flow solvers PUFFIn (top) ISIS-CFD lowest fidelity (middle) ISIS-CFD highest fidelity (bottom) and; pressure coefficient (left) and wave elevation pattern (right), for the base geometry.

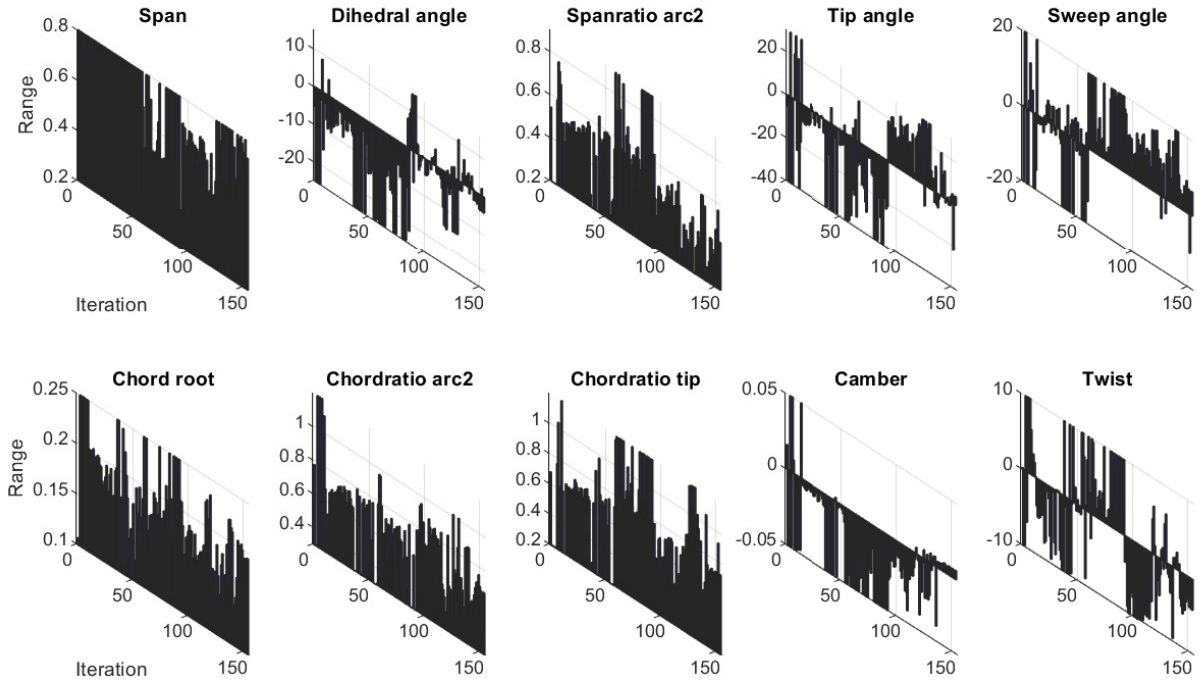
angle is unsuitable for the requested forces. This is taken into account for the surrogate model, by assigning an arbitrary high value (about 10 times the base geometry resistance) to the training point.

### 6.1 MF framework A: RANS based

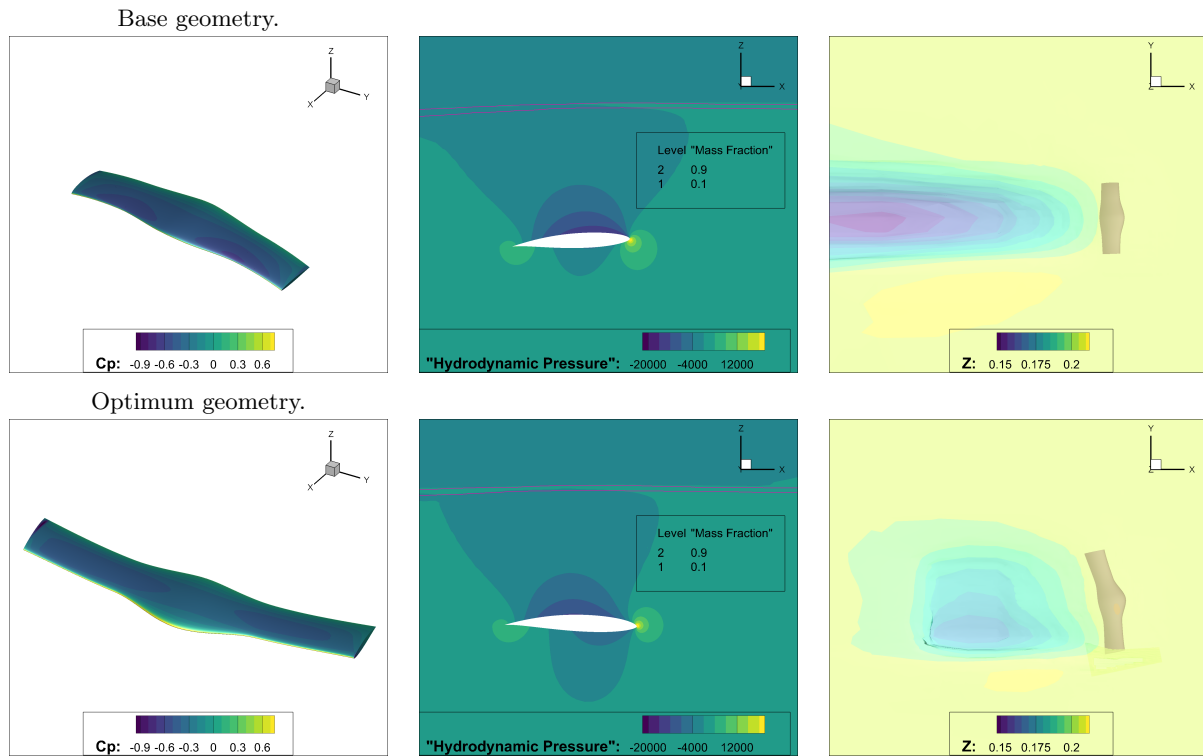
Framework A is based on the RANS flow solver with four levels of fidelity and it is terminated at  $CC = 50.11$ . The evolution of the optimum design is presented in Fig. 13. The surrogate-predicted optimum at each iteration is depicted for each geometrical parameter individually showing that global convergence is not achieved since the parameters are still oscillating.

Table 3 summarizes the optimization results. The adaptive sampling used 178 low-fidelity samples and 33 samples evaluated with the fidelity level 3. This framework does not provide a convergence to an optimum which is also indicated in Fig. 13. Furthermore, the design space still needs to be better explored since the prediction error is high (the predicted drag is negative!).

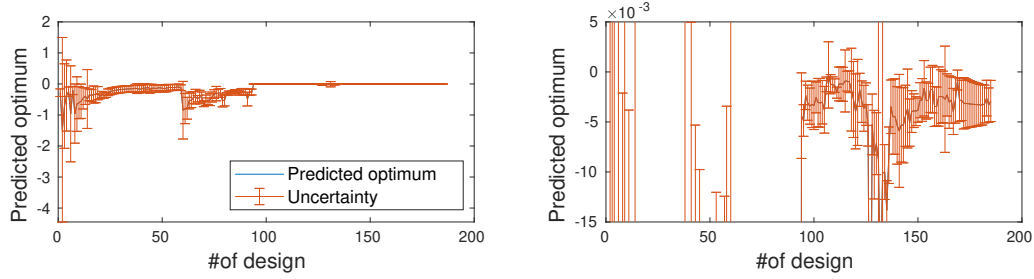




**Figure 13:** MF framework A: variation of the metamodel-predicted optimum for the iterations up to CC 50.



**Figure 14:** MF framework A: Optimum kitefoil comparison with base geometry at CC 50. Top (base), bottom (optimum); from left to right; pressure coefficient, hydrodynamic pressure at  $Y = 0$ , surface elevation.



**Figure 15:** MF framework A: evolution of the optimum prediction with associated uncertainty.

The evolution of the predicted minimum resistance is presented in Fig. 15. The associated uncertainty abruptly diminishes around 100 iterations. However, the right plot which uses the same scale as for Framework B (Fig. 16) indicates there still exists a considerable amount of uncertainty. And this uncertainty is underestimated, since the uncertainty band contains only negative values while the resistance should be positive. These are the main reasons why no optimum design is identified.

The ‘optimum’ design is validated with an ISIS-CFD computation at the highest fidelity level and compared with the base geometry in Fig. 14. The figure presents the pressure coefficient, hydrodynamic pressure around the foil and surface elevation respectively. Since the same view is adopted, geometrical differences may easily be distinguished. The optimised geometry has a wider span and thus a smaller span ratio. In addition, it has a higher chord root, lower twist angle and the rest of the parameters are close to the base geometry, even though the planform is clearly inefficient. The wave elevation (right column) shows that the modified hydrofoil affects a wider free surface area which, despite the shallower wave trough, may indicate a higher drag value.

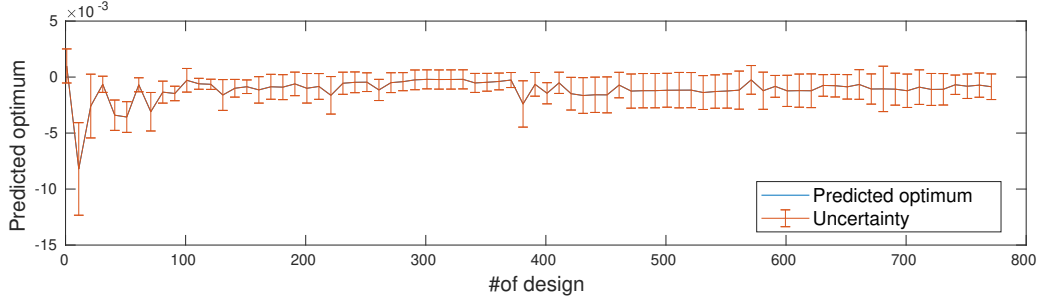
## 6.2 MF framework B: Potential flow-RANS based

The framework B is based on RANS and potential flow solvers. The optimization has been terminated after two weeks and 800 simulations, although the corresponding CC only reached 5.93 which is a very premature stage. The CC hardly increased since there is approximately 1 : 50 ratio in terms of wall-clock time between a PUFFIn simulation and the lowest fidelity ISIS-CFD. Therefore, as the number of sampled points increases, the optimization and adaptive sampling stage itself starts to dominate the design optimization procedure, requiring more time than potential flow simulations.

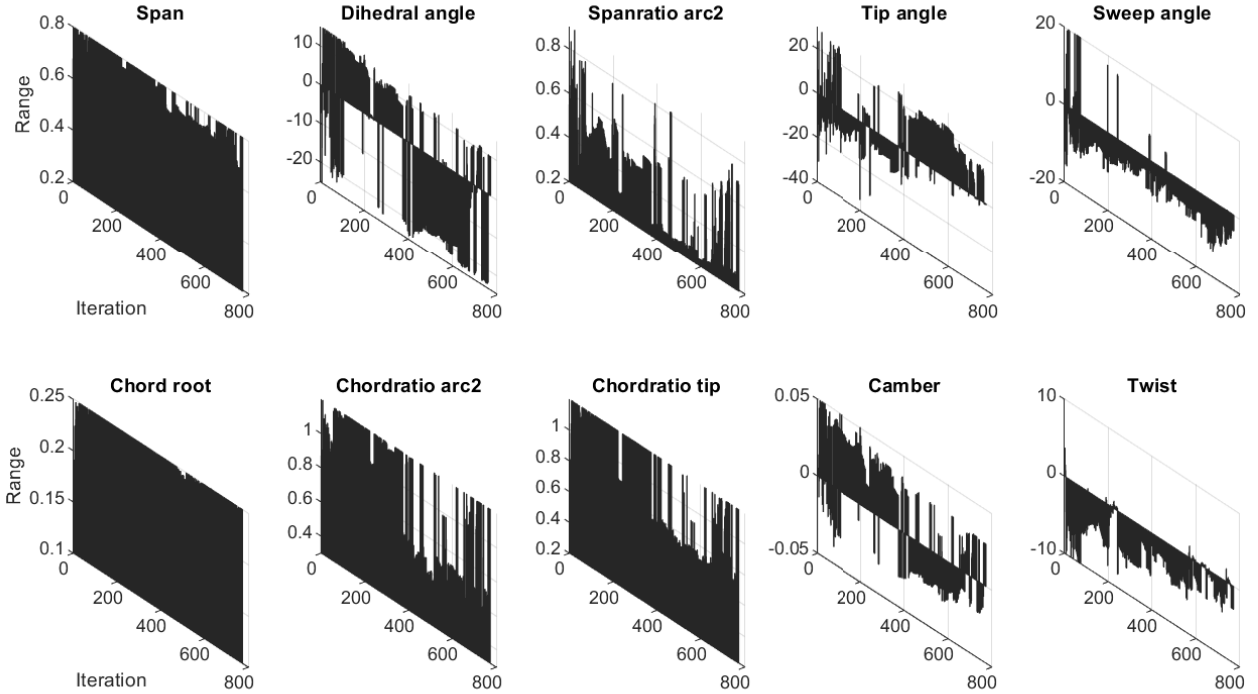
In framework B, the  $\Delta x$  value is greater than in framework A (Tab. 3) and a high prediction error (with negative predicted drag) is observed. This framework used 796 low-fidelity training points and one point from each higher fidelity for the base geometry. The lack of viscous resistance may play a role in the high prediction errors.

The evolution of the optimum design is presented in Fig. 17 based on each dimension individually. The optimization is not converged since each of the parameters is still oscillating. The evolution of the predicted optimum resistance is depicted for every tenth design case, with associated uncertainty estimation, in Fig. 16. The prediction becomes stable relatively rapidly but the uncertainty does not diminish in successive iterations. Also, the predicted optimum drag is negative which shows that the surrogate model is unreliable.

In Fig. 18 the base geometry (top) is compared with the optimum (bottom) based on the pressure coefficient, hydrodynamic pressure around the foil and surface elevation respectively. In terms of geometrical differences with the base form, the optimum design has lower values only for the sweep



**Figure 16:** MF framework B: evolution of the optimum resistance prediction with associated uncertainty.



**Figure 17:** MF framework B: variation of the metamodel-predicted optimum for the iterations up to  $CC=5.93$ .

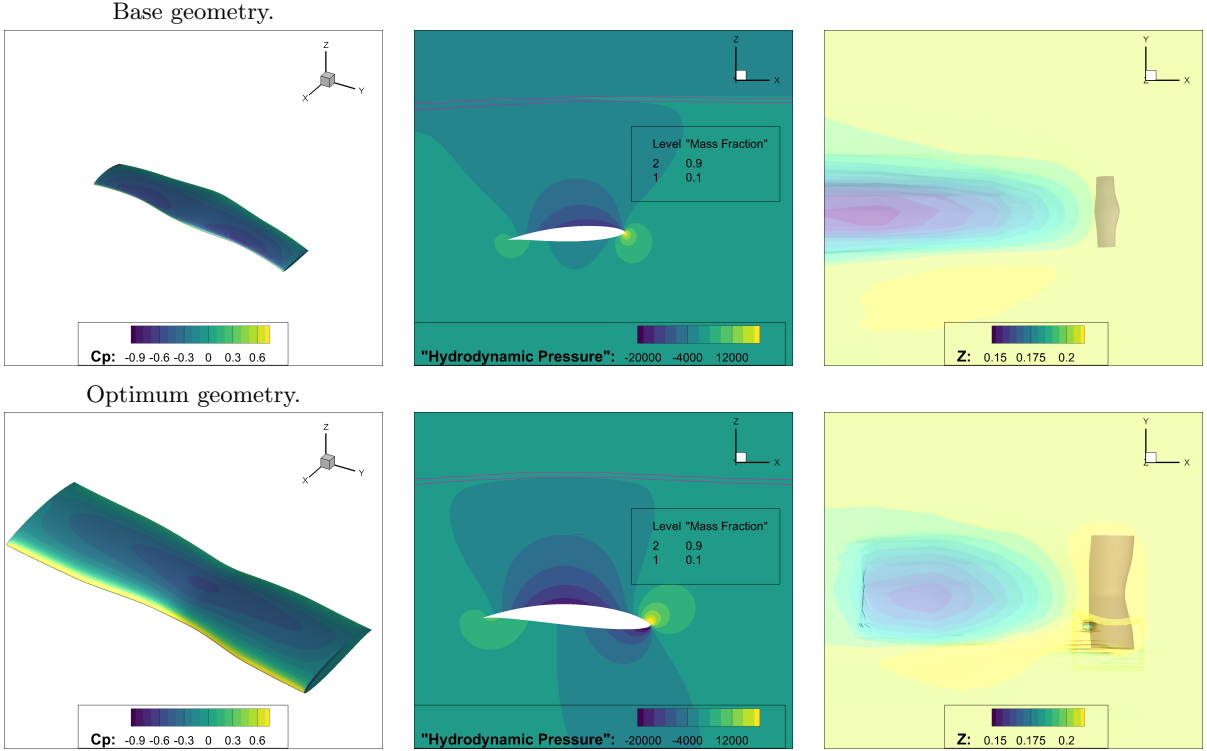
angle, twist and larger values for the rest of the variables. In addition, the optimum geometry takes upper bound values for span, chord root and chord ratio arc2. Since the same viewpoint is used for these comparisons, the pressure coefficient shows the geometrical difference from the base geometry.

The optimum has a tendency to have lower or upper bound values for the design parameters (Fig. 17) which results in an unrealistic shape. It is possible that the unreliable surrogate model favors the boundaries of the domain.

### 6.3 Discussion

Based on the presented results, different practical ways to improve the optimization can be identified. These are discussed here and will be implemented in the coming months.

A first issue is the appearance of negative resistance predictions from the surrogate models, even though all the simulated values are positive and the power-law SRBF kernels (Sect. 2.3) are well-suited for representing these results. The likely cause for this is that, when a simulation fails to



**Figure 18:** MF framework B: Optimum kitefoil comparison with base geometry at CC 5.93. Top (base), bottom (optimum), from left to right; pressure coefficient, hydrodynamic pressure, surface elevation.

converge, a high drag value (0.02) is assigned to discourage the optimization algorithm to explore this region. This approach may negatively affect the behavior of the MF surrogate, since it results in non-smooth training data. Therefore, an additional test is conducted on the framework B dataset, where the assigned value in case of a crash is reduced to 0.0025, i.e. only slightly more than the highest observed training values. The results are compared in Tab. 4: for the same dataset, the prediction and validation errors decrease significantly, the predicted resistance is now positive. Therefore, a better penalization approach for the optimization should be considered to preserve the surrogate quality; this can be either through a better choice of the training set values for crashes, or through a dedicated constraint metamodel to exclude these zones.

**Table 4:** The effect of the training point value assigned to crashed simulations, for Framework B.

	$D$	$\Delta_x\%$	$\Delta_f\%$	$E_p\%$	$\mathcal{J}_1$	$\mathcal{J}_2$	$\mathcal{J}_3$	$\mathcal{J}_4$	$\mathcal{J}_5$
Original (0.02)	11	39.3	87.9	-236.6	1	1	1	1	796
Updated (0.0025)	11	27.7	47.5	-111.6	1	1	1	1	796

A second problem is the underfitting of the data due to the noise filtering. In Tab. 5 the numbers of RBF centers are presented for each fidelity level. Compared to the available number of training points and given the high dimension of the design space ( $D = 11$ ), the number of centers is too low, especially since PUFFIn data are not very noisy. The low number of centers likely indicates clustering of the training points instead of exploration, which may affect the surrogate accuracy. To overcome the clustering, a novel noise filtering approach is under development (Wackers *et al.*, 2023).

**Table 5:** Number of SRBF centers  $M$  for each fidelity level.

Framework	1	2	3	4	5
A	1	1	23	24	-
B	1	1	1	1	24

A final issue is that the potential flow code is much faster than the viscous flow solver, so it can rapidly add large numbers of training points. However, in framework B the cost of the adaptive sampling itself starts to dominate the optimization procedure, deteriorating the advantage of such a rapid solver. To take advantage of the potential flow solver, it is necessary to accelerate the adaptive sampling stage; by diminishing the number of RBF kernels used to evaluate the means in (2.4) from 1000 to 100 and by using the shared-memory parallelization of the adaptive sampling tool, its cost can probably be reduced by a factor 10 – 50.

## 7. CONCLUSIONS & FUTURE WORK

A comparison of two multi-fidelity frameworks for the high-dimensional optimization of a realistic hydrofoil with a limited budget of function evaluations is presented and discussed. The paper both evaluates the suitability of the flow solvers for automated multi-fidelity shape optimization and it investigates the optimization itself.

The presented automated simulation strategy based on RANS simulations appears to be ready for industrial multi-fidelity optimization without user intervention. It is demonstrated that ‘sock mesh’ overset domains can be generated automatically, which provides efficient high-quality meshes that are perfectly aligned with the geometry. In addition, automated mesh adaptation generates both coarse and very fine meshes with hardly any user intervention, as seen from the figures in section 3.2.

The potential flow solver shows excellent correlation with the RANS solutions. Over the whole range of geometries considered here, both solvers have similar tendencies which is important for efficient multi-fidelity optimization. Furthermore, the cost of a potential-flow solution is less than 0.5% of a high-fidelity RANS simulation, which means that potential-flow low-fidelity simulations have a great potential to reduce the computational cost of hydrofoil optimization.

For the optimization, several issues were identified. First, dynamic positioning for different hydrofoil geometries invariably leads to non-converged simulations (as mentioned also by Ploé (2018)). The high objective value currently assigned to non-converged training points may suppress the actual behavior of the surrogate and cause unreliable predictions, as seen in section 6.3. Optimization with a dedicated constraint metamodel to exclude non-stable regions will be considered to overcome this issue.

Another issue is data clustering in the adaptive sampling, which is indicated by the low observed number of kernel centers. To overcome this issue, a new noise filtering approach is presented in Wackers et al. (2023) which will also be tested on the current frameworks as a future study.

Free-surface proximity is a challenge when using potential-flow solvers. For efficient exploration, the adaptive sampling has to sample boundary values which correspond to the most uncertain regions of the design space. This results in the creation of extreme geometries, which may operate very close to the surface. These cases may have low correlation between the potential flow solver and RANS.

When a rapid flow solver is available, as the number of training points increases, the adaptive sampling

stage starts to require more time than a simulation. Solutions for this are available, as outlined in Sect. 6.3. However, a general conclusion is that for realistic optimization, the cost of adaptive sampling and optimization may not be negligible.

Finally, despite the use of multi-fidelity surrogate models, the dimension of the design space may limit the convergence of the optimization. The use of design-space dimensionality reduction methods (Serani and Diez, 2023) was planned for this work and will be performed in the near future. The goal is to see if for a fixed computational time, a reduced design space which still retains most of the geometrical variation leads to a better optimum than the full design space.

### ACKNOWLEDGEMENTS

This work received funds from the Instituts Carnot ARTS and MERS in the projects OptiFoil and ORUP, which is gratefully acknowledged. CNR-INM is partially supported by the Office of Naval Research through NICOP grant N62909-21-1-2042, administered by Dr. Elena McCarthy and by Dr. Woei-Min Lin of the Office of Naval Research Global and the Office of Naval Research, respectively.

### AUTHORS' CONTRIBUTION

HPS performed all simulations and the main analysis, assisted by JW. The surrogate model was developed by RP, AS, MD and JW. PUFFIn is developed by PP, MS, JBL, BA, FH and PB. HPS and JW wrote the first draft of the manuscript. All authors discussed the results and commented on the manuscript.

### REFERENCES

- P. S. Beran, D. Bryson, A. S. Thelen, M. Diez, and A. Serani. Comparison of multi-fidelity approaches for military vehicle design. In *AIAA AVIATION 2020 FORUM*, page 3158, 2020.
- J. Katz and A. Plotkin. *Low-speed aerodynamics*. Cambridge University Press, 2001.
- M. C. Kennedy and A. O'Hagan. Predicting the output from a complex computer code when fast approximations are available. *Biometrika*, 87(1):1–13, 2000.
- A. Leroyer and M. Visonneau. Numerical methods for RANSE simulations of a self-propelled fish-like body. *Journal of Fluids and Structures*, 20(7):975–991, 2005.
- S. Lloyd. Least squares quantization in PCM. *IEEE transactions on information theory*, 28(2):129–137, 1982.
- F. R. Menter. Two-equation eddy-viscosity turbulence models for engineering applications. *AIAA journal*, 32(8):1598–1605, 1994.
- J. Nagawkar and L. Leifsson. Multifidelity aerodynamic flow field prediction using random forest-based machine learning. *Aerospace Science and Technology*, 123:107449, 2022.
- G. Ortali, N. Demo, and G. Rozza. A Gaussian process regression approach within a data-driven POD framework for engineering problems in fluid dynamics. *Mathematics in Engineering*, 4(3):1–16, 2022.
- R. Pellegrini, J. Wackers, R. Broglia, A. Serani, M. Visonneau, and M. Diez. A multi-fidelity active learning method for global design optimization problems with noisy evaluations. *Engineering with Computers*, pages 1–24, 2022.

- P. Perali, M. Sacher, J.-B. Leroux, J. Wackers, B. Augier, F. Hauville, and P. Bot. Comparaisons de conditions aux limites linéarisées pour l'étude d'un hydrofoil submergé à l'aide d'une approche potentielle. In *25e Congrès Français de Mécanique*, Nantes, France, 2022.
- L. Pernod, M. Sacher, J. Wackers, B. Augier, and P. Bot. Free-surface effects on two-dimensional hydrofoils by RANS-VOF simulations. *Journal of Sailing Technology*, 8(01):24–38, 2023.
- P. Ploé. *Surrogate-based optimization of hydrofoil shapes using RANS simulations*. PhD thesis, École Centrale de Nantes, 2018.
- P. Queutey and M. Visonneau. An interface capturing method for free-surface hydrodynamic flows. *Computers & Fluids*, 36(9):1481–1510, 2007.
- J. Richeux. Towards the automation of adaptive RANS simulations for hydrofoils. Master's thesis, École Centrale de Nantes, 2022.
- P. Robin, A. Leroyer, J. Richeux, D. de Prémorél, and J. Wackers. Starting off the right foot with foil sock approach and AGR criterion. In *24th Numerical Towing Tank Symposium (NuTTS 2022)*, Zagreb, Croatia, 2022.
- A. Serani and M. Diez. Parametric model embedding. *Computer Methods in Applied Mechanics and Engineering*, 404:115776, 2023.
- A. Serani, R. Pellegrini, J. Wackers, C.-E. Jeanson, P. Queutey, M. Visonneau, and M. Diez. Adaptive multi-fidelity sampling for CFD-based optimisation via radial basis function metamodels. *International Journal of Computational Fluid Dynamics*, 33(6-7):237–255, 2019.
- A. Serani, F. Stern, E. F. Campana, and M. Diez. Hull-form stochastic optimization via computational-cost reduction methods. *Engineering with Computers*, 38(Suppl 3):2245–2269, 2022.
- S. Volpi, M. Diez, N. J. Gaul, H. Song, U. Iemma, K. K. Choi, E. F. Campana, and F. Stern. Development and validation of a dynamic metamodel based on stochastic radial basis functions and uncertainty quantification. *Structural and Multidisciplinary Optimization*, 51:347–368, 2015.
- J. Wackers, G. B. Deng, E. Guilmineau, A. Leroyer, P. Queutey, M. Visonneau, A. Palmieri, and A. Liverani. Can adaptive grid refinement produce grid-independent solutions for incompressible flows? *Journal of Computational Physics*, 344:364 – 380, 2017.
- J. Wackers, G. B. Deng, C. Raymond, E. Guilmineau, A. Leroyer, P. Queutey, and M. Visonneau. Adaptive grid refinement for ship resistance computations. *Ocean Engineering*, 250:110969, 2022a.
- J. Wackers, H. Pehlivan Solak, M. Visonneau, R. Pellegrini, A. Serani, and M. Diez. Adaptive multi-fidelity surrogate modelling for high-quality shape optimization. In *Research workshop AVT-354 on Multifidelity methods for military vehicle design*, Varna, Bulgaria, 2022b.
- J. Wackers, R. Pellegrini, M. Diez, A. Serani, and M. Visonneau. Improving active learning in multi-fidelity hydrodynamic optimization. In *34th Symposium on Naval Hydrodynamics*, Washington, DC, 2022c.
- J. Wackers, H. Pehlivan Solak, R. Pellegrini, A. Serani, and M. Diez. Error estimation for surrogate models with noisy small-sized training sets. In *Adaptive Modelling and Simulation (ADMOS 2023)*, Gothenburg, Sweden, 2023.