Bangladesh University of Engineering and Technology
Dhaka, Bangladesh

Efficient Hydrodynamic Shape Optimization Using Artificial Neural Network-Based Surrogate Model with Adaptive Sampling Strategy

Authors:

Nayma Hassan Emu naymahassanemu@gmail.com
Tayyib Ahsan tayyibahsan17@gmail.com
Md. Abdul Kader 0424122005@name.buet.ac.bd
Muztaba Rafid Rhythm muztabarafid137@gmail.com
Tajimur Rafi tajimurrafi@gmail.com
Md Ebrahim Hossen ebrahimhossenebu1775@gmail.com

* Adapted and extended from the paper titled “Efficient Foil Shape Optimization Using Artificial Neural Network-Based Surrogate Model with Adaptive Sampling Strategy.” In *Proceedings of the 13th International Conference on Electrical and Computer Engineering (ICECE)*, Dhaka, Bangladesh, 2024, pp. 533–538, doi: 10.1109/ICECE64886.2024.11024718.

Abstract

In recent times, hydrofoil-based marine craft have drawn considerable interest due to their ability to generate hydrodynamic lift with reduced hull-water interaction, leading to lower resistance and improved efficiency of high-speed waterborne transportation. Small changes in thickness distribution, camber-like deformation, pressure-side curvature, and suction-side curvature can significantly alter lift, drag, and pressure distribution, making foil section geometry critical to craft performance. Nonetheless, the improvement of lift cannot be regarded as an unconstrained objective in practical hydrofoil design, since excessive drag, unreasonable geometry, increased material usage, and unfavorable pressure distribution may limit the practical applicability of the optimized section. Therefore, efficient foil-shape optimization needs to prioritize improved lifting performance while preserving important design constraints such as drag and sectional area. This requirement is especially crucial for early-stage engineering design in resource-constrained environments, including developing economies such as Bangladesh, where access to large experimental facilities and affordability of high-fidelity simulations along with computing resources may be limited.

To overcome the limitations of experimental and numerical design processes in terms of time, cost, efficiency, and uncertainty, researchers are developing optimization frameworks utilizing surrogate models based on Artificial Neural Networks (ANNs). Further enhancement of these frameworks is achieved by integrating adaptive sampling methods, which improve efficiency by reducing the number of training samples needed to reach optimal solutions. The present study develops an Artificial Neural Network (ANN)-based surrogate model optimization framework to maximize the lift performance of a baseline NACA0012 foil section while adhering to set hydrodynamic constraints. The foil shape is parameterized with Hicks-Henne bump functions that provide a way to deform the upper and lower surfaces in a smooth and controlled manner, using relatively few design variables while preserving the leading and trailing-edge positions. Eight design variables are used, corresponding to four control locations on each surface of the foil. XFOIL is utilized as an inexpensive numerical tool to compute lift and drag coefficient data for the baseline and deformed foil sections at a constant Reynolds number of 1×10^6 and an angle of attack of $5°$. The ANN surrogate model is trained to approximate the nonlinear relationship between the Hicks-Henne design variables and the hydrodynamic coefficients. Then the Sequential Least Squares Programming (SLSQP) is used to optimize this trained surrogate model, subject to constraints that drag coefficient and sectional area remain equal to those of the baseline foil, with an objective of maximizing lift coefficient. Under the same final sampling budget, two sampling strategies are examined and compared: traditional One-Shot Sobol Sampling and Optimization-Based Adaptive Sampling. In the case of One-Shot Sampling, all 173 samples are generated in a single shot to cover the design space uniformly using a Sobol sequence. In the Adaptive Sampling approach, an initial Sobol dataset of 128 samples trains a first surrogate model, after which additional samples are iteratively placed around predicted optimum regions to improve surrogate accuracy where it matters most, rather than distributing sampling effort uniformly across the design space.

The results show that both sampling strategies provide an improved lift performance compared to the baseline foil section, however using the same total number of samples incurs a significantly improved optimum with the Optimization-Based Adaptive Sampling scheme. The One-Shot Sampling model increases the lift coefficient by 28.16% relative to the baseline, while the Adaptive Sampling model achieves a 58.68% lift increase. The drag coefficient remains nearly constant across the baseline and optimized cases, since drag is explicitly constrained during optimization. The result is that the improvement in foil efficiency ratio closely tracks the increase in lift coefficient. The geometry optimized using adaptive sampling shows more effective surface deformation and the corresponding pressure distribution indicates a larger pressure difference between the suction and pressure sides, which explains the improved lift generation. The comparison between optima predicted by ANN and calculated values from XFOIL further indicates that the Adaptive Sampling model gains higher accuracy at the optimum. Despite the fact that the global error metrics over the full design space are broadly comparable for both adaptive and one-shot models, the adaptive model becomes more accurate in the optimum region, highlighting that its advantage lies in concentrating the sampling effort where it is most useful for optimization. This framework can serve as an early-stage screening and optimization tool for hydrofoil craft, extendable to other lift-generating marine devices such as rudders, stabilizer fins, and marine control surfaces, and can be integrated with higher-fidelity CFD, cavitation analysis, structural constraints, and experimental validation to create a multi-fidelity design pipeline.

Index Terms

Artificial Neural Network, Surrogate model, Hydrofoil, Shape Optimization, Sobol Sampling, Adaptive Sampling, XFOIL, Hicks-Henne Function, SLSQP

I. Introduction

Hydrofoils are lift-generating marine appendages designed to produce dynamic lift as a vessel advances through water [1]. This lift is more than just an auxiliary hydrodynamic effect in a hydrofoil craft; it contributes significantly to unloading the hull, reducing wetted surface area, and ultimately contributing to reduced resistance while enhancing high-speed marine performance [2]. Recent system-based hydrofoil optimization for AC75 racing yachts indicates that vessel performance, trim, and operating state are strongly influenced by hydrofoil geometry since foil-generated lift allows the hull to operate out of the water, enabling much higher speeds than conventional displacement operation [3]. Thus, hydrofoil sections need to be designed according to the requirement of producing maximum lift with respect to vessel displacement and in conjunction with controlling drag, pressure distribution, structural feasibility, and operational limitations such as cavitation [4].

The hydrodynamic behavior of a hydrofoil is strongly influenced by its sectional geometry. Small variations in camber, thickness distribution, leading-edge shape, and pressure and suction side curvatures result in drastic changes in the lift coefficient, drag coefficient, and pressure distribution [5]. As a hydrofoil obviously needs to offset a substantial portion of the vessel's weight, maximizing lift under realistic constraint is a primary design goal [2]. However, lift can only be improved if accompanied by design constraints due to the growth of drag, geometric feasibility, and pressure distribution because excessive drag leads to lower transport efficiency and adverse pressure peaks may additionally promote cavitation [6]. Recent Computational Fluid Dynamics (CFD)–Genetic Algorithm (GA) optimization of hydrofoil leading-edge geometry further demonstrates that local geometric modification can improve pressure distribution and delay incipient cavitation, although lift-to-drag variation must still be carefully checked before practical application [7]. This explains why hydrofoil section optimization is typically posed as a constrained design problem rather than an unconstrained search for maximum lift.

As a result, more recent hydrofoil optimization studies have used high-fidelity RANS-based numerical frameworks to account for viscous flow, cavitation, separation effects, and three-dimensional geometric effects. Such optimization revealed that planform, cross-sectional, and junction geometry variables can be tuned to achieve reduced drag while cavitation is delayed and separation avoided [2]. Likewise, cavitation-constrained optimization of America's Cup AC75 hydrofoil sections showed that pressure-based constraints, flap effects, and multipoint operating conditions are important to high-performance hydrofoil design [8]. Hydrostructural optimization of composite hydrofoils further demonstrated that considerations of hydrodynamic performance must be combined with cavitation and structural constraints when material response and foil deformation affect the final design [9].

In hydrofoil craft, the fundamental behavior of foil-section is paramount because lift production directly affects hull emergence, powering requirement, ride quality, and overall vessel efficiency. Recent work on machine-learning-enabled foil design tools has demonstrated the potential for data-driven models to enable rapid forward prediction and inverse design of foil sections with sufficiently accurate predictability for early-stage design exploration [10]. Recent marine and ocean-engineering studies identify data-driven surrogate models as a very promising strategy for achieving more efficient shape optimization by reducing the number of simulations during the design process [11].

Traditional manual design methods, such as parametric analyses and cut-and-try techniques, database-matrix approaches, and other intuitive approaches, heavily rely on data and expertise. These methods are generally unsystematic, and often inefficient practices, which provide discrete designs in the design space, leaving considerable uncertainty in optimality [12]. Whereas methods like Computational Fluid Dynamics based Iterative Design and Analysis require a great deal of processing power which can be both time-consuming and costly. To tackle these problems, researchers have developed optimization algorithms based on surrogate models, including the kriging model [13], the support vector machine model [14], and the neural network model [15]. By eliminating the necessity of extensive numerical computations, the surrogate models have gained popularity in fluid dynamic shape optimization. Hydrofoil design has started to devise surrogate-assisted and multi-fidelity optimization strategies that can significantly reduce the overall numerical cost, whilst still serving as a guide for improvement of hydrodynamic performance [16]. Similar trends are also illustrated in marine foil optimization, where neural-network based prediction and Bayesian optimization have been implemented to speed up the multi-point foil design. [17].

A practical way to reduce the expense of early-stage foil design is to utilize a computationally cheaper surrogate model that has been trained using a small sample of numerical data points, replacing repeated high-fidelity CFD calls with low-cost alternatives. Once a suitable training dataset is generated, Artificial Neural Networks (ANNs) are capable of approximating nonlinear relationships between geometric design variables and hydrodynamic coefficients. Data-driven frameworks have been integrated with geometric parameterization methods such as Hicks-Henne bump functions to explore feasible shape variations while keeping the number of design variables manageable in foil shape optimization [18]. Such methods are especially appealing in the area of hydrofoil section design, since they allow the suction and pressure surfaces to be varied without having to redefine the entire foil geometry. This means that surrogate modeling in this work is not limited to prediction, but rather as a low-cost rapid search approach for exploring the foil design space ahead of validation at higher fidelity.

In this work, the numerical analysis tool used to consider lift and drag coefficient data for the section of baseline foil and deformed foil is XFOIL. XFOIL is a widely used tool for rapid preliminary analysis of two-dimensional foil sections based on a combination of panel-method formulation with boundary-layer analysis [19]. Although XFOIL cannot replace high-fidelity CFD or experimental validation for final hydrofoil-craft design, it can offer an extremely useful low-fidelity screening tool at early stages of optimization when computational resources are constrained. Previous comparative work has also shown that XFOIL can yield reasonable predictions for low Reynolds number foil section analysis before more expensive CFD validation is performed [20]. Therefore, the role of XFOIL in this work is best understood as a resource-efficient early-stage design tool rather than a full-fledged replacement for high-fidelity hydrodynamic evaluations.

Although the use of surrogate models to design fluid-dynamic problems is expanding, their predictive performance varies for different conditions. Zhang et al. [21] showed that Support Vector Machine regression for small sample prediction is less accurate, while a single hidden layer RBF neural network does better; however, this may not be sufficient when the training sets have high non-linearity and irregularity. Among the various shallow-learning models tested, Kriging yielded the most consistent prediction accuracy for small sampling engineering applications. Nevertheless, training these surrogate models requires extensive experimental or numerical data generation, which defeats the purpose of shifting towards surrogate models in the first place. Therefore, it makes sense to further reduce the number of training samples with an aim to reach the global optimum of an optimization problem as fast as possible.

To address this issue, surrogate models are typically coupled with efficient sampling strategies to enhance the model in an iterative method while minimizing redundant simulations. Kim and Boukouvala [22] compared various subset selection methods for surrogate model-based optimization and found that adaptive sampling based on optimization can yield better performance when the number of samples is limited. Other sampling approaches have also been studied. Jiangtao et al. [23] proposed the RMSE and crowdness enhance (RCE) adaptive sampling, a method that provides higher surrogate accuracy with fewer samples. Li et al. [24] obtained realistic foil geometries from a small training database with the help of a method based on deep convolutional generative adversarial network (DCGAN). More recently, Wang et al. [25] adopted ensemble surrogate model with hybrid sampling strategy and reduced the number of CFD calls compared to Non-dominated Sorting Genetic Algorithm-II (NSGA-II) and Particle Swarm Optimization

(PSO). The purpose of exploring these sampling strategies is to develop a framework that can find the global optimum using as few samples as possible, thereby reducing computational time and cost. This study establishes an ANN-based surrogate optimization framework to maximize lift with drag and sectional area constraints, and investigates the benefit of Optimization-Based Adaptive Sampling over One-Shot Sampling in terms of seeking the global optimum with fewer samples.

The motivation to develop such a low-cost hydrofoil design framework has more significance in the developing economies like Bangladesh. Bangladesh has a vast network of inland waterways and inland water transport contributes significantly to national freight and passenger movement. Efficient waterborne transport is directly relevant to improved national mobility and logistics in Bangladesh, as inland waterways account for a significant share of cargo and passenger traffic [26]. The design of a low-carbon inland-waterway strategy is also important since greener and more reliable water transport can lower emissions by improving resilience and facilitating trade connectivity [27]. At the same time, large-scale experimental hydrodynamics facilities, repeated high-fidelity CFD studies, and high-performance computing access may not always be available to universities and new designers in the resource-constrained research environment of Bangladesh. The national R & D expenditure has also remained relatively low as a percentage of GDP [28], further substantiating the importance of computationally economical engineering design methods.

This low-cost foil optimization is not only a computational convenience; it is also an overall strategy that Bangladesh can exploit to use its very limited technical resources as effectively as possible. A low-fidelity-to-surrogate validated workflow allows students, researchers, and local designers to screen hundreds of candidate hydrofoil sections before investing time and cost in expensive CFD or experimental testing. This method may be useful for preliminary design of hydrofoil craft, systematic optimization of marine appendages, and other lift-generating marine devices for efficient generation of lift in an acceptably short time and at a reasonable cost. In particular, given the resource constraints posed by developing economies, it is important to explore efficient water transport technologies without presuming access to unlimited advanced facilities.

In this context, the present study uses the NACA0012 section as a baseline foil because of its geometric simplicity, symmetry, and extensive use as a benchmark shape in fluid-dynamic studies. This study parameterizes the foil shape using the Hicks-Henne bump functions [29] because of its effectiveness with a low number of design variables [30]. Moreover, Hicks-Henne functions also provide sufficiently smooth geometries, along with adequate freedom in shape deformation without altering the positions of the leading and trailing edges [31]. The optimization objective is focused on the hydrofoil relevance where lift is maximized while drag and sectional area are constrained, so that the optimized section yields a greater lift without increasing material usage. Since the sectional area is constant, for the same material, the optimized and baseline sections have unchanged material volume per unit span, which gives about constant sectional weight. In order to make comparisons with the baseline foil in a fair manner, the flow condition is kept unchanged throughout the optimization so that lift and drag change due to geometric modification rather than due to Reynolds number or angle of attack. Motivated by this, the current work develops an ANN-based optimization framework for maximizing foil-section lift. The workflow generates design points using Sobol sampling and optimization-based adaptive sampling, deforming the foil shape with Hicks-Henne functions, running XFOIL for obtaining lift and drag coefficients, and optimizing the trained surrogate model using Sequential Least Squares Programming (SLSQP). An adaptive sampling strategy is proposed where the surrogate model is trained near the optimum, as opposed to spending sampling effort equally across less relevant areas of the design space. This makes the framework more appropriate for resource-limited design environments where every numerical evaluation must add useful content to the quest for an improved hydrofoil section.

The primary contribution of this work is an optimization framework for foil shape specific to hydrofoils, focusing on maximizing lift while being subject to operational constraints to drag and area. This method shows how the combination of a low-fidelity solver and an ANN-based surrogate model with adaptive sampling can lead to a significantly improved lifting section using relatively few simulations. This study presents a more suitable approach for hydrofoil-focused assessment and resource-aware engineering design by framing the optimization objective around lift generation for hydrofoil craft specifically and by linking the method to practical constraints in Bangladesh's marine research environment.

II. Methodology

A. Shape Parameterization

This study utilizes the Hicks-Henne bump function [29] to parameterize the shape of the original foil. This method has shown good results in parameterization of foils with a low number of design parameters [30]. Moreover, Hicks-Henne functions also provide sufficiently smooth geometries, along with adequate freedom in shape deformation without altering the positions of the leading and trailing edges [31]. Since the optimization is performed for a specific flow condition, the angle of attack should be kept constant. This means that the positions of the leading and trailing edges cannot be changed. Considering these requirements, the Hicks-Henne bump function has been chosen as the shape parameterization method for this study.

In this function, a linear combination of n augmented sine functions is added to the original coordinates of the airfoil in the form of a bump. With this parameterization, the upper and lower surfaces of the airfoil are deformed independently. The function can be written in this form:

$$y_{\text{mod}} = y_0 + \sum_{i=0}^{n} a_i \sin^{w_i}\left(\pi x^{\frac{\ln(0.5)}{\ln\left(x_i^M\right)}}\right) \tag{1}$$

Where, y_0 = initial coordinates of the upper or lower surface of the foil, y_{mod} = final coordinates of the upper or lower surface of the foil, n = number of bumps for each one of the upper or lower surfaces, x_i^M = x coordinate of the bump or x coordinate of the control point, w_i = bump width, a_i = bump intensity or coordinate's values of the control point. For this foil deformation process, all other parameters (x_i^M, w_i, n) except for the bump intensity (a_i) are kept fixed or unchanged. Therefore, the design variable is the bump intensity (a_i). Figure 2 illustrates a few examples of deformed shapes with different combinations of bump intensities (a_i).

In the Figure 1 above, the positions of the bumps relative to the control points (red) are marked with different color codes. By adding the bump functions to the original foil (blue), we can obtain the deformed foil (orange).

In this case, four control points have been chosen for each of the upper and lower surfaces, resulting in a total of eight design variables or bump intensities $(a_i = x_1, x_2, x_3, \ldots\ldots, x_8)$ required for this problem. The design variables are points on the foil surface taken at 20%, 40%, 60%, and 80% of the chord on the suction side and pressure side. A Python code was written to create the Hicks-Henne function and generate deformed foil shapes.

B. Sampling Strategies

After defining the design variables and surrogate model type, it is crucial to determine how to vary the parameters within the design space. The sampling method significantly affects the effectiveness of the surrogate model, influencing both the outcome and the number of samples required. Sampling for surrogate models can be categorized into two types. Standard Surrogate Model Optimization (SSMO) typically uses a one-shot sampling approach, such as Latin hypercube sampling (LHS), Sobol sampling,

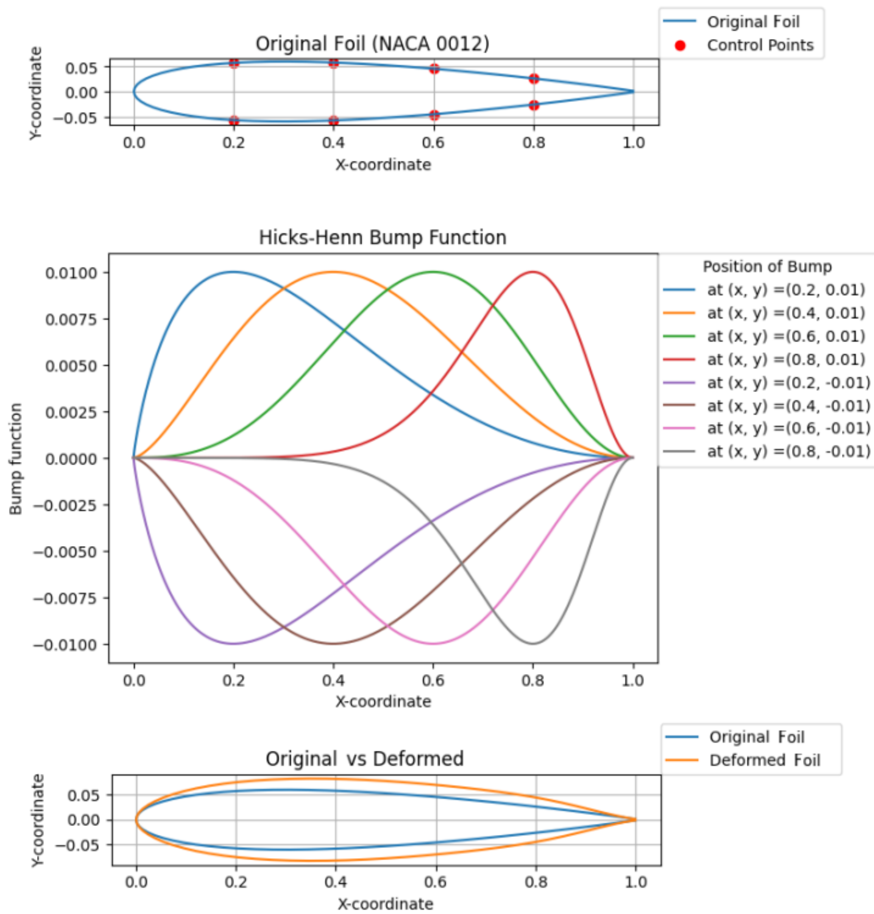

Fig. 1: Representation of foil shape deformation using Hicks-Henne bump function

or random sampling, where the number of computer simulations is fixed at the start of the process. In contrast, Iterative Surrogate Model Optimization (ISMO) employs sequential sampling techniques to build a sufficiently accurate surrogate using as few samples as possible.

In design optimization problems, such as shape optimization, it is essential to develop a surrogate model that accurately predicts the region of interest—such as minima or maxima—with high precision, using a small number of samples. Since data collection is computationally expensive and time-consuming, an incremental approach to placing samples efficiently is preferred. As the goal of surrogate-based optimization (SBO) is to find the global optimum, points near the optima are more critical than others. Thus, building a surrogate that uniformly fits the entire design space would be inefficient given the sample constraints. Therefore, newly added data points should be chosen selectively to ensure the surrogate accurately fits the region near the optimum. This strategy effectively focuses on areas with a higher likelihood of containing the global optimum. It should be emphasized that this study aims to generate a sufficiently accurate surrogate for predicting the region of the optimum, rather than one that fits the entire design space. Consequently, the final model may have lower accuracy in regions of lesser interest but will be highly accurate around the optimum.

This study explores two different sampling techniques. One is the quasi-random, low-discrepancy sequence known as Sobol sampling, which fills the design space uniformly. The other method is an optimization-based adaptive sampling strategy, where an initial set of samples is generated using Sobol,

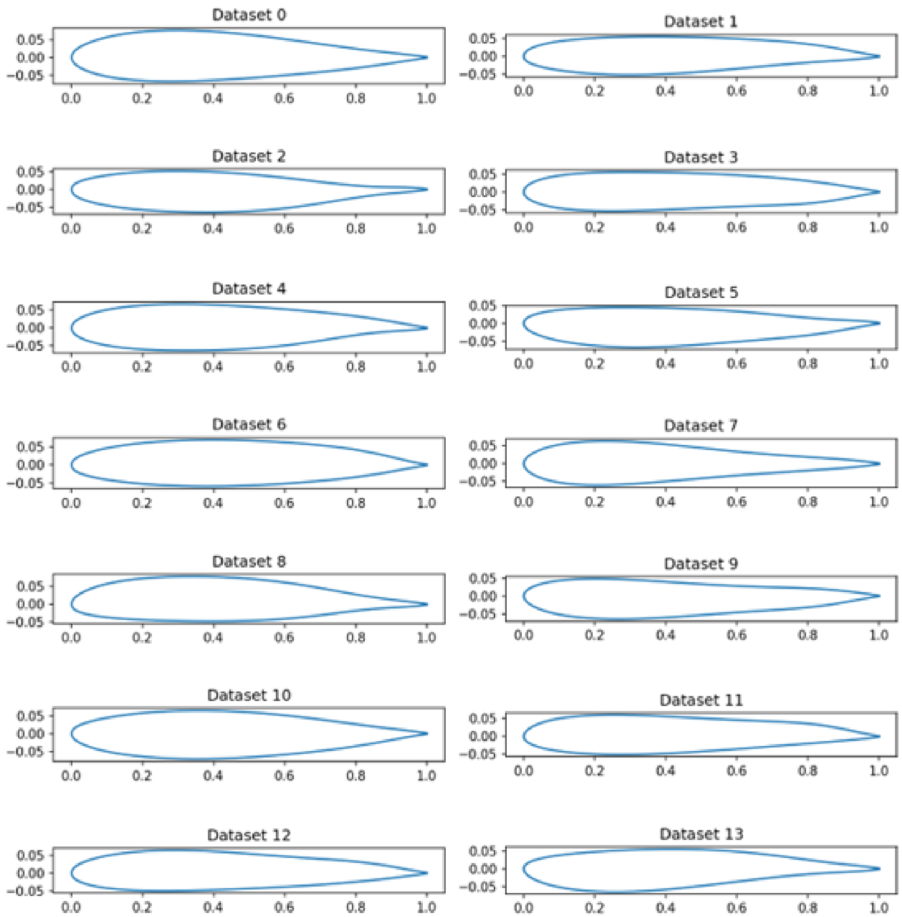

Fig. 2: Example of Deformed foils

followed by the iterative addition of a selectively chosen sample set.

1) Sobol Sampling: Sobol sequences [32] represent a type of quasi-random, low-discrepancy, and low-dispersion sequence. These sequences, which fall under Quasi-Monte Carlo methods, exhibit a faster convergence rate compared to standard Monte Carlo algorithms [33]. In this study, Sobol sequencing was chosen as one of the sampling methods due to its ability to evenly distribute points in space by minimizing empty regions. As previously mentioned, eight design variables were selected, representing the bump intensities of the Hicks-Henne function for the upper and lower surfaces of the foil. Using Python's Quasi-Monte Carlo submodule (scipy.stats.qmc), samples of design variables were generated (Figure 3).

2) Optimization-Based Adaptive Sampling: Optimization-based adaptive sampling is an iterative technique that begins by generating an initial set of samples using a space-filling method, such as Sobol sampling. A surrogate model is then built based on these samples. Due to the limited number of initial samples, the first surrogate model usually has low accuracy. To refine the initial surrogate model, it is optimized, even in its inaccurate state, to identify potential optima. To locate both global and local optima, multiple optimizations are performed using different initial points. Once the optima are found, new sample points are generated around these regions, forming clusters near the optima. These clusters are created using Sobol sampling within smaller, defined boundaries. The surrogate model is then updated with the newly acquired samples, and the process is repeated iteratively. The procedure continues until one of the following criteria is met: (1) the solution does not improve over a consecutive set of iterations, (2) the

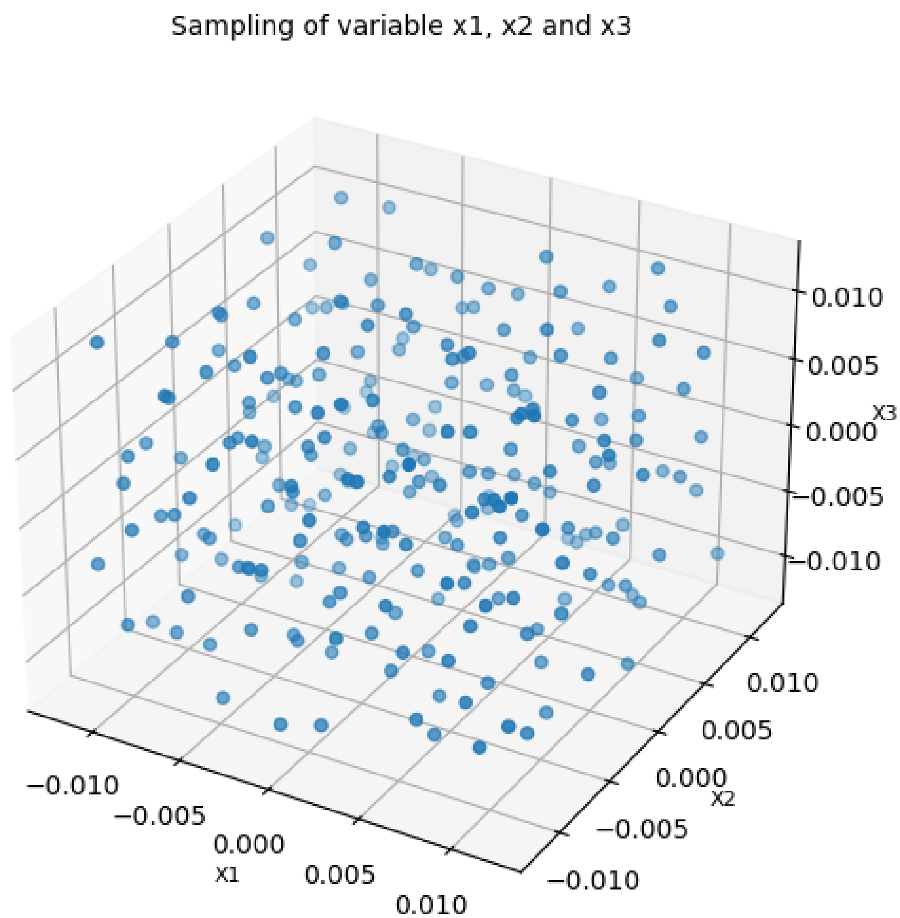

Fig. 3: Example of Sobol sampling of design variables

solution shows minimal variation with different initial points, or (3) a feasible solution is reached with sufficiently low loss. Figure 4 illustrates the overall process of optimization-based adaptive sampling for an unknown function.

C. Numerical Simulation

The numerical simulation of the base foil NACA0012 and the deformed foils was performed using an interactive program designed for the analysis and design of subsonic isolated airfoils, called XFOIL [19]. XFOIL can be used for both viscous and inviscid analyses, as well as for airfoil design and inverse design. The inviscid formulation utilizes a linear-vorticity panel method, while to account for viscous effects, XFOIL combines the potential flow solution with boundary layer analysis [34].

In this study, the shape of the NACA0012 airfoil was optimized under fixed flow conditions. The lift and drag coefficients of the base and deformed foils were extracted from XFOIL simulations using an automated Python script. The flow conditions were: $Re = 1 \times 10^6$, $\alpha = 5°$. The viscous solution iteration limit (ITER) in XFOIL was increased to allow up to 20 Newton iterations to ensure convergence. This adjustment was made to avoid missing results and non-converging calculations with more intricate foil designs. The number of panel nodes was kept at 160.

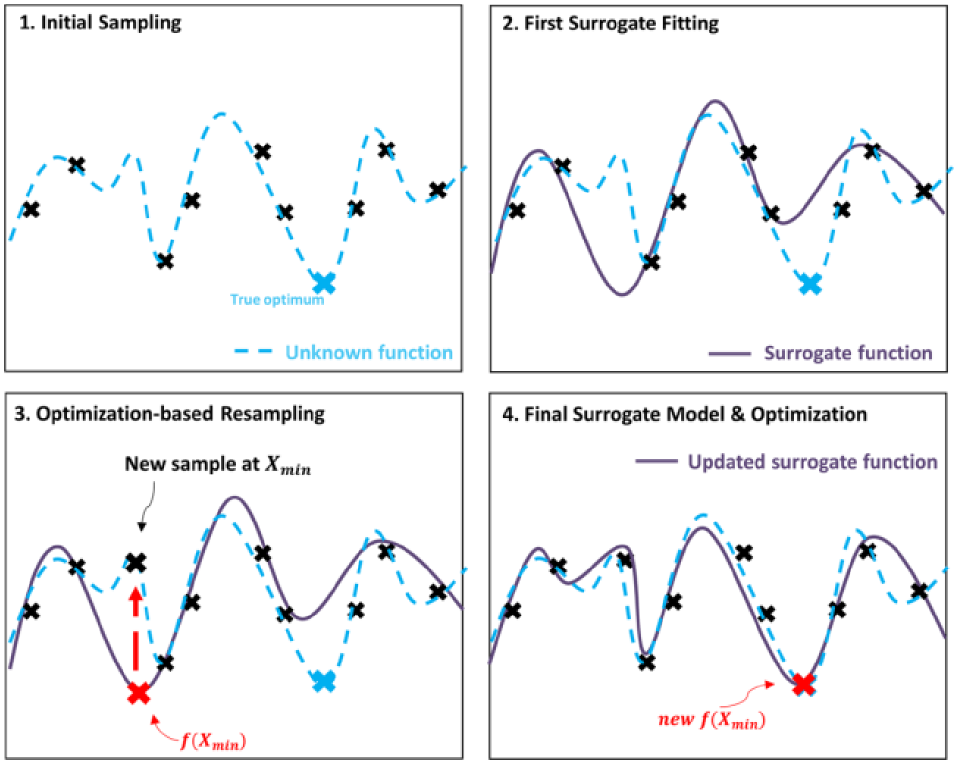

Fig. 4: Optimization-based adaptive sampling [22]

D. ANN Model Training

In this research, the surrogate model was developed by implementing an ANN model structured as a Feed Forward Neural Network, where each neuron is connected to every neuron in the next layer. The input neurons represent the design variable values, and the output neurons represent the lift and drag coefficient values. The model training was implemented in Python, utilizing the TensorFlow framework [35], with Keras [36] for the machine learning-specific components.

Figure 5 shows the dataset generation process for ANN-based surrogate model training. Before the training process, the data was cleaned to remove any missing values and duplicates. To improve learning performance, all inputs were normalized to fall within the range $[-1, 1]$ before the start of the training phase. The total dataset was divided into 5 folds for cross-validation. This cross-validation set was used to estimate the generalization error of the surrogate model. During the training process, the loss function was set to Mean Absolute Error (MAE), and the metrics evaluated were Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). The Adam optimizer was used with a learning rate of 0.01, and the tanh activation function was applied to the hidden layers. L2 regularization was implemented to prevent overfitting. The loss function values and metrics were evaluated by varying the number of neurons in the hidden layer using K-fold cross-validation. The architecture that produced the lowest loss was selected as the final surrogate model, as detailed in Table I. Although the feasibility of using a deep model was also investigated, the shallow model demonstrated superior performance. Consequently, the finalized ANN model consists of 8 input nodes representing the design variables, a single hidden layer with 14 nodes, and two output nodes representing the lift and drag coefficients. Figure 6 illustrates this ANN model architecture, and Figure 7 shows the validation curve of the model across 200 epochs, demonstrating that the loss steadily diminishes as the epoch number increases.

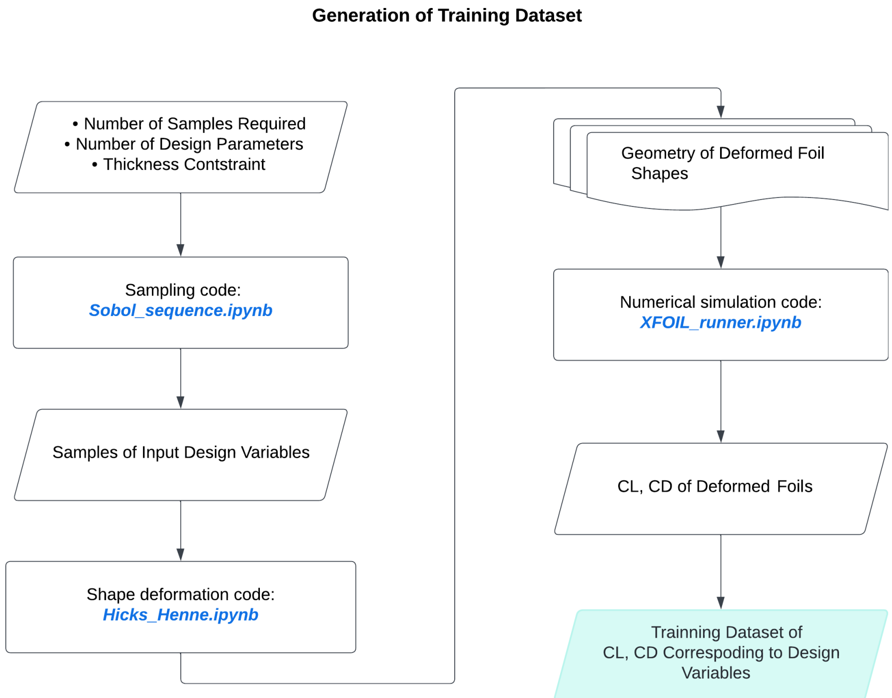

Fig. 5: Training dataset generation process

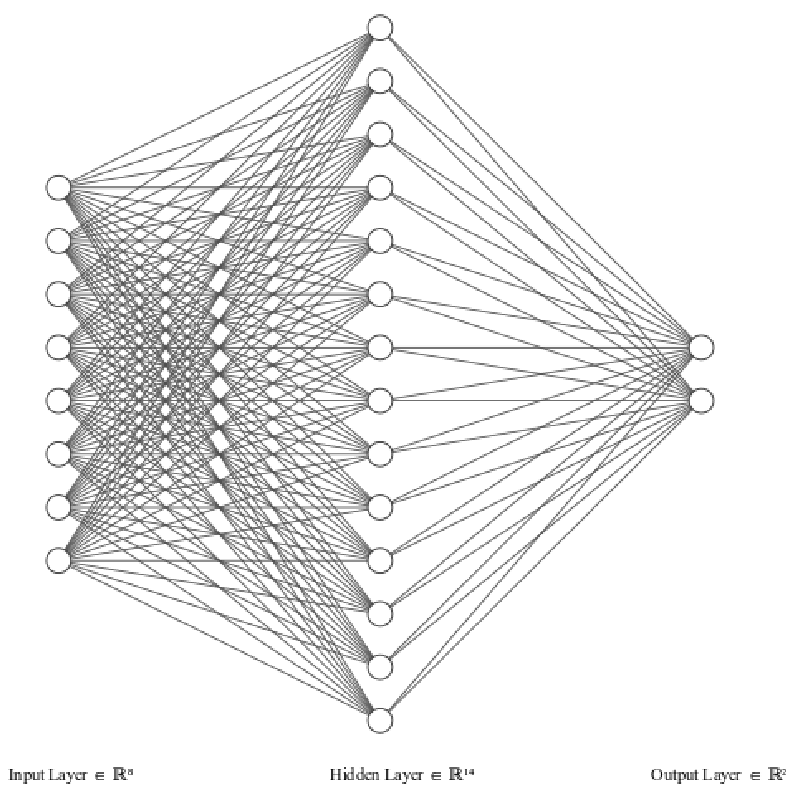

Fig. 6: ANN model for training

TABLE I: Training Parameters for ANN Model

Input Layer	Hidden Layer	Output Layer	Regularizer	Optimizer	Learning Rate
8	14×1	2	$L^2 : 0.01$	Adam	0.01
Loss Function	Metrics			Activation Function	Validation Set
MAE	MSE	RMSE	MAPE	Tanh	20%

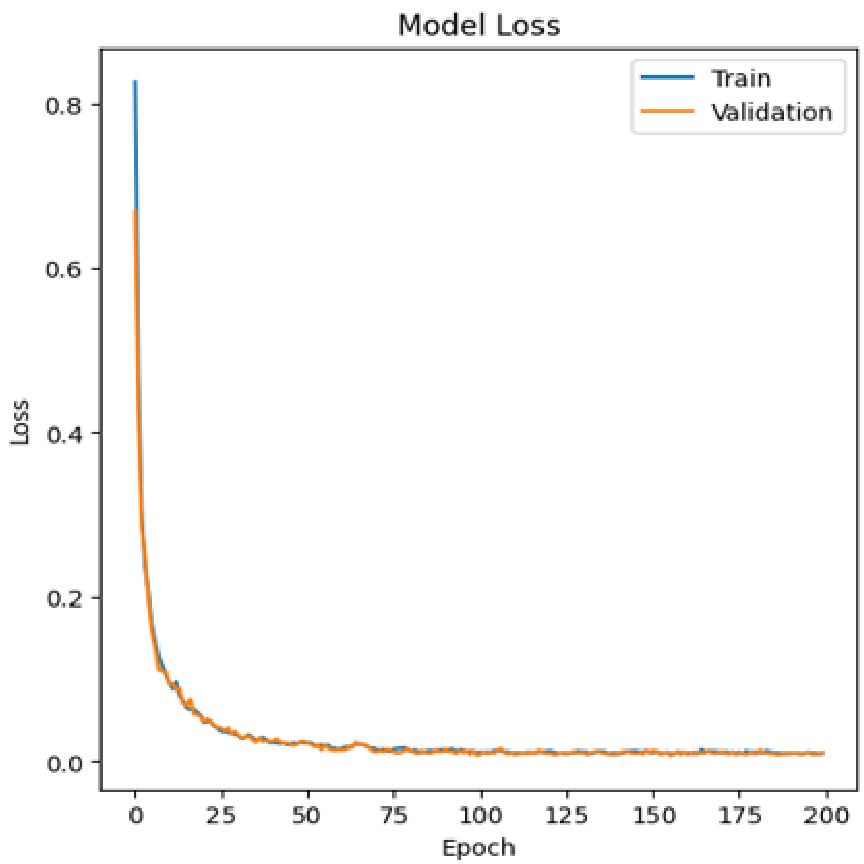

Fig. 7: Validation curve of MAE for 128 samples

E. Surrogate Model Optimization

Using the Python module scipy.optimize.minimize, the SLSQP optimizer was employed to implement the optimization of the surrogate model. Sequential Least Squares Programming (SLSQP) minimizes a function of several variables, allowing for any combination of bounds, equality, and inequality constraints. The objective of this optimization problem is to maximize the lift coefficient (C_L) while keeping the drag coefficient (C_D) constant and maintaining a constant area with respect to the design variables. As previously mentioned, the base foil is NACA0012, and the number of design variables is 8. To eliminate unrealistic airfoil shapes, the design variables are constrained to $\pm 10\%$ of the airfoil's thickness. We restrict the boundaries of each variable to $[-0.012, 0.012]$. The design space lies between the regions of the green and orange curves shown in Figure 8.

Objective: $\min(-C_L)$
Constraint 1: $C_D = C_{D_{\text{original}}}$
Constraint 2: $A = A_{\text{original}}$
Flow condition: $Re = 1 \times 10^6$, $\alpha = 5°$.

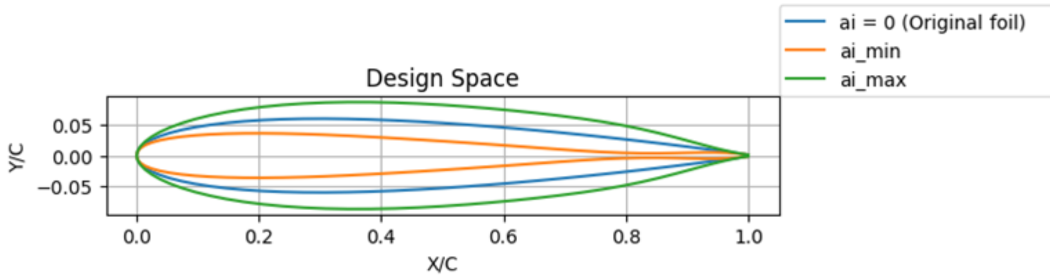

Fig. 8: Design space for optimization

1) Surrogate Model Optimization with Adaptive Sampling: In the case of the adaptive sampling method, the initial samples were generated using a low-discrepancy Sobol sequence. Using the Hicks-Henne function, the deformed foil geometry was generated based on the sample set of design variables. XFOIL simulations of the deformed foils were performed to extract the lift and drag coefficients (C_L, C_D). A training dataset of C_L and C_D, corresponding to the design variables, was created. An ANN-based surrogate model was then trained. Optimization of the surrogate model was performed to find both the local and global optima. In this case, three different optimizations were carried out using three initial points, which yielded different solutions. If the convergence criteria were not met, four more points around each optimum were added using Sobol, resulting in an additional 15 points for each iteration. The new points were added to the previous sample set, and the entire process was repeated until convergence. In this paper, 128 initial points were generated using Sobol sampling, and samples were added over three iterations, leading to a total of 173 samples by the final iteration. The process flowchart of this method is illustrated in Fig. 9.

2) Surrogate Model Optimization with One-shot Sampling: In the one-shot sampling case (Figure 10), the entire optimization is completed in a single sequential process. The total number of samples is generated at the initial stage using a space-filling sampling method, such as the Sobol sequence. Subsequently, shape parameterization, numerical simulation, dataset generation, model training, and optimization are performed sequentially. To compare with the adaptive sampling method, two optimizations were performed separately using 128 and 173 samples generated via Sobol sampling.

III. Results and Discussion

The Optimization-Based Adaptive Sampling strategy was compared with the conventional One-Shot Sobol Sampling approach to find the optimum solution while reducing training samples. Fig. 11 presents the lift coefficient (C_L), drag coefficient (C_D), and lift-to-drag ratio (C_L/C_D) for the baseline NACA0012 section, the ANN model trained with One-Shot Sampling, and the ANN model trained with Adaptive Sampling at the same final sample size of 173. The ANN model with Adaptive Sampling resulted in a 58.68% increased lift-to-drag ratio compared to the baseline foil section, while the One-Shot Sampling model achieved a 28.16% increase. Since both models used the same number of samples, this result shows that the adaptive method directed the optimizer toward a better optimum in the design space compared with the one-shot approach that only reached a local optimum. Because drag was explicitly constrained during optimization, the drag coefficient is nearly constant across the three cases. The slight reduction in

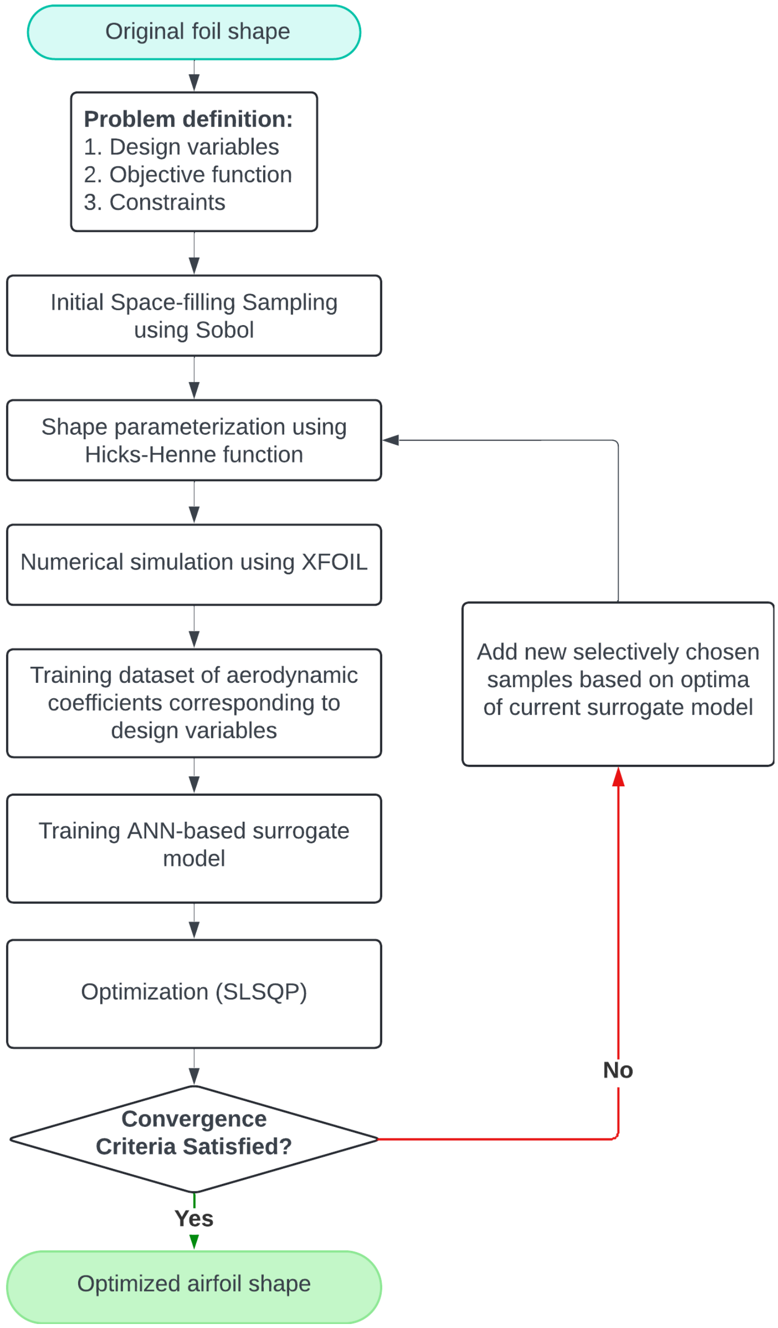

Fig. 9: Optimization Flowchart for Adaptive Sampling Method

the one-shot sampling model indicates its proclivity to underestimate the drag. Consequently, the increase in the lift-to-drag ratio chiefly follows the increase in lift coefficient, because the drag coefficient was nearly kept constant.

The optimized foil geometries are shown in Fig. 12. Both optimized shapes present significant deformation at the suction and pressure surfaces with respect to those of the baseline NACA0012 section, while maintaining overall section constraints. The resultant geometry acquired from Adaptive Sampling portrays a stronger shape modification in comparison with the One-Shot Sampling outcome, which is consistent with its higher lift coefficient. The improvement in lift is achieved without an increase in the material volume per unit span because the sectional area was constrained during optimization. This is useful for hydrofoil section design because a higher lifting capability is acquired while keeping the section

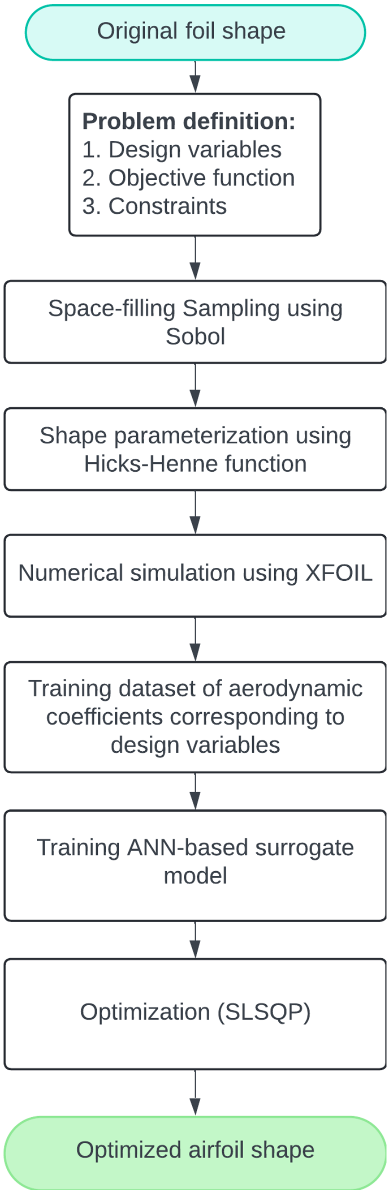

Fig. 10: Optimization Flowchart for One-shot Sampling Method

size and approximate material usage unchanged.

Fig. 13 compares the pressure coefficient distributions of the baseline and optimized sections. The pressure distribution explains the physical reason behind the lift improvement observed in Fig. 11. For adaptively sampled optimized section, the suction side pressure is reduced and the pressure on the lower surface is increased compared to the baseline section. This larger pressure difference between the two surfaces enhances the circulation around the foil section which in turn, results in a higher lift coefficient. The pressure redistribution is also consistent with the increased asymmetry observed in the optimized geometry. Since the Reynolds number and angle of attack were kept fixed throughout the optimization, this improvement in lift can be attributed to the geometric modification rather than to flow condition changes.

Surrogate model prediction at the optimum point has to be accurate. Fig. 14 shows the percentage error between the ANN-predicted optimum and the corresponding XFOIL calculated value. In the case of Adaptive Sampling, the error in drag coefficient reduces steadily as new samples are added around the

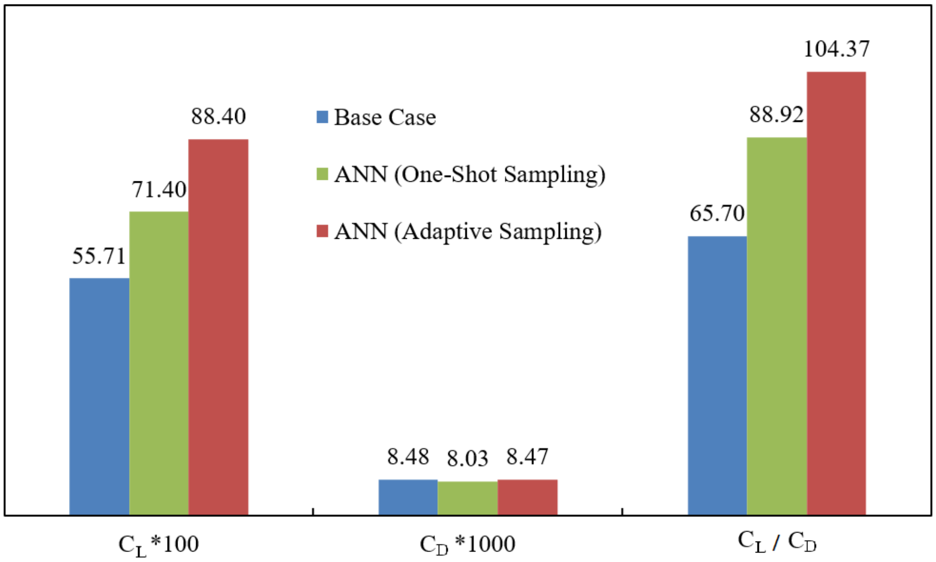

Fig. 11: Comparison of the hydrodynamic coefficients of the optimum for different sampling techniques

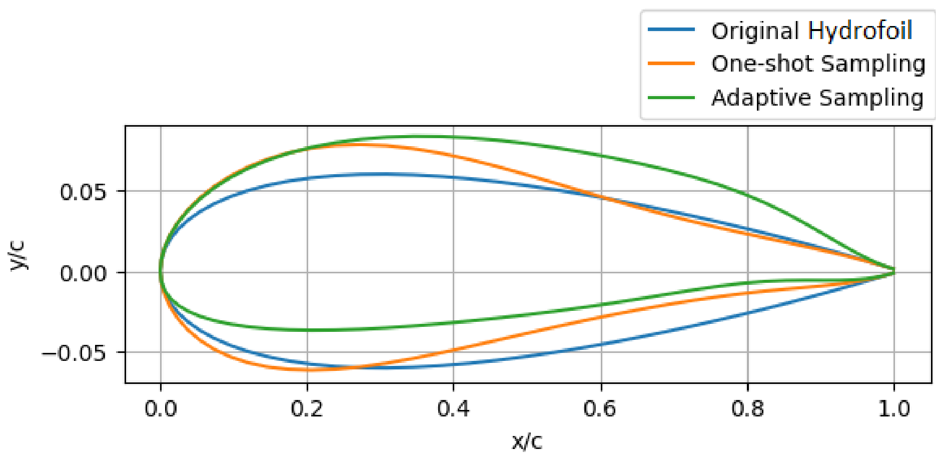

Fig. 12: Geometries of the baseline and the optimal foil

predicted optimum. It starts at about 22.5% for 128 samples, decreases to around 18% at 143 samples, then to about 7% at 158 samples, and finally becomes nearly negligible at 173 samples. The lift coefficient error follows a less monotonic trend, increasing from 5.33% to 9.75% in the first adaptive iteration, then dropping to 8.47%, before reducing to less than 2% at the final sample size. This final reduction indicates that the additional adaptively selected samples contribute to higher accuracy of the surrogate model near the optimum region.

On the other hand, even though the One-Shot sampling model is trained with the same total number of 173 samples, it does not reach the same level of optimum-point accuracy. The drag coefficient error falls down to 5.61%, but the lift coefficient error improves comparatively less for this case. This comparison illustrates that adaptive sampling usually does not just find a better optimum, but also improves accuracy of the surrogate model around the optimum. This is expected because Adaptive Sampling places new samples around promising optimum regions instead of uniformly distributing all samples across the design space.

The overall prediction performance of the ANN surrogate models using MAE, MSE, RMSE, and MAPE

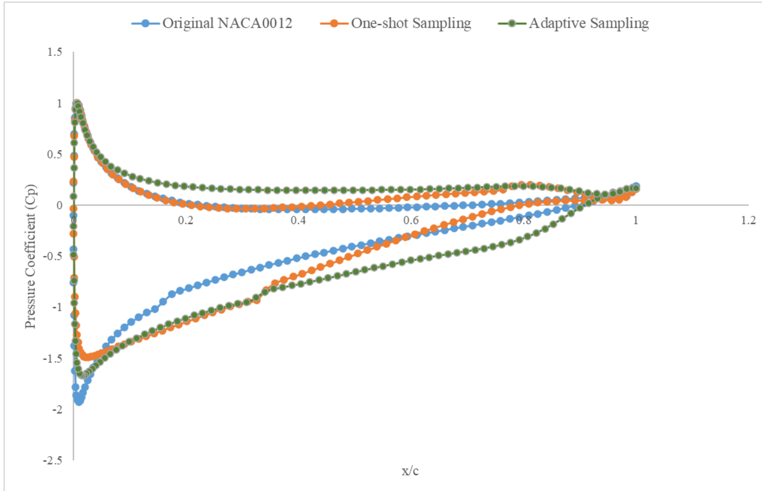

Fig. 13: Comparison of pressure coefficients between baseline and optimal foil

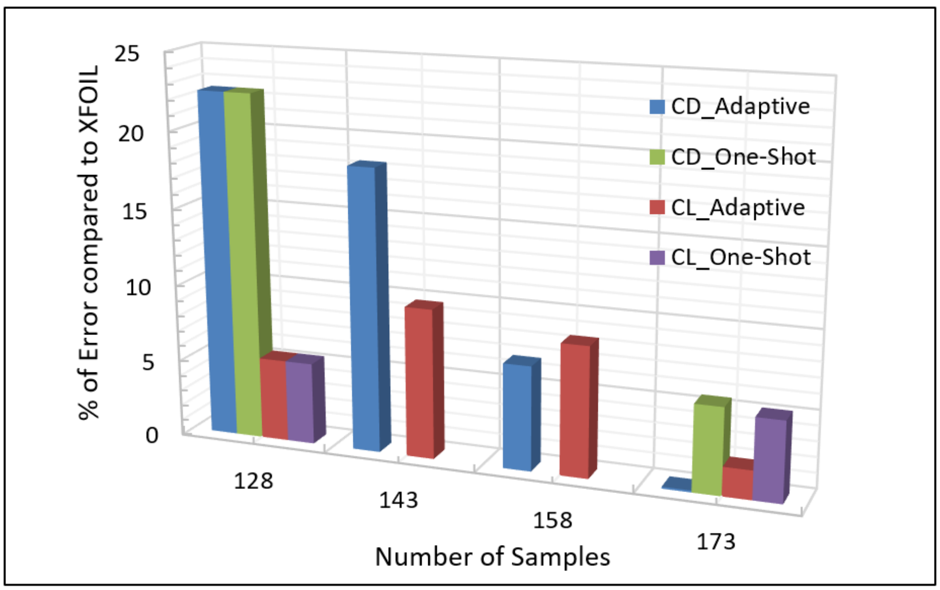

Fig. 14: Percentage of error of the predicted optimum compared to XFOIL

is reported in Table II. For the Adaptive Sampling model, the table reports the performance at four sample sizes: 128, 143, 158, and 173. The One-Shot Sampling model is reported at the final sample size of 173 for direct comparison. At 173 samples, the Adaptive Sampling model gives MAE $= 1.26 \times 10^{-2}$, MSE $= 3.89 \times 10^{-4}$, RMSE $= 1.97 \times 10^{-2}$, and MAPE $= 4.74\%$, and the One-Shot Sampling model gives MAE $= 1.35 \times 10^{-2}$, MSE $= 5.78 \times 10^{-4}$, RMSE $= 2.40 \times 10^{-2}$, and MAPE $= 5.42\%$. These values indicate that the overall prediction performance of both models over the full design space is similar, with the adaptive sampling model achieving higher accuracy in the optimal region.

The combined interpretation of Fig. 14 and Table II is important. Table II shows that the two surrogate models perform similarly over the design space, and Fig. 14 shows that the Adaptive Sampling model is significantly more accurate in the vicinity of the optimum point. Therefore, the benefit of Adaptive Sampling is concentrated in the most important region of the design space rather than being uniformly distributed everywhere. This agrees with the purpose of optimization-based adaptive sampling, where new samples are added near regions that are more likely to contain the optimum instead of being spread equally

TABLE II: Performance Metrics of ANN Model

Iteration	ANN with Adaptive Sampling	Sample count	MAE	MSE	RMSE	MAPE
0	Sobol (one-shot)	128	1.21E-02	3.30E-04	1.82E-02	1.27E+01
1	Sobol + selective (Adaptive)	143	1.27E-02	4.94E-04	2.22E-02	4.57E+00
2	Sobol + selective (Adaptive)	158	1.24E-02	3.52E-04	1.87E-02	6.79E+00
3	Sobol + selective (Adaptive)	173	1.26E-02	3.89E-04	1.97E-02	4.74E+00
Case	**ANN with One-Shot Sampling**	**Sample count**	**MAE**	**MSE**	**RMSE**	**MAPE**
1	Sobol (one-shot)	173	1.35E-02	5.78E-04	2.40E-02	5.42E+00

across less useful regions. This is consistent with the goal of optimization-based adaptive sampling; instead of being uniformly distributed across less useful regions, new samples are added predominantly in areas that are more likely to contain the optimum.

Overall, these results suggest that Optimization-Based Adaptive Sampling is a much superior surrogate-assisted foil optimization framework to One-Shot Sobol Sampling for the same limited sampling budget. The adaptive model finds a better optimum, improves lift by 58.68% with respect to the baseline section, maintains the drag and sectional area constraints, and predicts the optimum with higher accuracy. This is important for early-stage hydrofoil section design since each additional numerical evaluation requires time and computational effort. Therefore, the proposed ANN-based surrogate model with Adaptive Sampling provides a feasible and resource-efficient means of improving hydrofoil lift performance under fixed design constraints.

IV. Conclusions

This study explores machine learning techniques to tackle design optimization problems, which, in practical cases, can be used to significantly reduce the computational cost and time of hydrofoil craft design, specifically, in developing economies like Bangladesh where computational resources are scarce. In this paper, an ANN-based surrogate model optimization framework has been developed to enhance foil efficiency by maximizing lift while constraining drag and sectional area to remain constant. Moreover, two sampling strategies were investigated: One-Shot Sampling and Optimization-Based Adaptive Sampling. Both sampling strategies resulted in a significant increase in lift and hydrodynamic efficiency. However, for the same number of samples, the Optimization-Based Adaptive Sampling was able to find the global maximum, whereas the One-Shot Sampling model remained adhered to a local maximum. The ANN model with Adaptive Sampling led to an optimized foil shape that exhibits 58.68% more lift relative to the base case. This selective sampling method considerably increased the accuracy of the surrogate model at the optimum point. In contrast, the One-Shot Sampling did not achieve the same level of accuracy at the optimum point. However, both models showed similar performance across the entire design space.

Although the proposed sampling technique performed better than the conventional sampling method, this study has several limitations that need to be addressed. While the study aimed to reach the global maximum, it has not been verified whether the true global maximum was achieved. The computational speed of the surrogate model relative to traditional methods has also not been quantified. The optimization was carried out for a single base hydrofoil under a single fixed flow condition, without exploring broader design spaces, multiple base foil geometries, or a range of operating conditions. Furthermore, the framework constrains drag and sectional area only, omitting practically relevant constraints such as cavitation and structural feasibility. Finally, the results rely on XFOIL, a low-fidelity solver, and have not been validated against high-fidelity CFD simulations or experimental data.

Future studies will verify that the true global optimum is obtained by comparing results against an exhaustive search of the design space. The computational speed of the surrogate model will also be measured directly and compared to traditional optimization methods. Future studies will validate the optimized foil shapes through experimental testing and CFD simulations. This framework can also be extended to a wider design space, including more base foil shapes and different operating conditions, such as varying Reynolds numbers and angles of attack. Finally, additional constraints, such as cavitation and structural limits, can be incorporated to make this framework more applicable in real-world scenarios. Ultimately, this framework will provide naval architects and ship designers with a cost-effective screening tool for evaluating candidate hydrofoil sections prior to committing to expensive CFD analyses or experimental testing.

V. Acknowledgement

First of all, we would like to convey our heartiest thanks to our faculty advisor, Professor Dr. Zobair Ibn Awal, for his valuable guidance, support, and encouragement throughout this project. We also thank the Bangladesh University of Engineering and Technology (BUET) for providing us with the necessary facilities and resources. Finally, we would like to extend our sincere thanks to everyone who helped in the successful completion of this work, both directly and indirectly.

References

[1] A. Steinert, A. Hug, P. Medl, T. Schulze, and F. Holzapfel, "Modeling and dynamic analysis of a "flying" fully submerged hydrofoil craft for control," *Aerospace Systems*, 2026. [Online]. Available: https://doi.org/10.1007/s42401-025-00445-y

[2] Y. Liao, A. Yildirim, J. R. Martins, and Y. L. Young, "RANS-based optimization of a T-shaped hydrofoil considering junction design," *Ocean Engineering*, vol. 262, p. 112051, 2022. [Online]. Available: https://doi.org/10.1016/j.oceaneng.2022.112051

[3] R. Tannenberg, K. Hochkirch, A. Walther, S. R. Turnock, and S. W. Boyd, "Development of an adjoint system-based hydrofoil optimisation framework using algorithmic differentiation," *Optimization and Engineering*, vol. 27, pp. 1397–1423, 2026. [Online]. Available: https://doi.org/10.1007/s11081-025-10055-4

[4] N. Garg, G. K. Kenway, J. R. Martins, and Y. L. Young, "High-fidelity multipoint hydrostructural optimization of a 3-d hydrofoil," vol. 71, pp. 15–39. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S088997461630144X

[5] N. Garg, B. W. Pearce, P. A. Brandner, A. W. Phillips, J. R. Martins, and Y. L. Young, "Experimental investigation of a hydrofoil designed via hydrostructural optimization," vol. 84, pp. 243–262. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0889974617307934

[6] J. E. Kerwin, "13.04 LECTURE NOTES HYDROFOILS AND PROPELLERS."

[7] X. Xie and R. Tao, "Optimization of hydrofoil leading-edge geometry for delaying incipient cavitation using a CFD–GA approach," vol. 109, no. 2, p. 00368504261458949. [Online]. Available: https://journals.sagepub.com/doi/10.1177/00368504261458949

[8] G. W. Ng, A. Yildirim, H. Youngren, A. Lamkin, and E. Jonsson, "Design optimization of america's cup AC75 hydrofoil sections with flaps."

[9] G. W. Ng, Y. Liao, A. Yildirim, and J. R. Martins, "Hydrostructural optimization of subcavitating cambered and symmetric composite foils," vol. 351, p. 118545. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0263822324006731

[10] K. V. Kostas and M. Manousaridou, "Machine-learning-enabled foil design assistant," vol. 11, no. 7, p. 1470. [Online]. Available: https://www.mdpi.com/2077-1312/11/7/1470

[11] J. M. Walker, A. Coraddu, and L. Oneto, "A review on shape optimization of hulls and airfoils leveraging computational fluid dynamics data-driven surrogate models," *Ocean Engineering*, vol. 312, p. 119263, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0029801824023618

[12] O. Baysal, "Aerodynamic shape optimization: Methods and applications," *SAE Transactions*, vol. 108, pp. 794–802, 1999. [Online]. Available: http://www.jstor.org/stable/44729475

[13] S. Jeong, M. Murayama, and K. Yamamoto, "Efficient optimization design method using kriging model," *Journal of Aircraft*, vol. 42, no. 2, pp. 413–420, 2005.

[14] Q. Wang, W. Qian, and K. He, "Unsteady aerodynamic modeling at high angles of attack using support vector machines," *Chinese Journal of Aeronautics*, vol. 28, no. 3, pp. 659–668, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1000936115000552

[15] D. F. Kurtulus, "Ability to forecast unsteady aerodynamic forces of flapping airfoils by artificial neural network," *Neural Computing and Applications*, vol. 18, pp. 359–368, 2009.

[16] H. Pehlivan Solak, J. Wackers, R. Pellegrini, A. Serani, M. Diez, P. Perali, M. Sacher, J.-B. Leroux, B. Augier, F. Hauville, and P. Bot, "Hydrofoil optimization via automated multi-fidelity surrogate models," in *Proceedings of the International Conference on Hydroelasticity in Marine Technology*, 2023.

[17] S. Van Reen, B. Serbülent, and H.-D. Yao, "Machine learning-based multipoint optimisation for improving aerodynamics of symmetrically cambered wing sails in wind-assisted ship propulsion," vol. 342, p. 122829. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0029801825025120

[18] J. Jung and G. X. Gu, "Data-driven airfoil shape optimization framework for enhanced flutter performance," vol. 36, no. 10, p. 101706. [Online]. Available: https://pubs.aip.org/pof/article/36/10/101706/3316363/Data-driven-airfoil-shape-optimization-framework

[19] M. Drela, "Xfoil: An analysis and design system for low reynolds number airfoils," vol. 54, 06 1989.

[20] J. Morgado, R. Vizinho, M. Silvestre, and J. Páscoa, "XFOIL vs CFD performance predictions for high lift low reynolds number airfoils," vol. 52, pp. 207–214. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1270963816300839

[21] C. Zhang, Z. Wei, J. Wu, G. Li, and K. Zhu, "Application of shallow learning in aerodynamic prediction of wind tunnel test," in *2023 2nd International Conference on Machine Learning, Cloud Computing and Intelligent Mining (MLCCIM)*, 2023, pp. 6–17.

[22] S. H. Kim and F. Boukouvala, "Machine learning-based surrogate modeling for data-driven optimization: a comparison of subset selection for regression techniques," *Optimization Letters*, vol. 14, no. 4, pp. 989–1010, 2020.

[23] J. Huang, Z. Gao, Z. Zhou, and K. Zhao, "An improved adaptive sampling and experiment design method for aerodynamic optimization," *Chinese Journal of Aeronautics*, vol. 28, no. 5, pp. 1391–1399, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1000936115001570

[24] J. Li, M. Zhang, J. Martins, and C. Shu, "Efficient aerodynamic shape optimization with deep-learning-based geometric filtering," *AIAA Journal*, vol. 58, pp. 1–17, 07 2020.

[25] S. Wang, Q. Xu, and N. Wei, "Airfoil aerodynamic optimization design using ensemble learning surrogate model," *Journal of Aerospace Engineering*, vol. 37, no. 6, p. 04024093, 2024. [Online]. Available: https://ascelibrary.org/doi/abs/10.1061/JAEEEZ.ASENG-5410

[26] World Bank, "Better inland waterways leading to more domestic and regional connectivity," Feature Story, The World Bank, Jul. 2019, accessed: June 10, 2026. [Online]. Available: https://www.worldbank.org/en/news/feature/2019/07/08/better-inland-waterways-leading-to-more-domestic-and-regional-connectivity

[27] L. Farrell, X. H. Chiu, E. Nora, and R. Rohatgi, "Reviving Bangladesh's in-land waterways for low-carbon and resilient transport," Feature Story, The World Bank, Jul. 2020, accessed: June 10, 2026. [Online]. Available: https://www.worldbank.org/en/news/feature/2020/07/13/reviving-bangladeshs-in-land-waterways-for-low-carbon-and-resilient-transport

[28] *Report on research and development (r & d) survey 2022*. Bangladesh Bureau of Statistics (BBS), Statistics & Informatics Division (SID), Ministry of Planning, Government of the People's Republic of Bangladesh, OCLC: 1533589449.

[29] R. M. Hicks and P. A. Henne, "Wing design by numerical optimization," *Journal of Aircraft*, vol. 15, no. 7, pp. 407–412, 1978. [Online]. Available: https://doi.org/10.2514/3.58379

[30] P. Castonguay and S. Nadarajah, *Effect of Shape Parameterization on Aerodynamic Shape Optimization*. [Online]. Available: https://arc.aiaa.org/doi/abs/10.2514/6.2007-59

[31] M. Zanichelli, "Shape optimisation of airfoils by machine learning-based surrogate models - master's thesis - marco zanichelli - politecnico di milano," 12 2021.

[32] I. M. Sobol', "Distribution of points in a cube and the approximate evaluation of integrals (in Russian)," *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, vol. 7, pp. 784–802, 1967.

[33] F. Y. Kuo, W. T. M. Dunsmuir, I. H. Sloan, M. P. Wand, and R. S. Womersley, "Quasi-monte carlo for highly structured generalised response models," *Methodology and Computing in Applied Probability*, vol. 10, no. 2, pp. 239–275, Jun. 2008. [Online]. Available: https://doi.org/10.1007/s11009-007-9045-3

[34] M. Drela and M. GILES, "Viscous-inviscid analysis of transonic and low reynolds number airfoils," *Aiaa Journal - AIAA J*, vol. 25, pp. 1347–1355, 10 1987.

[35] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, software available from tensorflow.org. [Online]. Available: https://www.tensorflow.org/

[36] F. Chollet *et al.*, "Keras," https://keras.io, 2015.