A REDUCED ORDER DATA-DRIVEN METHOD FOR RESISTANCE PREDICTION AND SHAPE OPTIMIZATION OF HULL VANE

Cihad Çelik[a,*], Devrim Bülent Danışman[a], Shahroz Khan[b], Panagiotis Kaklis[b]

[a] *Istanbul Technical University, Faculty of Naval Architecture and Ocean Engineering, 34469, Istanbul, Turkey*

Email addresses: celikciha@itu.edu.tr; bulent.danisman@itu.edu.tr

[b] *Department of Naval Architecture, Ocean and Marine Engineering, University of Strathclyde, Henry Dyer Building, 100 Montrose Street, Glasgow G4 0LZ, Scotland, UK*

Email addresses: shahroz.khan@strath.ac.uk; panagiotis.kaklis@strath.ac.uk

Corresponding Author: Cihad ÇELİK[a,]

[a] *Istanbul Technical University, Faculty of Naval Architecture and Ocean Engineering, 34469, Istanbul, Turkey*

E-mail address: celikciha@itu.edu.tr

ABSTRACT

Hull Vane (HV) is an energy-saving appendage introduced by Hull Vane BV company to reduce total ship resistance. Shapewise, HV is a hydrofoil wing transversely fixed at the transom bottom of the hull.

In this paper, a data-driven shape optimization method is proposed for HV. To avoid the time-consuming resistance evaluation of designs via a viscous flow solver, we develop a Machine-Learning (ML) based model that predicts the hull's total resistance in the presence of an HV. For this purpose, Principal-Component Analysis (PCA) is first implemented to reduce the dimensionality of the problem, and then the prediction model is trained with the most influential of the Principal Components (PCs). Given that these PCs capture the maximum geometric variance of the original design space, higher accuracy can be achieved at the expense of a few training samples. After the training phase, the model is integrated with an optimizer, which searches in a dimensionally-reduced design space for the optimal design of the HV. The obtained results achieved a 70% dimensionality reduction with the aid PCA and an approximately 98% accuracy for predicting total resistance. Compared with the reference HV, the optimized one reduced the total resistance by 1.2%.

1. INTRODUCTION

The societal awareness of the need to reduce carbon emissions is constantly increasing. New restrictive regulations on carbon emissions will be implemented by the International Maritime Organization (IMO) in the coming years. Over the last decade, significant attention has been devoted to the reduction of fuel consumption by ships, which can be achieved by employing various approaches, such as using alternative energy resources, decreasing hull resistance and increasing the efficiency of the ships' main engine. In this work, we focus on reducing the hulls' total resistance by using energy-saving appendages.

Such appendages have been widely used to decrease a ship's resistance by controlling the direction of streamlines and advantageously changing the pressure distribution on the ship hull. One of these appendages is Hull Vane (HV), which is a hydrofoil wing transversely fixed at the transom bottom of ships. Its working principle is based on the fact that the negative pressure zone on the suction side of the HV absorbs the high-pressure zone behind the ship, and as a result, it reduces the ship's hull resistance. The overall design of an HV takes into consideration several parameters, such as position, angle of attack, span- and chord lengths, which must be optimized to reduce the hull resistance further. The inventors of HV have conducted various experimental and computational studies **(Uithof et al., 2017; Hou et al., 2020)**, but these studies are mostly kept confidential; therefore, no detailed research on the design of HV is available in the literature.

A range of computational tools, including potential-flow solvers and viscous-flow solvers,is used to evaluate engineering models' physical properties, such as the total resistance of hull forms and the drag of cars and aeroplanes. Potential-flow codes cannot, however, interpret viscous effects. Therefore, viscous-flow solvers, based on Reynolds-averaged Navier–Stokes (RANS) equations, are adopted, but solving these nonlinear equations for complex free-form shapes, such as ship hulls, poses a challenge due to their computational complexity, especially when these solvers have to be used in a shape-optimization loop. Different approaches have been developed in this connection to overcome the computational burden caused by these solvers during an optimization process. Generally, the problem of computational cost is tackled in two different research directions. One approach aims to reduce the complexity by investigating the structure of the design space via sensitivity analysis or reduced-order representation. The other approach appeals to data-driven model-agnostic surrogate methods **(Diez et al., 2019)** or multi-fidelity metamodels **(Serani et al., 2019)** in conjunction with

models which couple physics with high-fidelity geometrical models, such as isogeometric analysis (IGA) **(Kostas et al., 2015)**, in order to achieve accuracy at a reasonable cost.

Recently, the expansion of Industry 4.0 has leveraged data-driven techniques, such as Machine Learning (ML), which have been proven capable to bypass the need for Computational Fluid Dynamics (CFD) solvers when evaluating, e.g., the resistance of a ship-hull **(Margari et al., 2018; Yu et al., 2019; Danışman, 2014)**, the drag coefficients of a car **(Gunpinar et al., 2019)** or an aerofoil design **(Li et al., 2019)**, within a specified accuracy level. In engineering design, despite the proven efficiency of these techniques, the problem of high computational cost still exists, given that training precise ML-based prediction models requires large training datasets containing both the design's parametric representation and its physical results **(Masood et al., 2021)**. Therefore, some researchers **(Hamdia et al., 2019)** use IGA to evaluate and construct large datasets for ML. However, for free-form shapes, IGA provides a combination of low-fidelity physics models with high-fidelity geometrical models, which may result in the construction of an ML model with a relatively poor prediction efficiency.

In this work, we propose a data-driven shape optimization approach to optimize the cross-sectional profile of the HV of a given motor yacht in order to reduce its overall resistance. First, the parametric profile of HV is constructed based on a reference shape, namely the NACA4412 foil, which was parameterized using **Kostas et al.'s (2017**) technique using eight shape parameters. To start the optimization, we first constructed a design space using the upper and lower bound of the design parameters, which are set around the reference HV profile. A subspace representation of the original design space is then created using Principal Component Analysis (PCA) **(Wold et al., 1987)**, which not only facilitates the optimizer to avoid excessive exploration of design space for global optima but also cancels the need of constructing a large training dataset for relatively precise prediction models.

Subsequently, designs from the subspace are sampled to generate a training dataset, and the total resistance for these designs is evaluated using a viscous flow solver after projecting them back onto the original design space. The design sampling in the original and the subspace is performed using the so-called *Latin Hypercube Sampling* **(Stein, 1987)**, which ensures even distribution of Design of Experiments (DoE). The dataset used to train the ML-based prediction model contained Principal Components (PCs) as independent parameters and their resistance as a dependent parameter. Finally, the prediction model is integrated with an optimizer, which explores the subspace to find the optimal cross-sectional profile of the HV in order to reduce

the total hull resistance. **Figure 1** shows the overall workflow of the proposed optimization framework.

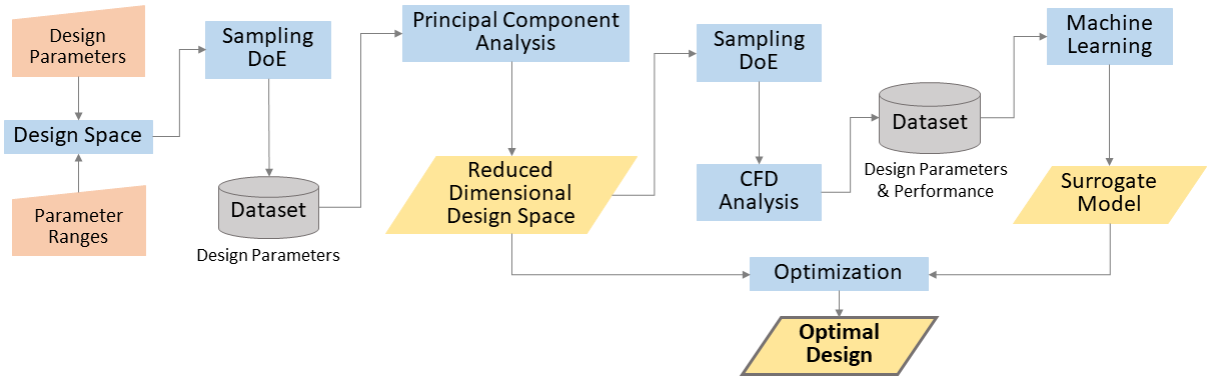

Figure 1: Overall workflow of the data-driven approach used for the HV optimization.

The rest of the paper is structured as follows. Section 2 gives an overview in the area of HV design and data-driven techniques used in maritime and other engineering fields. The Numerical modelling and the CFD setting for the problem in question are presented in Section 3. Section 4 then provides details on the construction of a data-driven prediction model and optimization. Results of CFD analysis, prediction model and optimization are presented in Section 5. Finally, concluding remarks and directions for future work are discussed in Section 6.

2. RELATED WORKS

2.1 Hull Vane

Uithof et al. (2014) described the advantages of HV as additional thrust force, trim correction, reduction of waves behind the vessel and motion damping in head seas. Once the horizontal component of the lift force of HV is greater than the horizontal component of the drag force, the resulting horizontal force then provides an additional thrust force, and the resulting vertical force changes the vessel's trim, thereby advantageously affecting the total resistance of the vessel. A negative pressure zone, which helps to reduce the stern wave, appears on the suction side of the HV due to the accelerated flow from the aft of the hull. **Uithof et al. (2014)** also examined the effect of HV located under the ship's hull, which created a negative pressure zone on its suction side, resulting in an additional pressure resistance on the hull HV was installed behind the transom of the ship. The position of HV has also optimized with a conclusion that a

sufficient vertical distance is essential in order to prevent the flow interaction between the HV and the ship's hull.

A series of towing tank experiments were conducted by **Hou et al. (2020)** on DTMB 5415 hull to study the effect of the angle of attack and location of the HV in calm water and head sea conditions. This study showed that the angle of attack is more influential than the HV location in reducing the trim response by 26% at Froude number (Fr)=0.413. **Celik et al. (2020)** performed a numerical study on the impact of HV on a motor yacht hull and the influence of its chord length on the model scale. It was observed that HV had more effect on wave resistance among the resistance components, which was reduced by 31% compared to the bare hull. The chord length of the HV was varied from 1.75% to 3.5% of the waterline length, and it was noted that the total resistance of the motor yacht decreased as the chord length increased. However, the large chord length of HV had an adverse effect on its structural strength.

Along with resistance, HV also affects the seakeeping characteristics of a ship. For example, **Bouckaert et al. (2016)** performed CFD based experiments on a 108m Holland Class OPV ship in order to investigate HV's effect on the seakeeping. Both HV installed, and bare hulls were tested at 2m and 4m wave heights with a wave period of 8 seconds. In the presence of HV, the vertical acceleration on the stern area was decreased by 13.1%, and 11.7% at 2m and 4m wave heights, respectively. Moreover, the pitch motion acceleration was reduced by 8.1% and 6.8% and the added resistance, which is the difference between the average resistance in waves and the resistance in calm water, was also decreased by 5.7% and 4.9%, respectively, in 4m and 2m wave height. In another study, **Uithof et al. (2016a)** also examined the effects of HV on the seakeeping behaviour of ferries and RoPax vessels. The results of this study showed that the HV reduced the roll motion by 0.7% and increased the natural period of pitch motion by 7.0%.

Uithof et al. (2017) also studied the influence of HV, interceptors, trim wedges and ballasting systems on a 50m AMECRC Series #13 patrol vessel with CFD simulations, which were carried out without the struts of the HV as they increase the number of computational meshes and have little effect on the total resistance. It was observed that the change in the position of HV in the vertical direction had a slight effect on the total resistance. **Hagemeister et al. (2017)** compared the effects of the ship's overall length and HV on the annual fuel consumption. Results showed that the total annual fuel saving for a ship with an HV installed was 15.1%, while the gain for the extended ship hull was 6.4%. **Bouckaert et al. (2016)** performed a life-cycle cost analysis of an offshore patrol vessel. In this study, HV was tested at speed with the highest annual fuel consumption. CFD simulations were conducted with struts and actuator disks in order to model

both the real configuration of HV and the propellers, respectively. As a result of this study, the annual fuel cost was decreased by 12.5%.

In addition to the effect of HV on the resistance and seakeeping, some of its unquantifiable contributions and benefits have been demonstrated by **Uithof et al. (2016b)**. These can be listed as reducing the size of the engine room, the initial investment costs with lower engine power, the size and costs of auxiliary machines and the tank volumes, which helped to create more usable space.

2.2 Data-Driven Methods

In the last decade, data-driven methods have gained tremendous attention in various fields of science and engineering. Multiple techniques have been proposed, modified, and implemented on a different set of problems in these fields, which otherwise were difficult, if not impossible, to solve. In this section, we discuss some of the existing data-driven techniques, closely related to the present work, proposed and implemented to reduce the computational cost of *shape optimization*. In particular, we mainly focus on reviewing applications of parametric dimension reduction and surrogate modelling with ML in the maritime industry.

2.2.1 Dimension reduction in shape optimization of ships

In shape optimization design, parametrization plays an important role, which can be performed with various techniques **(Samareh, 2001)**. The n number of the design parameters representing a shape defines the dimensionality of the design space, which provides a search domain in $\mathbb{R}^n$ for optimizers to find an optimal design. The dimensionality of the design space plays a crucial role in optimization, as high-dimensional spaces possess a high potential for finding the most optimal design. However, the high-dimensional design space causes high computational cost, which increases exponentially with dimensionality. Therefore, parametric dimension reduction techniques have been widely utilized, which generally identify a lower-dimensional latent space that captures most of the design space's design variability. The most commonly used dimension-reduction techniques are the Karhunen–Loève expansion (KLE) **(Fukunaga and Koontz, 1970)** (whose discrete representation is PCA), the active-subspaces method **(Constantine et al., 2015)**, and the autoencoder approach **(Hinton and Salakhutdinov, 2006)**.

In the area of naval architecture and marine engineering, **Diez et al. (2015)** proposed a dimension-reduction technique based on KLE/PCA to design a high-speed catamaran to reduce the wave component of calm-water resistance. **D'Agostino et al. (2020)** used KLE for the assessment of geometric variability retained by a different type of shape parameterization

methods such as Free-Form Deformation (FFD), Radial Basis Functions (RBF) and Global Modification Functions (GMF), used in shape optimization of the ship hull. It was observed that the highest design space dimensionality reduction was achieved by FFD, followed by the GMF and RBF, while retaining 95% of the geometric variability of the original design space. Design space dimension reduction-based shape optimization of marine propellers was performed by **Gaggero et al. (2019)**. In this work, a 23-dimensional design space was created to include the INSEAN-E779A propeller, whose dimensionality was later reduced to 15 while retaining 98% of the geometric variance.

Autoencoder is an ML-based nonlinear dimension reduction that extracts the latent lower-dimensional manifolds of a high-dimensional design space. **D'Agostino et al. (2018)** implemented the autoencoder approach in the context of shape optimization of the USS Arleigh Burke-class destroyer for total resistance and compared autoencoder results with PCA. Results showed that a higher dimensionality reduction of 56% was achieved with autoencoders, where PCA was only able to provide 22% of dimensionality reduction. **Serani et al. (2019)** also compared the performance of nonlinear kernel-based PCA with typical PCA. They showed that even though high-dimension reduction can be achieved with a nonlinear version, it cannot provide the most optimal design.

Compared to other approaches, the method of active subspaces is a comparatively new approach that has been proven promising for various applications. It builds a lower-dimensional representation based on the fact that the functional performance does not vary the same along each dimension of the design space. Therefore, it extracts the *latent directions* of the design space, which substantially varies the design's performance. Design directions are extracted using the gradients of the function representing the optimization criterion. **Tezzele et al. (2018)** used this technique to optimize the US Navy Combatant DTMB 5415 for wave resistance, which was parameterized via FFD with eight parameters. The results of the active-subspace approach showed that the latent direction was sufficient to model the wave resistance of the tested model with an error of only 4.5%. Recently, **Khan et al. (2021)** proposed a two-step feature-to-feature learning methodology to discover a lower-dimensional latent space based on the combination of geometry- and physics-informed principal component analysis and the active subspace method.

2.2.2 ML in ship performance prediction

Recently, in the maritime industry, ML is being used to solve diverse problems, including energy-efficient route planning **(Zhang et al., 2019)**, structural integrity **(Mikulić & Parunov, 2019)**, seakeeping **(Cepowski, 2007)**, overall fuel consumption **(Wang et al., 2018)**, and resistance reduction **(Margari et al., 2018**. Among these applications, hull resistance reduction prediction with Artificial Neural Networks (ANN) has received more attention during shape optimization. **Margari et al. (2018)** developed an ANN model to predict the residual resistance coefficient of MARAD Systematic Series. The training data was constructed with towing tank experiments, and the overall shape was represented with four design parameters taken as input to the model. An ANN and particle swarm optimization-based shape optimization was developed by **Palmer et al. (2015)** for the optimization of a TriSWACH, a novel trimaran hull with a small waterplane area centre, and two small side hulls, for residual resistance coefficient. Recently, **Yu et al. (2019)** used Deep Neural Networks (DNN) for optimal trimaran configuration for best calm-water transportation efficiency. In this work, the designs in the training dataset were evaluated using a potential-flow code called Multi-hull Simple-source Panel Method (MSPM). Finally, **Danışman (2014)** used ANN with Sequential Quadratic Programming (SQP) to optimize the catamaran hulls' wave resistance. **Diez et al. (2019)** presented a novel ML method for resistance reduction, which combined a dynamic Radial Basis Function (RBF) surrogate model with a sequential Multi-Criterion Adaptive Sampling (MCAS) technique and achieved 9% resistance reduction with few simulations. **Serani et al. (2019)** proposed a multi-fidelity metamodel with adaptive sampling based on stochastic RBF for global design optimization. Other examples of implementing Kriging based surrogate models for ship hull form optimization are **(Liu et al., 2018)**, **(Scholcz et al., 2015)**, and **(Solak, 2020)**.

A significant difficulty in developing such performance prediction models, especially for free-form shapes, is the creation of training data. The computational cost of creating such data rises exponentially with input parameters, especially if the data has to be generated by running time expensive CFD simulations. Therefore, as explained previously, a common practice is to use low-fidelity physics simulation codes to evaluate designs in the training dataset. Recently, to overcome this computational difficulty, **Sun et al. (2020)** proposed a new type of DNN architecture, which does not mainly depend on the simulation data during training and was used for the prediction of incompressible flows around simple shapes. Although the model proposed by **Sun et al. (2020)** has proven to be efficient for simple problems; however, for complex problems, it still needs to be evolved. Therefore, the present work aims to handle this problem

by reducing the number of design variables used to represent the shape with dimension reduction. Moreover, to the best of the authors' knowledge, there is no existing data-driven technique of any kind used for hull resistance prediction in the presence of HV.

3. NUMERICAL MODELLING FOR CFD

In this work, the open-source software, OpenFOAM®,[1], is utilized as the flow solver, which offers the extensive capacity to solve different kinds of laminar, turbulent, and multi-phase fluid dynamics problems. Three-dimensional, incompressible, unsteady RANS equations are solved by implementing the finite volume discretization in order to calculate the ship resistance components on the model scale. Using Einstein's summation convention, the continuity and momentum equations can be written as below:

$$\frac{\partial u_i}{\partial x_i} = 0, \tag{1}$$

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = \frac{-1}{\rho}\frac{\partial p}{\partial x_i} + \vartheta \frac{\partial^2 u_i}{\partial x_j^2} + \frac{\partial}{\partial x_j}\left(\overline{-u_i' u_j'}\right), \tag{2}$$

where *i,j=1,2,3,* and the symbols: t, $x = (x_1, x_2, x_3)$, u, p, ρ, $\vartheta = \frac{\mu}{\rho}$, and μ, represent time, spatial coordinates, velocity, pressure, fluid density, kinematic viscosity and dynamic viscosity, respectively. The last term in the right-hand side of the momentum equation depicts the so-called Reynolds-stress term, which emerges from the Reynolds averaging procedure indicated by the overline symbol. The Reynolds term based on the turbulent viscosity hypothesis is modelled as:

$$\langle -u_i u_j \rangle = \frac{1}{2}\delta_{ij} k - \vartheta_t \left(\frac{\partial \langle U_i \rangle}{\partial x_j} + \frac{\partial \langle U_j \rangle}{\partial x_i} \right) \tag{3}$$

Here, ϑ_t and k indicate the turbulent viscosity and the turbulent kinetic energy **(Pope, 2002)**, respectively, while δ_{ij} is the Kronecker delta symbol. Finally, the Shear Stress Transport (SST) $k - \omega$ turbulence model is employed in order to solve the turbulence equations, where ω represents the specific dissipation rate.

The first-order implicit (Euler) scheme is applied to discretize the unsteady term in Navier-Stokes equations. PIMPLE algorithm, a combination of PISO (Pressure Implicit Split Operator) and SIMPLE (Semi Implicit Methods Pressure Linked Equations) algorithms, is used to solve

[1] https://www.openfoam.com/

the pressure-velocity coupling. Although the pseudo-transient PIMPLE method has a higher computational cost, it is more stable than the SIMPLE method. CFD simulations are carried out in two phases, air and water. Here, the pitch and heave motions of the ship hull are released to model real experimental conditions **(Delen et al., 2020)**. Therefore, "interDyMFoam" **(Hu et al., 2016)** is implemented in OpenFOAM®,[1] version 5 library, which is a solver for two incompressible isothermal immiscible fluids using the Volume of Fluid (VOF) phase fraction-based interface capturing approach with optimal mesh motion and adaptive re-meshing.

The time step, explained in detail in Section 5.3, is chosen equal to 10^{-2} seconds in accordance with the **ITTC (International Towing Tank Conference) (2011)** guidelines. To accurately capture the flow features, the recommended time interval is set equal to $[5 \cdot 10^{-3}]\frac{L}{U}$ seconds, where L and U are the reference length and design speed, respectively.

3.1 Ship Model and Hull Vane Configuration

A motor yacht model is utilized for the simulations in the current study. The waterline (L_{WL}) of the model is 3.5m and the Fr at service speed is 0.37. **Table 1** and **Figure 2** depict the main dimensions and the body plan of the yacht model, respectively.

The NACA4412 foil, which is broadly used in pertinent literature, is selected as the initial cross-section for the HV's shape optimization. The chord length of HV is 2.9% of the waterline length (L_{WL}) and the angle of attack is 0 deg with respect to the calm waterline. Moreover, HV's span length is taken equal to the model's breadth. The leading edge of HV is positioned from the transom corner at a horizontal/vertical distance equal to 2.29% and 1.66% of the L_{WL}, respectively (see **Figure 3 (a)**). These features are kept constant throughout the optimization process. CFD simulations have been conducted without struts connecting HV to the hull.

Table 1: Main particulars of the motor yacht in model scale

Length on waterline	L_{WL} (m)	3.5
Length between perpendiculars	L_{BP} (m)	3.5
Breadth	B (m)	0.727
Draught (midship)	T (m)	0.212
Displacement volume	∇ (m^3)	0.268
Displacement	Δ (ton)	0.268
Wetted surface area	S_M (m^2)	2.769
Block coefficient	C_B	0.533
Longitudinal centre of buoyancy	L_{CB} (m) (+ fwd)	-0.16
Longitudinal centre of floatation	L_{CF} (m) (+ fwd)	-0.348
Service speed	V_M	2.15 m/s
Froude number	Fr	0.37

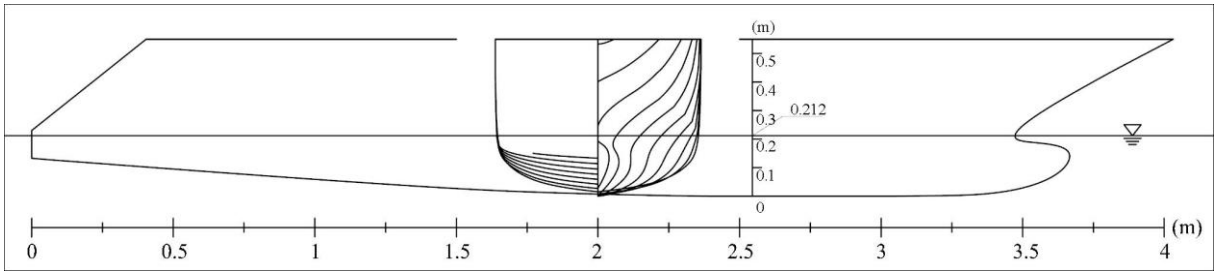

Figure 2: Body plan of the model yacht.

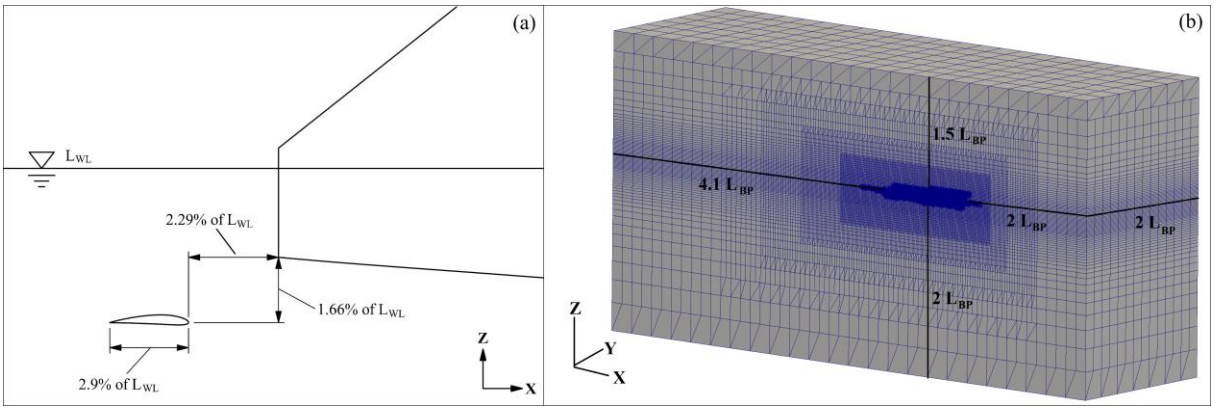

Figure 3: (a) HV configuration and (b) computational domain dimensions.

3.2 Computational Domain and Boundary Conditions

Inlet, bottom, and side boundaries are located $2L_{BP}$ away from the ship model and the outlet boundary is placed $4.1L_{BP}$ downstream direction (see **Figure 3(b)**). These boundary distances are within the range specified in the ITTC procedure and guidelines **(ITTC, 2011)** to avoid the wave reflection from the boundaries.

Boundary conditions are also determined according to the ITTC guidelines. The model hull is specified as the wall function. The sides and bottom boundaries are constrained with a symmetry condition. The inlet, outlet, and atmosphere boundary conditions are defined separately by considering the turbulence parameters and the flow characteristics. Details on the boundary conditions are presented in **Table 2**.

Here, the Dirichlet boundary condition is represented by a fixed value, whereas zero gradients express the homogenous Neumann boundary condition. The outlet phase means velocity, and pressure inlet-outlet velocity applies zero gradients for outflow. The fixed flux pressure sets the pressure gradient on the boundary by the velocity boundary condition. The total pressure is calculated by adding static pressure to dynamic pressure. The inlet-outlet condition provides a zero gradient outflow condition. The frequency, turbulence eddy viscosity, and turbulence kinetic energy wall functions are represented by omegaWallFunction, nutkWallFunction, kqRWallFunction, respectively. **Islam and Soares (2018)** have implemented the same

boundary conditions to estimate the hydrodynamic derivatives of a container ship using OpenFOAM® flow solver.

The table below provides the type of boundary conditions imposed on the different parts of the boundary of the control volume (computational domain) with regard to fluid properties and turbulence parameters depicted in the first column of **Table 2**. These include velocity (U), pressure (p_rgh), phase fraction (alpha.water), turbulent kinetic energy (k), wall function model (nut), turbulence specific dissipation rate (omega). As for the various acronyms used in the remaining columns of the table, their explanation goes as follows: FV: Fixed Value, ZG: Zero Gradient, PIOV: Pressure Inlet Outlet Velocity, MWV: Moving Wall Velocity, FFP: Fixed Flux Pressure, TP: Total Pressure and IO: Inlet-Outlet conditions.

Table 2: Boundary conditions according to the fluid properties and the turbulence parameters.

	Inlet	**Outlet**	**Atmosphere**	**Hull**
U	FV	ZG	PIOV	MWV
p_rgh	FFP	ZG	TP	FFP
αlpha.water	FV	ZG	IO	ZG
k	FV	ZG	FV	kqRWF
nut	FV	ZG	ZG	nutkWF
omega	FV	ZG	FV	omegaWF

3.3 Mesh Generation

A domain mesh within the control volume is created with the blockMesh utility in OpenFOAM®. TopoSet utility, which is the imaginary rectangular prism in the domain mesh, is applied five times to increase the mesh density up to the ship hull. Two nested control volumes are applied to refine the mesh around the free surface, bow, and aft of the hull to capture the flow characteristics adequately. One further control volume is used to refine the mesh around the HV (see **Figure 4**). Finally, the snappyHexMesh utility is executed to create a three-dimensional mesh. After a simulation, the computed y+ distribution along the hull and HV surfaces is around 30.

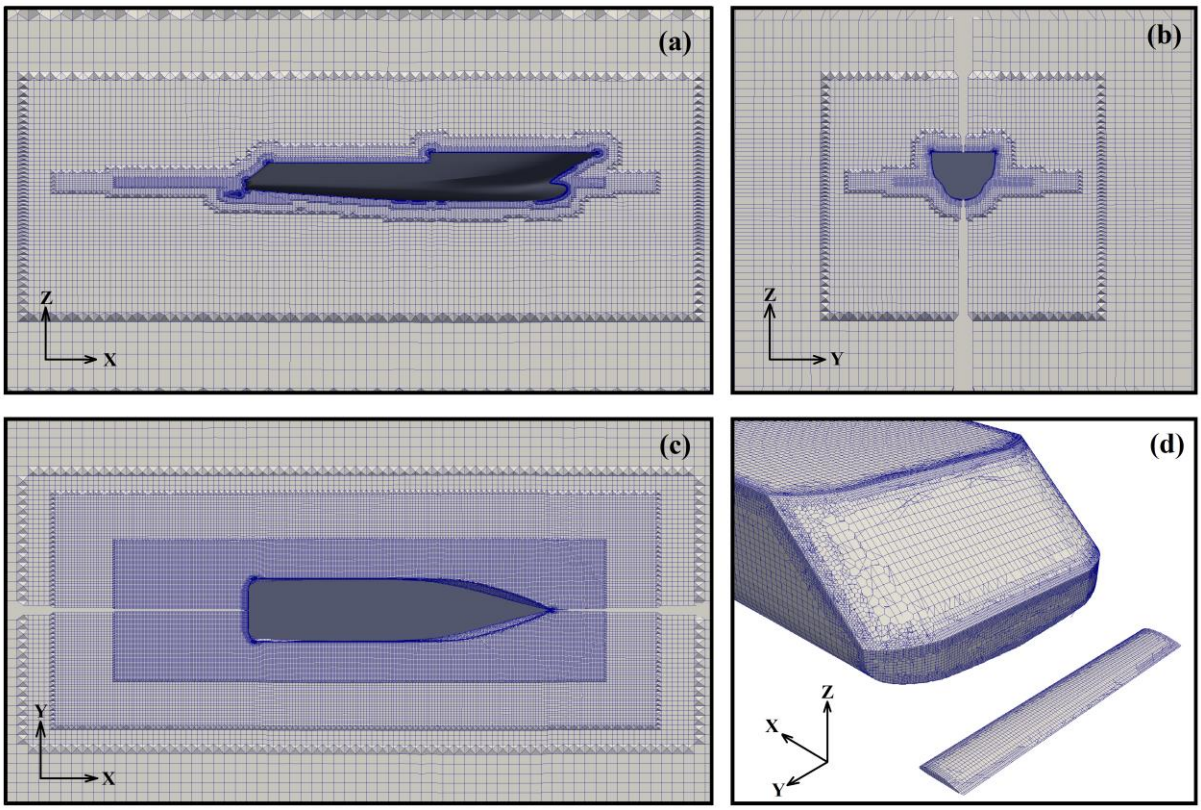

Figure 4: Detailed mesh views of the ship model from (a) profile, (b) section, (c) top, and (d) stern.

4. DATA-DRIVEN HULL VANE SHAPE OPTIMIZATION

In this section, along with the parametric model of the reference HV's cross-section, the overall workflow data-driven process is described, including sampling design of experiments (DoE) for the training dataset, background and implementation of PCA and training of ML model for resistance prediction.

4.1 Hull Vane Parametric Model

The parametric model of the hydrofoil is constructed using a technique described in **Kostas et al. (2017)**. For a given realization of the parameter vector $\mathbf{x} = (x_1, x_2, \ldots, x_7)$, the corresponding hydrofoil is constructed through an automatic geometric construction process, referred to as $G(\mathbf{x})$, described in section 4.1. This construction is based on the determination of the control vertices of four simple Bézier curves, a quadratic and a cubic one for the suction side, and a pair of quadratic simple Bézier curves for the representation of the camber line. The Bézier representation of the pressure side is then obtained by mirroring the suction profile along

the camber line. Conclusively, the geometry of the hydrofoil is represented as a G^1-continuous Bézier curve defined as below

$$H_i(u) = \sum_{k=0}^{3} d_{ik}\,(G;\mathbf{x})B_k^3(t),\ \ t = \frac{u-u_i}{h_i}\epsilon[0,1], h_i = u_{i+1} - u_i, \quad u\epsilon[u_i, u_{i+1}], i = 0, \dots 3, \tag{4}$$

where d_{ik} represents the kth control point of the ith simple Bézier curve. $B_k^3(u), k = 0, \dots, 3$, is the corresponding cubic Bernstein basis and $U = \{u_{i,} i = 0, \dots 4: u_i < u_{i+1}, i = 0, \dots, 3\}$ is a fixed user-specified knot vector. In view of (4), we deduce that, if we change the realization of the parameter vector from $\mathbf{x}$ to $\acute{\mathbf{x}}$, the geometric construction G will deliver a new set of control points, $d_{ik}(G;\acute{\mathbf{x}}), k, i = 0, \dots, 3$, which will produce a deformed hydrofoil geometry.

Initially, the profile is defined via eight parameters, namely the length L of the hydrofoil, the maximum width of the suction side with respect to chord (z_max), the maximum camber width with respect to chord (c_max), the longitudinal position of suction side's maximum width (x_z_max), the longitudinal position of camber side's maximum width (x_c_max), the suction side's angle at trailing edge with respect to chord (a_b), the camber angle trailing edge concerning chord (a_b_p) and, finally, the leading-edge form factor (tip).

Table 3 provides the description of the parameters, while Figure 5 gives their graphical illustration. Without loss of generality $L = 1$ in this work. In view of this choice and the formulae in the fifth column of **Table 3**, we first conclude that all remaining length parameters will take values in (0,1). However, it should be stressed that the seven parameters x_1:=z_max, x_2:=c_max, x_3:=x_z_max, x_4:=x_c_max, x_5:= a_b x_6:= a_b_p and x_7:= tip do not lie in the hyper-cube $[0,1]^7$, equivalently they are not linearly independent. A more careful look into the formulae in the fifth column of **Table 3** reveals that the lower (x_k^l) and upper (x_k^u) limits of a parameter x_k is, in general, depend on the previous parameters $x_1, \dots, x_{k-1}$, which is accurately expressed as:

$$x_k^l \le x_k \le x_k^u, \text{ where } x_k^l = f_{lk}(\mathbf{x}_k), x_k^u = f_{uk}(\mathbf{x}_k) \text{ and } \mathbf{x}_k = (x_1, \dots, x_{k-1}), \tag{5}$$

For example,

$$f_{u2}(\mathbf{x}_1) = 0.9x_1, f_{l4}(\mathbf{x}_3) = \frac{7x_3}{10}, f_{u4}(\mathrm{x}_3) = \frac{3}{10} + \frac{7x_3}{10}, f_{l5}(\mathbf{x}_4) = \arctan\left(\frac{x_1}{L-x_3}\right). \tag{6}$$

Table 3: Parameters' definition.

Parameter	Name	Description	Symbol	Actual range
	Length			
x_1	Max width	Maximum width of suction side w.r.t. chord	z_max	$\left[\frac{L}{500}, \frac{L}{5}\right]$
x_2	Camber width	Camber maximum width w.r.t. chord	c_max	$[0, 0.91x_1]$
x_3	Max-width position	Longitudinal position of suction side's max width	x_z_max	$\left[\frac{L}{5}, \frac{7L}{10}\right]$
x_4	Max-camber-width position	Longitudinal position of camber's max width	x_c_max	$\left[0, \frac{3L}{10}\right] + \frac{7x_3}{10}$
x_5	Suction-side angle	Suction side's angle at trailing edge w.r.t. chord	a_b	$\left[\arctan\left(\frac{x_1}{L - x_3}\right), 89\right]$
x_6	Camber angle	Camber angle at trailing edge w.r.t. chord	a_b_p	$[0, x_5]$
x_7	Tip	Leading edge form factor	tip	[0.1,0.9]

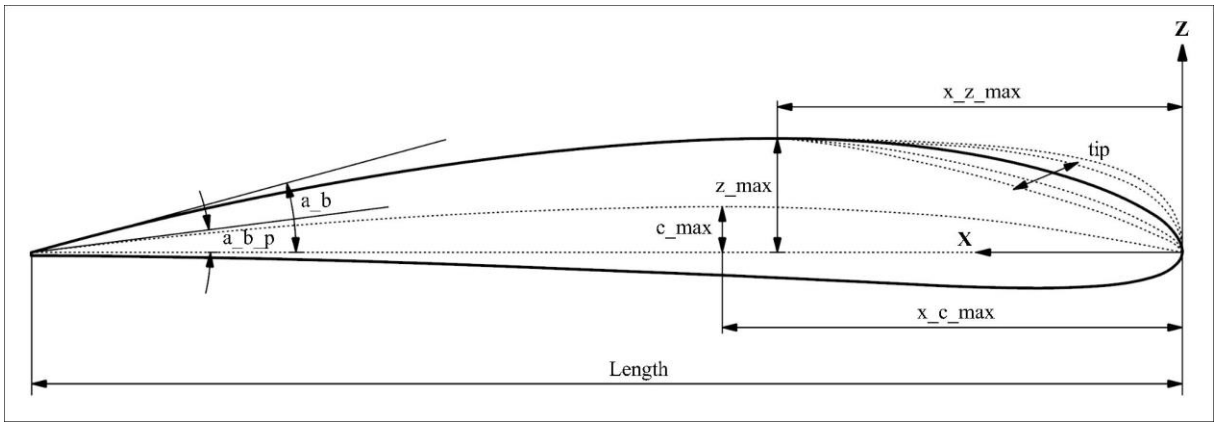

Figure 5: Hydrofoil's parametric representation.

The whole sectional profile consists of four simple Bézier curves for suction and pressure sides. The suction side consists of cubic and quadratic Bézier curves; the first and last control points of the cubic Bezier curve lay at $(x - axis, y - axis) = (0,0)$ and at (x_z_max, z_max), respectively, and the internal control points of this curve lie on a line between $(0, x\_max)$. The first control point of the quadratic Bézier curve coincides with the last control point of the cubic curve, and its last control point lies at $(L, 0)$. The internal control point of the quadratic curve has the coordinates $\left((L - \frac{z\_max}{tan(a,b)}), z\_max\right)$. The camber curve is generated with two quadratic Bézier curves. To obtain the pressure side curve, suction side curves are mirrored along the camber curve. Finally, all four Bézier curves are approximated by a single cubic B-spline curve. For further detail on the construction of parametric representation, readers should refer to **Kostas et al. (2017)**. The sectional profile is extruded to generate a three-dimensional surface model of HV and is transversely integrated behind the hull.

Once the parametric modeller is set up, then PCA is implemented in two different settings. In the first setting, PCA is applied to the design parameters, **x**, of the HV designs sampled from a design space, whose construction with be discussed in the subsequent sections. In the second setting, PCA is implemented on the discretizations of the HV designs. As explained earlier in this section, the design parameters do not lie in the unit-hypercube, so the PCA applicability prerequisite, regarding non-orthogonality and existence of correlation between the design parameters, is fulfilled. It should be stressed that the parameters in **x** are intrinsically involved in the geometry construction through G, so the principal components in the first setting capture the variance associated with designs via x. In the second setting, principal components capture the variance of points representing the discrete version of the HV designs.

4.2 Sampling Design of Experiment

In this study, first, two datasets containing only DoE are constructed for the implementation of PCA to generate a lower-dimensional subspace of the original design space. Afterwards, another dataset is created from the subspace, whose design instances are evaluated with the help of the viscous flow solver as described in Section 3. This dataset is then used for training the ML model. The total number of samples required for the ML model depends on the number of input variables and the number of hidden layers. The study carried out by **Alwosheel et al. (2018)** suggests that the minimum number of samples should be fifty times the number of weights that corresponds to connection points between the input variables and the nodes in the hidden layers. The evaluation of a large set of HV designs using CFD is a computationally expensive process; therefore, the main objective of dimensional reduction is to reduce the number of samples needed to train the ML model by reducing the input dimensionality of the dataset.

As described in the previous section, seven parameters represent the HV sectional profile, which creates a seven-dimensional design space to be explored during optimization. The bounding limits of the design space are defined by setting upper and lower limits of the design parameters centred around the reference design values. Implementation of PCA or ML requires a uniformly distributed and diverse set of samples, which covers all the design possibilities of given design space with only a few samples. One of the aims is also to train a reliable model while keeping the computational cost associated with the construction of the training dataset low. It is well known that the convergence of Monte Carlo sampling is slow, as it requires a large number of samples to cover the entire design space (**Khan and Gunpinar, 2018**). If a

small number of samples are taken with Monte Carlo, then there is a high possibility that these samples will be unevenly distributed or will be clustered at some region of the design space. Therefore, if sampling is performed entirely based on Monte Carlo, then training a surrogate model with a small number of samples may not generalize the whole problem. Therefore, in the present work, we use this simple strategy, which expands the design space variably in three steps and samples some proportion of the total number of designs during each step. This method prevents the clustering of samples in some regions of the design space.

In this approach, we first evaluate the centroid of the given design space and then we create a temporary design space (X') whose upper and lower limits are evaluated as 20% percentage of increment and decrement of the centroid, respectively. From this space, we sample 50% of the total number of designs. Afterwards, we create another design space (X'') whose limits are composed as 30% of increment and decrement of the centroid. From this space, we sample 30% of the total designs; however, at this stage, these designs are constrained to be different from those present in X'. In the final step, the third design space (X''') is created as 40% of increment and decrement of the centroid, and the remaining 20% of designs are sampled from this space. These designs are constrained to be different from those in both X'' and X'''. In this way, we try to distribute the samples over the design space evenly. A detailed algorithmic formulation of sampling is given in **Algorithm 1**.

Algorithm 1

1: A parametric design parametrized with n design parameters $x = \{x_k, k = 1,2, \dots, n\}$.

2: Define the design space X with lower x^l and upper x^u bounds of n parameters, $X \coloneqq \left\{x_k^l \le x_k \le x_k^u, \forall k = \{1,2, \dots, n\}\right\}$.

3: Initialize the total number of designs (L) to be sampled from X.

4: Evaluate the centroid of the design space $c = \{c_k, k = 1,2, \dots, n\}$.

5: Create a temporary design space X' as $X' \coloneqq \left\{c_k - 0.2c_k \le c_k \le c_k + 0.2c, \forall k = \{1,2, \dots, n\}\right\}$.

6: Sample $0.5N$ designs from X' and create dataset $D' = \{x_i, i = 1,2, \dots, 0.5N\}$

7: Create X'' as $X'' \coloneqq \left\{c_k - 0.3c_k \le c_k \le c_k + 0.3c, \forall k = \{1,2, \dots, n\}\right\}$.

8: Sample $0.3N$ designs from X'' and create dataset $D'' = \{x_i, i = 1,2, \dots, 0.3N | x_i \notin X'\}$.

9: Create X''' as $X''' \coloneqq \left\{c_k - 0.4c_k \le c_k \le c_k + 0.4c, \forall k = \{1,2, \dots, n\}\right\}$.

10: Sample $0.2N$ designs from X''' and create dataset $D''' = \{, i = 1,2, \dots, 0.2N | x_i \notin X''\}$.

11: Create a final training dataset $= D' \cup D'' \cup D'''$.

4.3 Dimension Reduction

4.3.1 Principal Component Analysis (PCA)

PCA reduces the dimensionality of a given design space by performing a projection of the data points, sampled from the design space, in a new linear subspace, which is defined by the eigenvectors (called the PCs) of the $[L \times L]$ covariance matrix (**D' Agostino et al., 2020**) represented as;

$$C = \frac{1}{L} D^T D \tag{7}$$

Here, D is a dataset consist of L samples from X, and each sample is of n dimensions, so $D \in R^{L \times n}$. d_k is the kth design (row) of D. This matrix is used to compute the covariance matrix, C. The eigenvectors have the properties of maximizing the (geometric) variance of points projected on them and minimizing the mean squared distance between the original points and the relative projections. The PCs are defined by the solution of the eigen problem as;

$$Cz_i = \lambda_i z_i, \qquad i = 1,2, \dots, L. \tag{8}$$

Moreover, the eigenvalues $\{\lambda_i\}_{i=1}^{L}$ with $\lambda_i > \lambda_{i+1}$ represent the variance resolved along the relative eigenvectors $\{z_{i=1}^{L}, (z_i^T z_i = 1)\}$. From this property, a subset of N eigenvectors is used to compute a reduced dimensionality representation x_k of the original vector d_k as;

$$x_k = Z^T d_k, \quad dim(Z) = N \times L, \tag{9}$$

where the matrix Z has dimensions $dim(Z) = N \times L$ and is composed of the first N largest variance PCs. The projection on the orthonormal basis given by the columns of Z is

$$\widehat{d_k} = Zx_k = ZZ^T d_k \tag{10}$$

Here the above equation estimates the design in reduced dimensional subspace to the original n-dimensional space. ZZ^T is the projection matrix which defines the linear transformation of d_k in the subspace defined by the column space of Z and the projection $\widehat{d_k}$ represents the minimum squared error approximation of the relative d_k.

4.3.2 Hull Vane Reconstruction with PCA

As explained earlier, we first implement PCA on the 8 design parameters used for the geometry construction through G, for which a dataset D consisting of $L = 1000$ samples is created. The

variance and cumulative distribution of the resulting PCs carrying geometric information are shown in **Figure 6**.

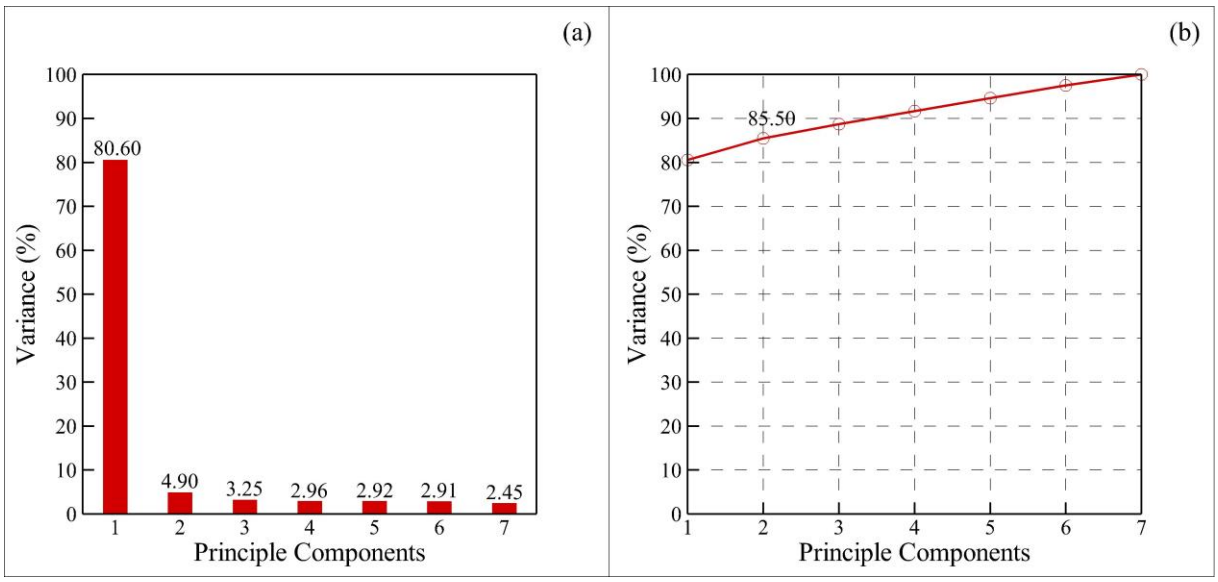

Figure 6: Variance of principal components (a) and their cumulative sum (b) when PCA is implemented on the dataset composed of design parameters defining the hydrofoil shape.

It can be seen that 85.5% geometric variation in the original design space is captured with the first two PCs. Therefore, a two-dimensional subspace based on these components can be generated, whose bounding limits are calculated by the scalar multiplication of the bounding limits of the original design space with the first two eigenvectors. The choice of this variance for the decision on the final number of eigenvectors depends on the problem. Although, for some applications, a variance of 90% or 95% is typically favoured; however, it will be shown with optimization results in Section 5.5 that, in our case, 85.5% gives us enough geometric variability to find an optimal design with good reconstruction accuracy. **Table 4** shows some of the PC-based design instances along with their full space projection. **Figure 7** shows some of HV's cross-sectional sample geometries with their corresponding full space values.

Table 4: Samples obtained from subspace and their corresponding full-scale representation in the original design space.

PCs			**Design parameters**						
No	PC_1	PC_2	**max_z**	**max_c**	**x_z_max**	**x_c_max**	**a_b**	**a_b_p**	**tip**
1	1.2655	-0.0008	0.518	0.483	0.343	0.643	0.148	0.480	0.562
2	1.1051	-0.0849	0.454	0.409	0.258	0.603	0.075	0.403	0.507
3	1.0932	-0.0589	0.449	0.408	0.268	0.584	0.091	0.403	0.497
4	1.3954	-0.0636	0.573	0.523	0.347	0.740	0.123	0.517	0.632
5	1.0603	-0.0015	0.434	0.405	0.287	0.539	0.124	0.402	0.471
6	1.0634	-0.0271	0.436	0.402	0.275	0.553	0.108	0.398	0.477

7	1.2080	-0.0530	0.496	0.453	0.302	0.639	0.108	0.448	0.546
8	1.2171	-0.0813	0.500	0.452	0.290	0.658	0.091	0.446	0.556
9	1.3896	-0.0732	0.571	0.519	0.341	0.742	0.116	0.513	0.631
10	1.2401	-0.0226	0.508	0.470	0.325	0.641	0.131	0.466	0.555

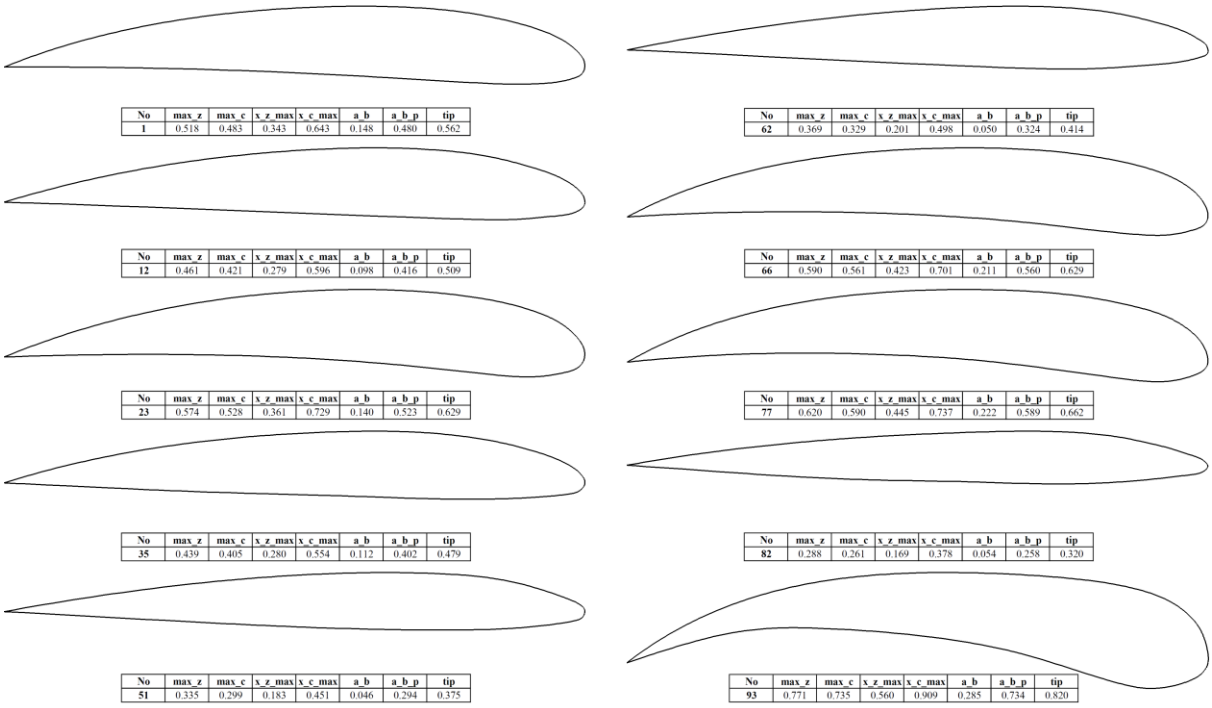

Figure 7: Cross-sectional geometries of HV samples of the training dataset.

Furthermore, along with implementing PCA on design parameter set, $\mathbf{x}$, we also implement PCA on the output of $G(\mathbf{x})$, which define an automatic geometric construction process for the hydrofoil for any given $\mathbf{x}$. In this case, each row of matrix D is composed of $\hat{n}$ points on the profile of the hydrofoil. As each point is two dimensional, so each instance (row) of D constitutes $2\hat{n}$ elements for all L samples, which forms $D \in R^{L\times 2\hat{n}}$, where $\hat{n} = 400$. **Figure 8** shows the results of the geometric variance resolved by the first ten eigenvectors. From these results, it can be seen that the first two components resolved 92% of the variance, which is higher than that obtained when dimensionality reduction is performed on the original design space. It is noteworthy that with this lower-dimensional design space, one can perform shape modification using only the eigenmodes, without the need of going back to the original design space as PCA is implemented on the discretization of the design sampled from the original parametric space. Therefore, it only provides the discretized version of the geometry, which in our case are points on the hydrofoils' profile. Consequently, an accurate and smooth reconstruction of the hydrofoil profile cannot be guaranteed, which is essential for our study due to the sensitivity of our high-fidelity hydrodynamic calculations to shape changes. In this

setting, PCA results are also sensitive to the resolution of this discretization. Even though shape deformation can be achieved at a low resolution, designs obtained via interpolation may miss essential shape features or introduce unwanted shape oscillations, resulting in uncertainty and false estimation of total resistance. This is prominent especially for surface and volume models, which are two- and three-dimensional objects, respectively, laying in a three-dimensional ambient space. With this increment in the object's dimensionality, the need for finer discretization increases exponentially to achieve stable results. Therefore, with such discretization, the matrix D can be high-dimensional, which implies that the computational cost to perform eigendecomposition is analogously high **(Sharma et al., 2013)**. For instance, for 5000 hydrofoil samples it takes 0.0485 seconds to implement PCA on D composed of eight design parameters and 27.6277 seconds to implement PCA on the same number of samples with each sample discretised via 200 points. Thus, in this work, the lower-dimensional design space constructed with the design parameters is used for prediction model training and running the shape optimization of HV.

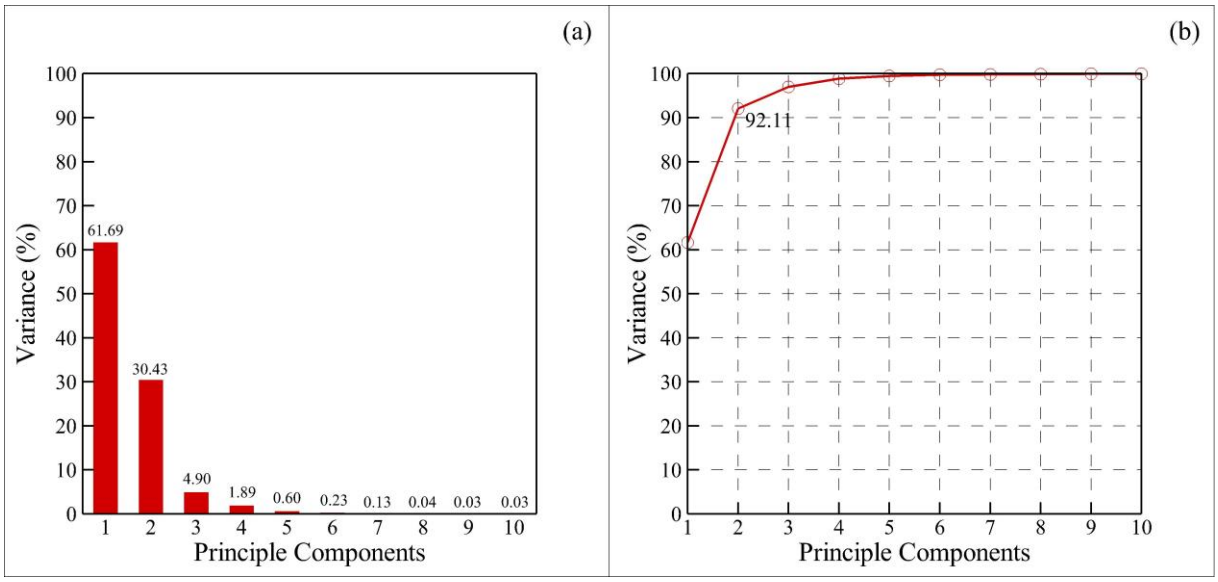

Figure 8: Variance of principal components (a) and their cumulative sum (b) when PCA is implemented on the dataset composed of points on the profile of the hydrofoil.

4.4 Resistance Prediction

As mentioned before, an ML-based prediction model is developed to evaluate the total resistance of the hull form in the presence of HV. In subsequent subsections, we discuss the training dataset and ML method used for total resistance (R_{TM}) prediction on model scale.

4.4.1 Dataset Construction

A training dataset consisting of $L' = 100$ hydrofoil designs are created. The hydrofoil designs are sampled from the subspace composed of the first two PCs by using the sampling method explained in Section 4.2. After that, designs in the training dataset are projected back to the original design space, and 3D geometries of HV are generated by utilizing the parametric modeller written in the Rhinoceros®,[2] script. Each HV design is integrated behind the vessel within the same parametric modeller, and the ship geometry with HV is exported from Rhinoceros® as the triangulated surface geometry in .stl file format. R_{TM} values of these designs are evaluated using the viscous solver, as explained in Section 3. Afterwards, a training dataset is constructed, which contained PCs as independent parameters and R_{TM} as dependent parameters. The R_{TM} values for the first ten samples of the training dataset are shown in **Table 5**.

Table 5: First ten samples of the training dataset.

No	PC_1	PC_2	$R_{TM}[N]$
1	1.2655	-0.0008	52.1590
2	1.1051	-0.0849	52.4288
3	1.0932	-0.0589	52.5504
4	1.3954	-0.0636	51.4698
5	1.0603	-0.0015	52.5742
6	1.0634	-0.0271	52.0222
7	1.2080	-0.0530	51.9176
8	1.2171	-0.0813	51.4288
9	1.3896	-0.0732	51.4324
10	1.2401	-0.0226	51.5902

4.4.2 Training the ML Model

Artificial neural networks (ANN) are inspired by biological neural networks in the nervous system and are used to tackle problems with highly complex relationships between covariates and response variables **(Rojas, 2013).** Based on the working principle of a nervous system, there are interconnected neurons, and they receive numerical data from other neurons. An architecture of a simple ANN is shown in **Figure 9**. The terms x_1 and x_2 indicate the inputs while w_1 and w_2 represent the respective weights of the input values. The summation, including the bias term (b), is transformed by an activation function, which gives an output value **(Arnold, 2016).** In this study, Matlab®,[3] Neural Network Toolbox is utilized.

[2] https://www.rhino3d.com/

[3] https://www.mathworks.com/

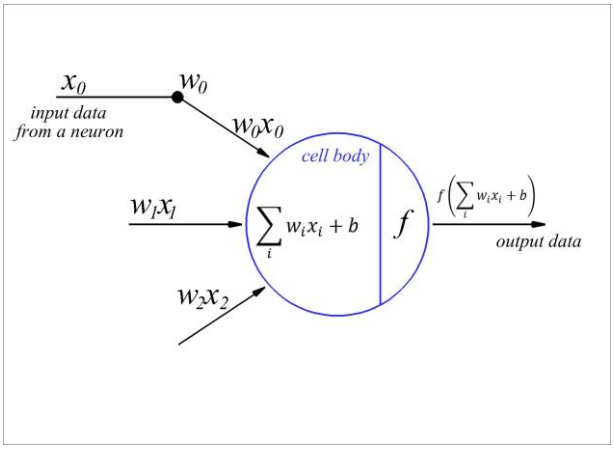

Figure 9: A typical neuron structure **(Arnold, 2016).**

ANN is employed as an ML model, which is trained with a dataset explained in Section 4.4.1 in order to predict the total resistance. The trained ANN is capable of imitating the flow solver by establishing a relationship between an adequate number of HV instances and the corresponding CFD results. The function obtained from the ANN is implemented as a cost function in the optimization process. The ANN architecture used in this study is shown in **Figure 10 (a)**, which consists of two inputs, one output and a single hidden layer with 20 neurons with sigmoid activation function. As we will show in subsequent sections of the paper, even with such a simple architecture, we are able to obtain significant prediction accuracy.

The training of ANN is carried out with the Levenberg-Marquardt **(Hagan and Menhaj, 1994)** network training function, and as depicted in **Figure 10 (b)**, the weights are adjusted by comparing the ANN's outputs with the targets, which are the total resistance values obtained from the solvers. The development of the prediction model with ANN comprises three fundamental phases: training, validation, and testing.

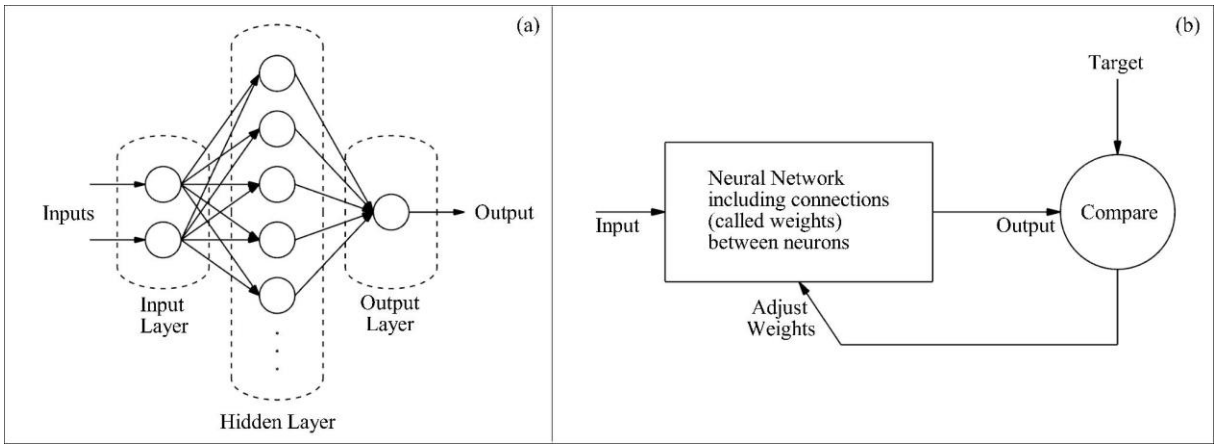

Figure 10: (a) ANN architecture and (b) workflow of the training process.

In the training process, the error between the ANN output and actual R_{TM} values of designs are evaluated to feedback the neural network for weight adjustment. After the training process is completed, a linear regression analysis is performed between the target and ANN outputs. The ideal case of this regression is that the slope of the regression curve is equal to 1, which means that the outputs and the targets overlap. The regression curve obtained from the training is compared with the ideal regression curve to interpret the system's success. In this study, 75% of the dataset is used for the training process.

While training, overfitting happens if the error gradually increases with the addition of a new dataset to the network, which results in inaccurate predictions. The validation process prevents the trained model from overfitting. In this process, the trained model is evaluated in order to keep the error at a minimum level by tuning the hyperparameters such as dropout (regularization technique), network weight initialization, activation function, and so on. Similar to the training phase, in the validation phase, the ideal regression curve is also compared with a linear regression between ANN outputs and actual values of R_{TM}. 15% of the overall dataset is reserved for the validation phase.

The testing phase is completely independent of the training and validation processes and 10% of the dataset is used at this phase. After a reliable training and validation process, the ANN is ready to replace the flow solver.

4.5 Optimization Method

The two-dimensional subspace is explored to find an optimal design for HV. During subspace exploration, the design's total resistance is evaluated with the trained prediction model. During this optimization, one can also assess designs with a CFD solver; however, this will significantly increase the computational cost because for high-dimensional problems, like one studied in this work, the commonly used stochastic optimization techniques require a large number of simulation runs (**Gunpinar and Khan, 2020**) and when each run is extensively computationally demanding then, it can be difficult, if not impossible, to exhaustively explore the design space for a global optimum. Although to overcome this, one could opt for more traditional approaches, such as gradient descent; however, they have their associated computational cost as they require the evaluation of gradients, which can be challenging to evaluate if the baseline simulation code cannot provide these gradients. Therefore, in our approach, we choose to perform shape optimization in connection with the surrogate model. Furthermore, as the subspace has lower dimensionality the original design space, it does not

require a large number of designs to be evaluated. This lower dimensionality also reduces the possibility for an optimizer to give local optima and reduces the computational cost of the whole optimization process significantly.

Moreover, as after dimensionality reduction, the problem becomes two-dimensional; therefore, one may estimate minima by analyzing the two-dimensional contour plot. However, such analysis may not help to evaluate the precise parametric values for the solution close to the global optima, especially if the problem is non-convex. Therefore, to ensure that we obtain the most optimal hydrofoil design, we run an optimization even in this two-dimensional subspace. An interior-point algorithm, available in Matlab® library, is chosen as the optimizer for subspace exploration. This algorithm satisfies the linear constraints for design space bounding limits at each iteration. Its success, especially in large-scale linear and nonlinear programming, has been proven compared to the other algorithms. Furthermore, different literature has that the interior point algorithm operates faster in problems than the other algorithms. For information on the detailed formulation of the interior-point algorithm, interested readers should appeal to Chapter 19 of **Nocedal and Wright (2006)**.

5. DISCUSSION OF RESULTS

5.1 Verification and Validation of the CFD

The spatial uncertainty estimation procedure for CFD applications proposed in **Celik et al. (2008)** has been implemented to the bare hull. The total number of mesh elements in the fine, medium, and coarse domains are 2.5, 1, and 0.41 million, respectively. Total resistance coefficients (C_T) obtained from CFD simulations are taken into consideration as uncertainty variables. The spatial uncertainty calculation steps are shown in **Table 6**. The fine Grid Convergence Index (GCI) is estimated to be 0.01%. The CFD simulations for each case are performed with the number of mesh elements used in the fine domain.

Table 6: Spatial discretization uncertainty.

Number of mesh elements	Fine	2503279
	Medium	1022266
	Coarse	413198
Refinement factor	r_{21}	1.348
	r_{32}	1.352
Total resistance coefficients	C_{TM1}	0.00981
	C_{TM2}	0.00982

	C_{TM3}	0.01003
Apparent order	P_{last}	9.2304
Extrapolated value	$f_{21\ ext}$	0.0098
Approximate relative error	(%) e_{21a}	0.1306
Extrapolated relative error	(%) e_{21ext}	0.0089
Fine grid convergence index	(%) GCI_{21fine}	0.01

For the bare hull, the towing tank experiments have been performed at Ata Nutku Ship Model Test Laboratory[4] of Istanbul Technical University. The results have been compared with the CFD simulation, which is carried out with the fine domain mesh features at a design speed of 2.15 m/s. CFD simulation results show good agreement with the experiment, showing the difference of 1.54% in R_{TM}.

5.2 Effect of the Hull Vane on the Resistance Components

In multi-phase CFD simulations, the frictional resistance (R_{FM}) and the residuary (R_{RM}) are computed separately to obtain the R_{TM}. Since R_{FM} and the viscous pressure resistance (R_{VPM}) do not depend on the wavy part of the phenomenon, their sum ($(1+k)R_{FM}$) is evaluated by simulating the flow in a single-phase domain (double-body approach). It should be noted that in the single-phase simulation, the force resulting from the negative pressure gradient on the transom of the ship model is not taken into consideration. Since, in the multi-phase CFD simulation, the transom area is completely dry at service speed, which means that there is no force on it. Finally, the wave-making resistance (R_{WM}) is obtained by subtracting $R_{FM} + R_{VPM}$ from R_{TM}.

Table 7: HV impact on the resistance components.

	$R_{FM}[N]$	$R_{VPM}[N]$	$R_{WM}[N]$	$R_{TM}[N]$
Bare Hull	20.49	5.84	36.45	62.78
Bare Hull w/HV (NACA 4412)	22.12	4.93	24.56	51.60
Difference (%)	+8.0	-15.7	-32.6	-17.8

[4] https://www.researchgate.net/lab/Ata-Nutku-Ship-Model-Basin-A-G-Avci

In **Table 7**, the total-resistance along with its components are compared, in model scale, with and without HV. Firstly, R_{FM} is expected to increase as a result of the fact that the total wetted surface area has increased with the installation of the HV. The CFD simulation results show good agreement with the expectations, showing an increment of 8% in R_{FM}. On the other hand, it is observed that R_{VPM} and R_{WM} have been decreased due to the presence of the HV. More specifically, a significant portion of the reduction in the wave-making resistance, R_{WM}, is contributed by the reduction in the R_{TM}.

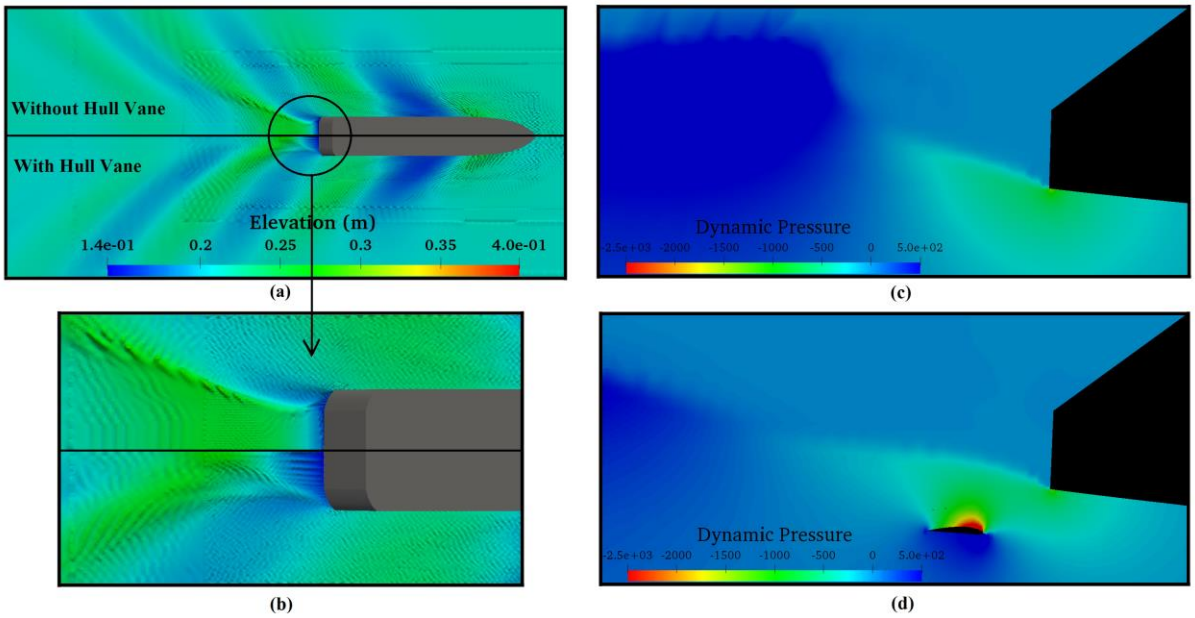

Figure 11: Wave elevation and dynamic pressure distribution around the stern area with and without the HV.

A negative pressure zone, which helps to reduce the stern wave, appears on the top side of the HV due to accelerated flow from the aft of the hull. The change in stern waves due to the influence of the HV is clearly depicted in **Figure 11 (a)** and **(b)**. Moreover, the pressure distribution of the stern profile section can be seen in **Figure 11 (c)** and **(d)**. As a result, R_{TM} and its component, R_W, have been decreased by 17.8% and 32.6%, respectively, on the model scale.

5.3 Computational Time

The technical specifications of the hardware server used for the code implementation and the materialization of the experiments scheduled in the context of the proposed approach are as follows; Intel Xeon Gold 6148 v5 type processor, 56 cores and Centos 7 operating system.

CFD simulations have been performed based on the unsteady RANS equations. Firstly, a simulation is started with the 0.005 seconds time step in accordance with the **ITTC (2011)** procedure, which continued until 5 computational seconds while the recorded Courant number set to around 0.02. The real computational time corresponding to 5 seconds is approximately 40 minutes. In the second stage, the time step is doubled, and the simulation continued for up to 23 computational seconds. The Courant number is around 0.04 between 5 and 23 seconds and the actual corresponding computational time is nearly 120 minutes. In the final stage, the time step is set back to 0.005 seconds in order to better capture the flow characteristics with a low time step. The simulation continued from 23 to 30 seconds, which corresponds to an actual computational time of about 60 minutes. The time step adjustments are prepared in a single script file that is inputted to the terminal of the server. In summary, the total computational time of a simulation is approximately 220 minutes.

5.4 Validation of the ML Model

The dataset used for the training, validation and testing of ANN consisted of 100 designs with their first two PCs as independent parameters and R_{TM} values as the dependent parameter, which are evaluated with a viscous flow solver. **Figure 12** shows the linear regression plot of the output (Y), namely the values of R_{TM} predicted by the trained ANN, versus the target values (T), which is the actual R_{TM} obtained from the viscous solver. The correlation between outputs and targets is represented by the correlation coefficient (R) for which the ideal case corresponds to 1. As depicted in **Figure 12**, the R values of the neural network model in the training, validation, and testing phases are very close to 1, which indicates a successful training of the prediction model.

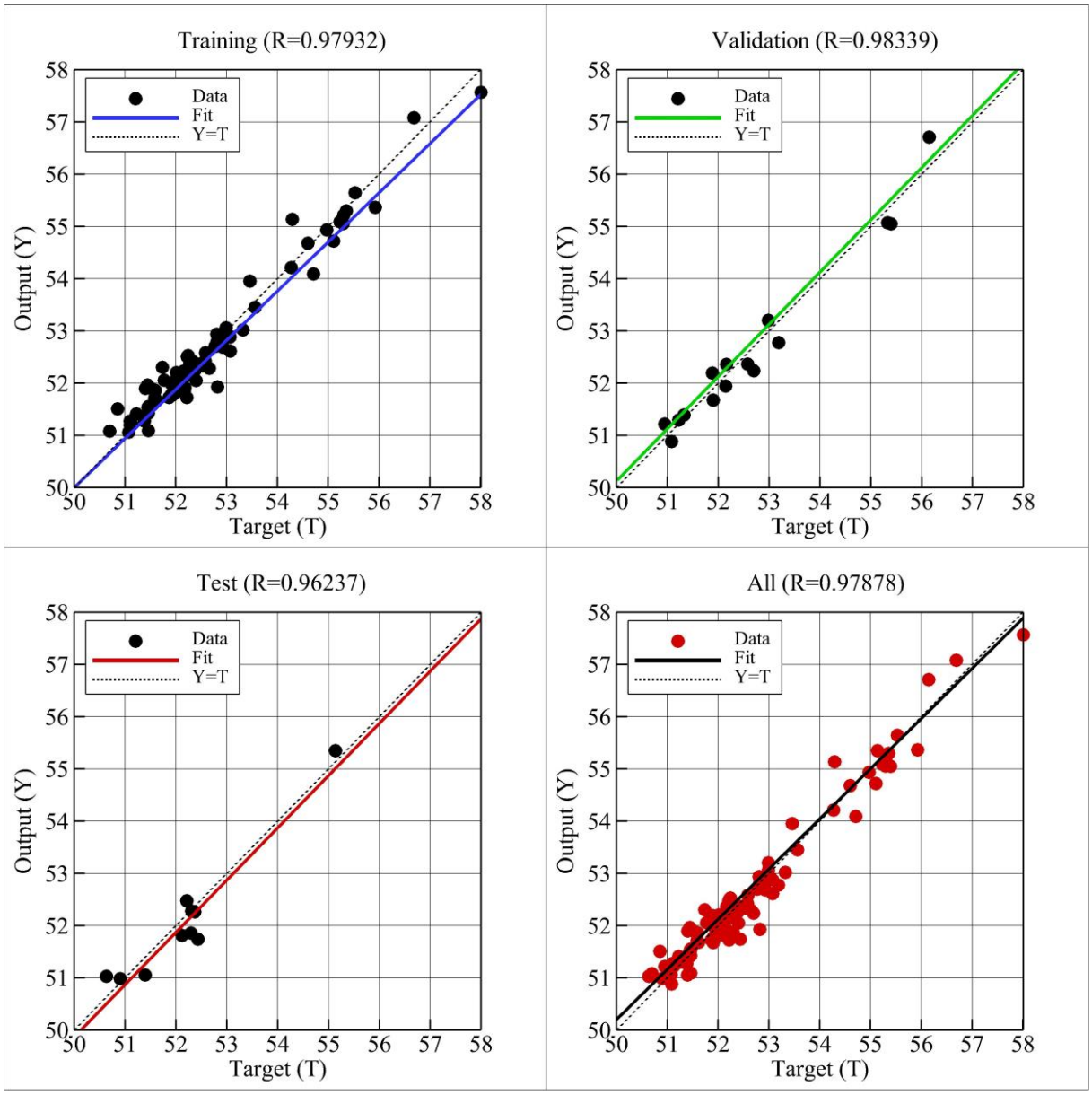

Figure 12: Comparison of the linear regression between predicted and target resistance values with respect to the ideal regression.

Furthermore, **Figure 13** depicts the distribution of the absolute errors between the simulation results of the HV designs and the predictions obtained from the ANN. It is seen that the errors in training, validation, and testing phases are distributed around zero in accordance with the Gaussian distribution, and the majority of the errors are smaller than ∓ 0.25 absolute errors.

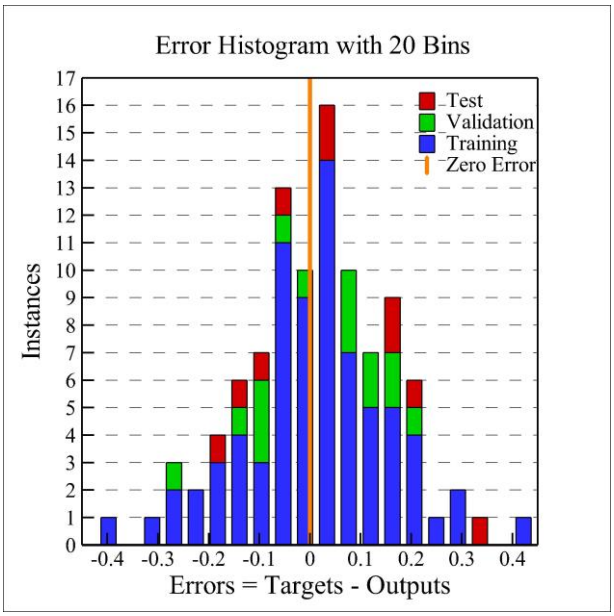

Figure 13: Error histogram of the trained model.

Finally, **Figure 14** depicts MSE (mean-squared error) values on the logarithmic scale versus Epochs that present the set of training vectors to a network one at a time. When the graph is interpreted in detail, the MSE values in the testing phase show good agreement with the MSE values in the validation phase. If there is an increase in MSE values during the testing phase, the trained model has an overfitting problem. The validation MSE is monitored during the training, and training is terminated when validation MSE starts to increase. The training is performed with 10 Epochs and the best validation performance is achieved with an MSE value of 0.01759 at 4th Epoch.

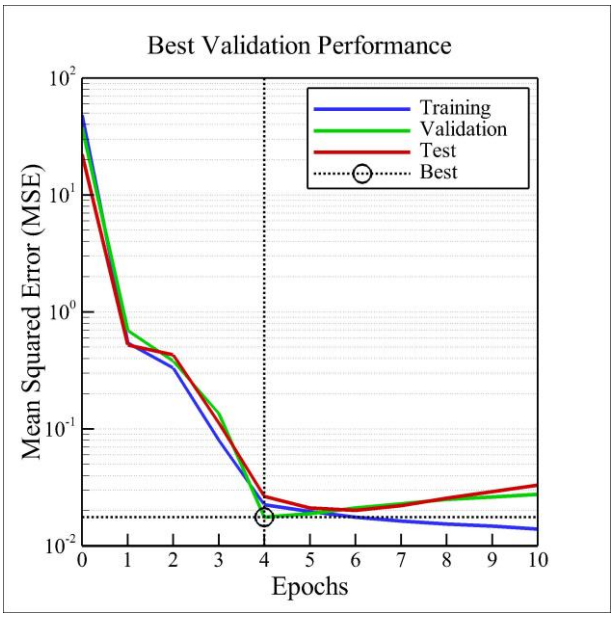

Figure 14: MSE values during the training, validation, and testing processes.

5.5 Optimization Results

In the present study, the shape optimization of the HV has been performed in order to reduce the total ship resistance. The objective function in the optimization process is an explicit function obtained after the trained ANN model. The interior-point algorithm, described in Section 4.5, has been used as the optimization algorithm. A series of initial values within the boundaries of the design space is defined, and local minimum values are searched to control the robustness of the optimization problem. It is observed that the local minimum values corresponding to the initial values always converge towards the same point with a R_{TM} of 50.72 N. **Figure 15** show the contour plot of the function obtained from ANN and the position of the optimal point.

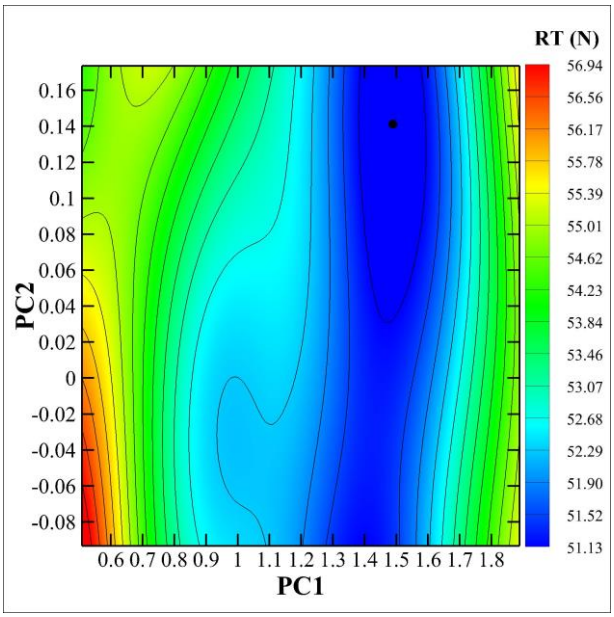

Figure 15: Two-dimensional contour plot of the objective function obtained from ANN.

The PCs of the optimized design are projected back to the original seven-dimensional design space to generate the three-dimensional shape of optimized HV. The initial and the optimized sectional profile of HV are shown in **Figure 16**.

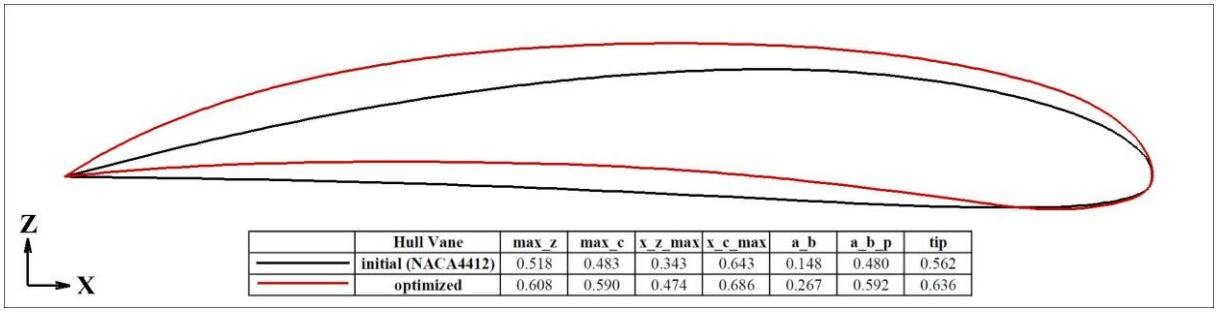

Figure 16: Shape of NACA4412 and optimized HV profile.

The optimum HV design is installed behind the vessel to compute the R_{TM} via a viscous flow solver. The R_{TM} estimation obtained from the simulation is 50.84 N. The absolute percentage error between the total resistance predicted from the ANN model and the resistance computed by the flow solver is 0.2361% (see **Table 8**). This shows that the ANN model trained with PCs can predict R_{TM} values for a hull with HV with satisfactory accuracy, even with a small dataset.

Table 8: ANN and CFD results of the optimized shape of HV.

PC_1	PC_2	$R_{TM-ANN}[N]$	$R_{TM-CFD}[N]$	**Absolute percentage error (%)**
1.4896	0.1412	50.72	50.84	0.2361

Table 9 shows the effect of optimized HV on the resistance components. The camber-curve width of the optimized HV has slightly increased compared to that of the initial shape. This has a slightly negative effect on the frictional resistance. The stern wave elevation and dynamic pressure distribution of the vessel without HV, with the initial and the optimized HV, are shown in **Figure 18** and **19**. The stern zone pressure distribution is further reduced (see **Figure 19 (b2)** and **(c2)**) due to the increased pressure difference between the suction and pressure side of the optimized HV (see **Figure 17**). As a result of this study, R_{TM} is further decreased by 1.2% with the optimized shape of HV; see **Table 9**.

Table 9: Resistance comparison of optimized and initial HV.

	$R_{FM}[N]$	$R_{RM}[N]$	$R_{TM}[N]$	**Difference (%)**
Bare Hull	20.49	42.30	62.78	
Bare Hull w/initial HV	22.12	29.49	51.60	-17.81
Bare Hull w/optimized HV	22.28	28.56	50.84	-19.02

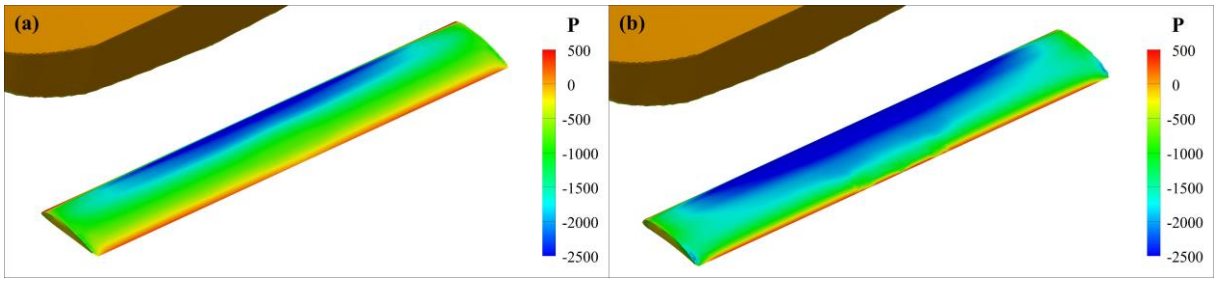

Figure 17: Dynamic pressure distribution along the (a) initial HV and (b) optimized HV.

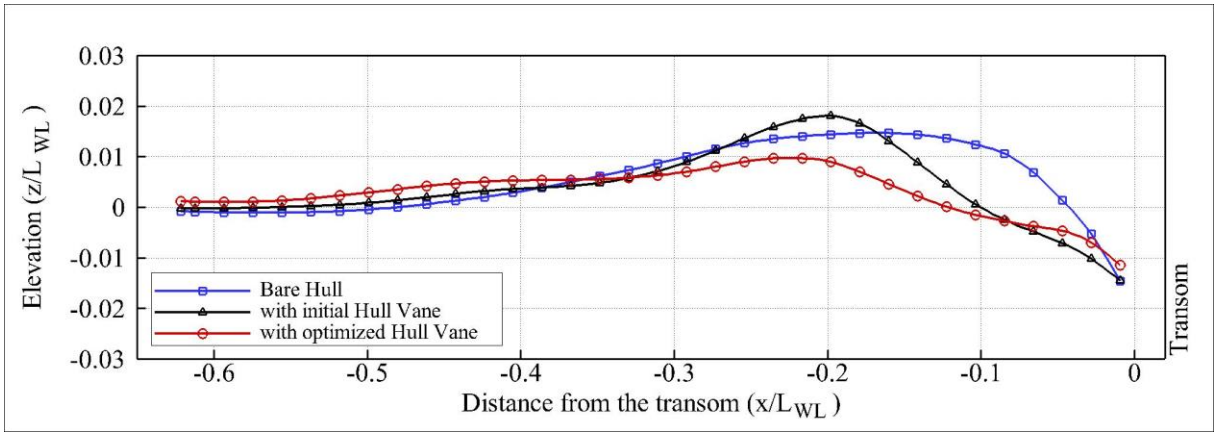

Figure 18: Comparison of wave elevations behind the transom taken from the vessel's centreline.

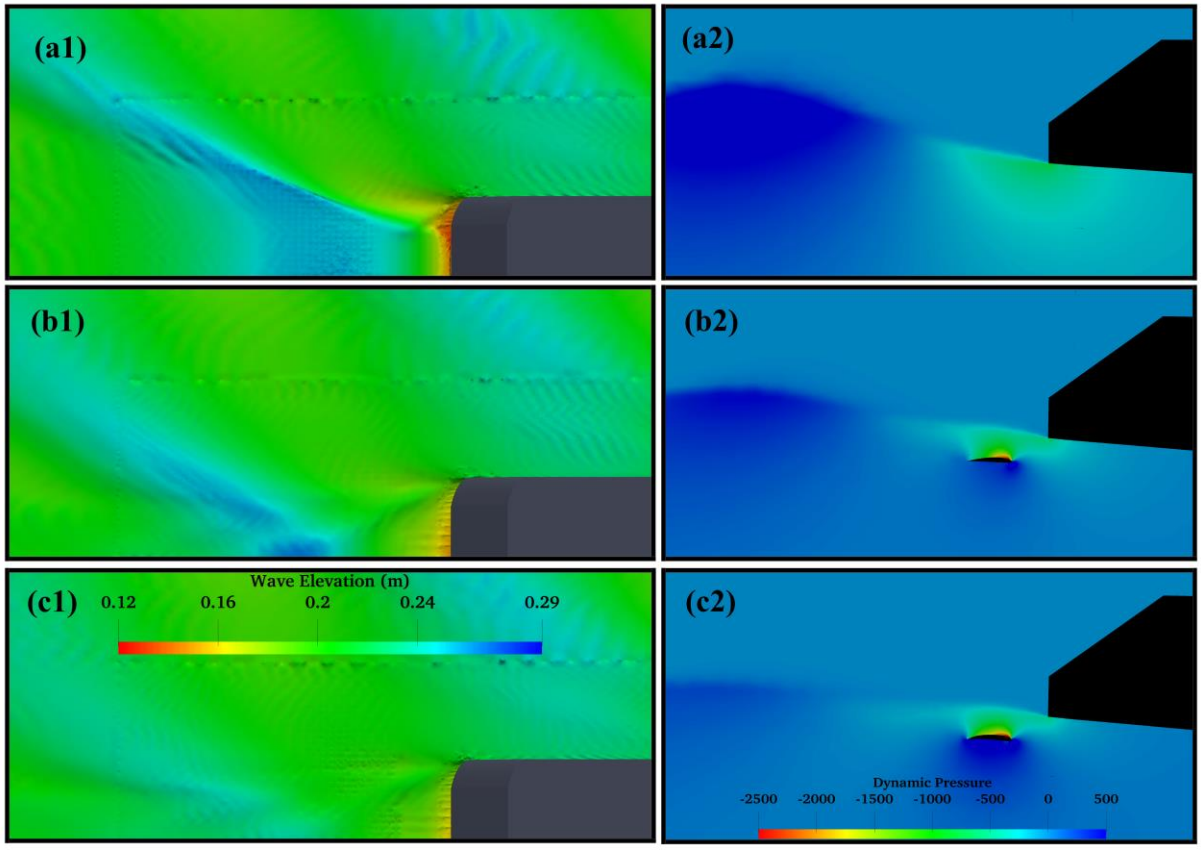

Figure 19: Wave elevation and dynamic pressure distribution of the vessel (a) without HV, (b) with initial HV and (c) with optimized HV.

Table 10 gives the trim and sinkage values of the vessel without HV (bare hull) and with the initial and optimized HV designs. The bare hull has 0.112 deg of bow down trim at $Fr = 0.37$. Due to the presence of the HV, the lift force occurs in the stern area of the vessel, which slightly increases the bow down trim of the vessel. In the presence of the optimized HV, the trim angle is 1.212 deg. The sinkage of the bare hull at the design speed is 0.0161m. The impact of the generated lift force at the stern is not only on the vessel's trim but also on the sinkage, which is further reduced with the optimized HV from 0. 0161m to 0.0098m on model scale.

Table 10: Comparison of the vessel's trim and sinkage.

	Trim (deg)	Sinkage (m)
Bare Hull	+0.112	-0.0161
Bare Hull w/initial HV	+0.915	-0.0117
Bare Hull w/optimized HV	+1.212	-0.0098

+: bow down

The impact of the initial and the optimized HV on the full-scale model has also been investigated. The scale ratio (λ) is 16.5, and the corresponding design speed (V_S) is 8.73 m/s from Froude similarity. Recently, **Hou et al. (2020)** carried out extrapolation from model scale to full scale and observed that HV directly impacts the hull form, hence on the change of the form factor. Therefore, in this study, the bare hull and the hull equipped with the initial and the optimized HV have been simulated on the model scale according to the double-body approach described in Section 5.2 in order to find the $(1 + k)$ and C_W, which are shown in **Table 11**. These values are the same for model and full scale due to Froude similarity. The C_{FS} values have been calculated in accordance with ITTC-1957 model-ship correlation.

$$C_{FS} = \frac{0.075}{(log_{10}Re - 2)^2} \tag{11}$$

The difference between C_{FS} values on the model and full scale are shown in **Table 11.** This difference is due to the fact that the chord length of the HV is also included when calculating the Re of the hull with HVs. Then, extrapolations are performed by implementing the standard ITTC procedure to estimate C_{TS} values. The corresponding R_{TS} values are calculated by the following formula,

$$R_{TS} = \frac{1}{2}\rho S_S V_S^2 C_{TS}, \tag{12}$$

where ρ is the density of the salt water and S_S is the wetted surface area of the ship hull.

It is observed that the effective power (P_E) decreases by 23.56% in the presence of initial HV and is further reduced by 3.42% with the optimized HV (see **Table 11**).

Table 11: Extrapolated results from model scale to full scale

	Model Scale				Full Scale						
	10^3C_{FM}	10^3C_{VP}	$1+k$	10^3C_W	10^3C_{FS}	$1+k$	10^3C_W	10^3C_{TS}	$R_{TS}[kN]$	$P_E[kW]$	**Diff. (%)**
Bare Hull	3.201	0.913	1.285	5.696	1.709	1.285	5.696	7.893	232.81	2033.21	-
Bare Hull w/initial HV	3.277	0.730	1.223	3.639	1.703	1.223	3.639	5.721	177.97	1554.25	23.56
Bare Hull w/optimized HV	3.299	0.973	1.295	3.256	1.703	1.295	3.256	5.461	169.99	1484.62	26.98

To calculate the brake power (P_B) and fuel consumption, one requires the propulsive efficiency (η_D) which will be different without HV and the initial-optimized HV. This is due to the fact that thrust deduction will be changed in the presence of the HV. Besides, the variables, such as engine load and specific fuel oil consumption (SFOC), contribute to fuel consumption and depend on the operating conditions at different engine loadings. Instead of calculating the fuel consumption directly, the percentage decrement in fuel consumption in terms of P_E has been evaluated with equation 13 **(Tezdogan et al., 2015)**, which can be considered the indication of fuel consumption assuming the efficiencies and SFOC remain constant.

$$\% \text{ decrement in fuel consumption in terms of } P_E = \frac{\Delta P_E}{P_{E,woHV}} = \frac{P_{E,woHV} - P_{E,wHV}}{P_{E,woHV}} \times 100. \quad (13)$$

Table 12: Estimation of fuel consumption in percentage.

	$P_E[kW]$	**Fuel Consumption (%)**
Bare Hull	2033.21	100.00
Initial (NACA4412)	1554.25	76.44
Optimized	1484.62	73.02

6. CONCLUSIONS and FUTURE WORK

The present paper proposes a data-driven design technique for shape optimization of an energy-saving appendage, namely HV (Hull Vane), for resistance reduction of a motor yacht hull. At the heart of this technique are PCA (Principal-Component Analysis), a dimension-reduction method, and an ML- (Machine Learning) based model, ANN (Artificial Neural Network), for hull resistance prediction in the presence of HV. PCA overcomes the hurdle of dimensionality and expedites the design space exploration during optimization, while ANN helps to bypass the need for time-consuming CFD simulations for design evaluation. The proposed method commences with an initial design of HV, whose cross-sectional profile is defined with a NACA4412 hydrofoil parameterized with seven design parameters. Next, a design space is formed around the baseline design, which was then sampled to create a dataset. PCA was implemented on this dataset, which created a two-dimensional latent subspace representation of the original design space while preserving 85.5% of the geometric variability. This

dimension reduction reduces the need for extensive simulation data for training ANN and helps to simplify its architecture. The two-dimensional design space was sampled again. The samples' total resistance was evaluated with our high-fidelity CFD (viscous) solver after projecting design to the original design space. The new dataset contained two PCs as independent variables and total resistance as a dependent variable and was used to train a feedforward ANN. The total resistance for the optimized design obtained from the prediction model are in strong agreement with the results obtained from the CFD solver, which validates that even when trained with few samples, the developed model can provide reliable results. This model was later integrated with an interior point optimization technique, exploring the two-dimensional subspace for an optimal design. The optimized design obtained from this process reduces the total resistance of the hull by 1.2% on the model scale compared to that of the baseline HV design, which is based on NACA4412 hydrofoil. The effective power is further decreased by 3.72% with the optimized HV on a full scale.

Although the uncertainty analysis and experimental benchmarks verified the robustness and accuracy of the flow solver, this study can be further strengthened via an empirical study, which will explore the HV influence on the various hydrodynamic characteristics of the ship, such as motions manoeuvring and propulsive efficiencies. Furthermore, in the presence of an optimized HV, the hull's shape can also be optimized in conjunction with parametric **(Khan et al., 2017)** and generative **(Khan et al., 2019)** tools.

REFERENCES

Alwosheel, A., van Cranenburgh, S., & Chorus, C. G. (2018). Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis. *Journal of Choice Modelling*, *28*(July), 167–182. https://doi.org/10.1016/j.jocm.2018.07.002

Arnold, T. B. (2016). *Introduction to neural networks* [PowerPoint presentation]. Yale University. http://euler.stat.yale.edu/~tba3/stat665/lectures/lec12/lecture12.pdf

Bouckaert, B., Uithof, K., Moerke, N., & Van Oossanen, P. G. (2016). *Hull Vane ® on 108m Holland-Class OPVs : Effects on Fuel Consumption and Seakeeping*. 12. http://www.hullvane.nl/wp-content/uploads/2017/01/MAST-2016-Paper-Hull-Vane-on-108-m-Holland-Class-OPV-effects-on-fuel-consumption-and-seakeeping.pdf

Celik, C., Danisman, D. B., Kaklis, P., & Khan, S. (2020). An investigation into the effect of the hull vane on the ship resistance in openfoam. *Sustainable Development and Innovations in Marine Technologies - Proceedings of the 18th International Congress of the International Maritime Association of the Mediterranean, IMAM 2019*, 136–141. https://doi.org/10.1201/9780367810085-17

Celik, Ghia, U., Roache, P. J., Freitas, C. J., Coleman, H., & Raad, P. E. (2008). Procedure for estimation and reporting of uncertainty due to discretization in CFD applications. *Journal*

of Fluids Engineering, Transactions of the ASME, *130*(7), 0780011–0780014. https://doi.org/10.1115/1.2960953

Cepowski, T. (2007). Application of artificial neural networks to approximation and identification of seakeeping performance of a bulk carrier in ballast loading condition. *Polish Maritime Research*, *14*(4), 31–39. https://doi.org/10.2478/v10012-007-0037-6

Constantine, P. G. (2015). *Active subspaces: Emerging ideas for dimension reduction in parameter studies*. Society for Industrial and Applied Mathematics.

D'Agostino, D., Serani, A., Diez, M., & Campana, E. F. (2018). Deep autoencoder for off-line design-space dimensionality reduction in shape optimization. *AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, 2018*, *210049*. https://doi.org/10.2514/6.2018-1648

D'Agostino, D., Serani, A., & Diez, M. (2020). Design-space assessment and dimensionality reduction: An off-line method for shape reparameterization in simulation-based optimization. Ocean Engineering, 197, 106852.

Danışman, D. B. (2014). Reduction of demi-hull wave interference resistance in fast displacement catamarans utilizing an optimized centerbulb concept. *Ocean engineering*, 91, 227-234.

Delen, C., Can, U., & Bal, S. (2020). Prediction of Resistance and Self-Propulsion Characteristics of a Full-Scale Naval Ship by CFD-Based GEOSIM Method. *Journal of Ship Research*, 1-16.

Diez, M., Campana, E. F., & Stern, F. (2015). Design-space dimensionality reduction in shape optimization by Karhunen-Loève expansion. *Computer Methods in Applied Mechanics and Engineering*, *283*, 1525–1544. https://doi.org/10.1016/j.cma.2014.10.042

Diez, M., Volpi, S., Serani, A., Stern, F., & Campana, E. F. (2019). Simulation-based design optimization by sequential multi-criterion adaptive sampling and dynamic radial basis functions. In *Advances in Evolutionary and Deterministic Methods for Design, Optimization and Control in Engineering and Sciences* (pp. 213–218).

Fukunaga, K., & Koontz, W. L. G. (1970). Application of the Karhunen-Loéve Expansion to Feature Selection and Ordering. *IEEE Transactions on Computers*, *C–19*(4), 311–318. https://doi.org/10.1109/T-C.1970.222918

Gaggero, S., Vernengo, G., Villa, D., & Bonfiglio, L. (2019). A reduced order approach for optimal design of efficient marine propellers. *Ships and Offshore Structures*, *0*(0), 1–15. https://doi.org/10.1080/17445302.2019.1606877

Gunpinar, E., Coskun, U. C., Ozsipahi, M., & Gunpinar, S. (2019). A Generative Design and Drag Coefficient Prediction System for Sedan Car Side Silhouettes based on Computational Fluid Dynamics. *CAD Computer Aided Design*, *111*, 65–79. https://doi.org/10.1016/j.cad.2019.02.003

Gunpinar, E., & Khan, S. (2020). A multi-criteria based selection method using non-dominated sorting for genetic algorithm based design. Optimization and Engineering, 21(4), 1319-1357. https://doi.org/10.1007/s11081-019-09477-8

Hagan, M. T., & Menhaj, M. B. (1994). Training feedforward networks with the Marquardt algorithm. IEEE transactions on Neural Networks, 5(6), 989-993.

Hagemeister, N., Uithof, K., Bouckaert, B., & Mikelic, A. (2017). HULL VANE ® VERSUS LENGTHENING - A comparison between four alternatives for a 61m OPV. *FAST Conference 2017*, 1–11.

Hamdia, K. M., Ghasemi, H., Bazi, Y., AlHichri, H., Alajlan, N., & Rabczuk, T. (2019). A novel deep learning based method for the computational material design of flexoelectric nanostructures with topology optimization. *Finite Elements in Analysis and Design*, *165* (August), 21–30. https://doi.org/10.1016/j.finel.2019.07.001

Hou, H., Krajewski, M., Ilter, Y. K., Day, S., Atlar, M., & Shi, W. (2020). An experimental

investigation of the impact of retrofitting an underwater stern foil on the resistance and motion. *Ocean Engineering*, *205*(November 2019), 107290. https://doi.org/10.1016/j.oceaneng.2020.107290

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *science*, *313*(5786), 504-507.

Hu, Z. Z., Greaves, D., & Raby, A. (2016). Numerical wave tank study of extreme waves and wave-structure interaction using OpenFoam®. *Ocean Engineering,* 126, 329-342.

Islam, H., & Guedes Soares, C. (2018). Estimation of hydrodynamic derivatives of a container ship using PMM simulation in OpenFOAM. *Ocean Engineering*, *164*(July), 414–425. https://doi.org/10.1016/j.oceaneng.2018.06.063

ITTC. (2011). Practical Guidelines for Ship CFD Applications. *ITTC – Recommended Procedures and Guidelines ITTC*, 1–8.

Khan, S., Gunpinar, E., & Mert Dogan, K. (2017). A novel design framework for generation and parametric modification of yacht hull surfaces. *Ocean Engineering*, *136* (November 2016), 243–259. https://doi.org/10.1016/j.oceaneng.2017.03.013

Khan, S., Gunpinar, E., & Sener, B. (2019). GenYacht: An interactive generative design system for computer-aided yacht hull design. *Ocean Engineering*, *191*(September), 106462. https://doi.org/10.1016/j.oceaneng.2019.106462

Khan, S., & Gunpinar, E. (2018). Sampling CAD models via an extended teaching–learning-based optimization technique. Computer-Aided Design, 100, 52-67. https://doi.org/10.1016/j.cad.2018.03.003

Khan, S., Serani, A., Diez, M., & Kaklis, P. (2021). Physics-Informed Feature-to-Feature Learning for Design-Space Dimensionality Reduction in Shape Optimisation. In AIAA Scitech 2021 Forum (p. 1235). https://doi.org/10.2514/6.2021-1235

Kostas, K. V., Ginnis, A. I., Politis, C. G., & Kaklis, P. D. (2015). Ship-hull shape optimization with a T-spline based BEM-isogeometric solver. *Computer Methods in Applied Mechanics and Engineering*, *284*, 611–622. https://doi.org/10.1016/j.cma.2014.10.030

Kostas, K. V., Ginnis, A. I., Politis, C. G., & Kaklis, P. D. (2017). Shape-optimization of 2D hydrofoils using an Isogeometric BEM solver. *CAD Computer Aided Design*, *82*, 79–87. https://doi.org/10.1016/j.cad.2016.07.002

Li, J., He, S., & Martins, J. R. R. A. (2019). Data-driven constraint approach to ensure low-speed performance in transonic aerodynamic shape optimization. *Aerospace Science and Technology*, *92*, 536–550. https://doi.org/10.1016/j.ast.2019.06.008

Liu, X., Wan, D., & Chen, G. (2018, June). Kriging-based surrogate model combined with weighted expected improvement for ship hull form optimization. In *International Conference on Offshore Mechanics and Arctic Engineering* (Vol. 51265, p. V07AT06A012). American Society of Mechanical Engineers.

Margari, V., Kanellopoulou, A., & Zaraphonitis, G. (2018). On the use of Artificial Neural Networks for the calm water resistance prediction of MARAD Systematic Series' hullforms. *Ocean Engineering*, *165*(July), 528–537. https://doi.org/10.1016/j.oceaneng.2018.07.035

Masood, Z., Khan, S., & Qian, L. (2021). Machine Learning-Based Surrogate Model for Accelerating Simulation-Driven Optimization of Hydropower Kaplan Turbine. *Renewable Energy, 173*, 827-848. https://doi.org/10.1016/j.renene.2021.04.005

Mikulić, A., & Parunov, J. (2019). A review of artificial intelligence applications in ship structures. *Trends in the Analysis and Design of Marine Structures*, 515.

Nocedal, J., & Wright, S. J. (2006). Numerical Optimization. In *Numerical Optimization*. https://doi.org/10.1007/978-0-387-40065-5

Palmer, S., Gorse, D., & Muk-Pavic, E. (2015). *Neural Networks and Particle Swarm Optimization for Function Approximation in Tri-SWACH Hull Design*. 1–6.

https://doi.org/10.1145/2797143.2797168

Pope, S. (2002). Turbulent Flows. In *European Journal of Mechanics - B/Fluids* (Vol. 21, Issue 1). https://doi.org/10.1016/s0997-7546(01)01166-9

Rojas, R. (2013). *Neural networks: a systematic introduction*. Springer Science & Business Media.

Samareh, J. A. (2001). Survey of shape parameterization techniques for high-fidelity multidisciplinary shape optimization. *AIAA Journal*, *39*(5), 877–884. https://doi.org/10.2514/2.1391

Scholcz, T. P., Gornicz, T., & Veldhuis, C. (2015). Multi-objective hull-form optimization using Kriging on noisy computer experiments. *In MARINE VI: proceedings of the VI International Conference on Computational Methods in Marine Engineering* (pp. 1064-1077). CIMNE.

Serani, A., Pellegrini, R., Wackers, J., Jeanson, C. E., Queutey, P., Visonneau, M., & Diez, M. (2019). Adaptive multi-fidelity sampling for CFD-based optimization via radial basis function metamodels. *International Journal of Computational Fluid Dynamics*, *33*(6–7), 237–255. https://doi.org/10.1080/10618562.2019.1683164

Serani, Andrea, D'Agostino, D., Campana, E. F., & Diez, M. (2019). Assessing the Interplay of Shape and Physical Parameters by Unsupervised Nonlinear Dimensionality Reduction Methods. *Journal of Ship Research*, *August*. https://doi.org/10.5957/josr.09180056

Sharma, A., Paliwal, K. K., Imoto, S., & Miyano, S. (2013). Principal component analysis using QR decomposition. International Journal of Machine Learning and Cybernetics, 4(6), 679-683. https://doi.org/10.1007/s13042-012-0131-7

Solak, H. P. (2020). Multi-dimensional surrogate based aft form optimization of ships using high fidelity solvers. *Brodogradnja: Teorija i praksa brodogradnje i pomorske tehnike,* 71(1), 85-100.

Stein, M. (1987). Large sample properties of simulations using latin hypercube sampling. *Technometrics*, *29*(2), 143–151. https://doi.org/10.1080/00401706.1987.10488205

Sun, L., Gao, H., Pan, S., & Wang, J. X. (2020). Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data. *Computer Methods in Applied Mechanics and Engineering*, *361*, 112732. https://doi.org/10.1016/j.cma.2019.112732

Tezdogan, T., Demirel, Y. K., Kellett, P., Khorasanchi, M., Incecik, A., & Turan, O. (2015). Full-scale unsteady RANS CFD simulations of ship behaviour and performance in head seas due to slow steaming. *Ocean Engineering*, 97, 186-206.

Tezzele, M., Salmoiraghi, F., Mola, A., & Rozza, G. (2018). Dimension reduction in heterogeneous parametric spaces with application to naval engineering shape design problems. *Advanced Modeling and Simulation in Engineering Sciences*, *5*(1). https://doi.org/10.1186/s40323-018-0118-3

Uithof, K., Bouckaert, B., Van Oossanen, P. G., & Moerke, N. (2016a). The effects of the hull vane on ship motions of ferries and ropax vessels. *RINA, Royal Institution of Naval Architects - Design and Operation of Ferries and Ro-Pax Vessels 2016*, *May*, 25–26.

Uithof, K., Bouckaert, B., Van Oossanen, P. G., & Moerke, N. (2016b). *A Cost-Benefit Analysis of Hull Vane Application on Motor Yatchs*. *November*, 1–10.

Uithof, K., Hagemeister, N., Bouckaert, B., Van Oossanen, P. G., & Moerke, N. (2017). A systematic comparison of the influence of the hull Vane®, inter-ceptors, trim wedges, and ballasting on the performance of the 50M amecrc series #13 patrol vessel. *RINA, Royal Institution of Naval Architects - Warship 2016: Advanced Technologies in Naval Design, Construction, and Operation*, *May*.

Uithof, K., Oossanen, P. van, Moerke, N., Oossanen, P. G. van, & Zaaijer, K. S. (2014). *AN UPDATE ON THE DEVELOPMENT OF THE HULL VANE*. 211–221.

Wang, S., Ji, B., Zhao, J., Liu, W., & Xu, T. (2018). Predicting ship fuel consumption based on

LASSO regression. *Transportation Research Part D: Transport and Environment*, *65*(October 2017), 817–824. https://doi.org/10.1016/j.trd.2017.09.014

Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and intelligent laboratory systems*, *2*(1-3), 37-52.

Yu, D., Wang, L., Zhong, Q., & Yeung, R. W. (2019, June). Evaluation and Optimization of Trimaran Configurations Using Deep Neural Network. In *International Conference on Offshore Mechanics and Arctic Engineering* (Vol. 58844, p. V07AT06A013). American Society of Mechanical Engineers.

Zhang, C., Zhang, D., Zhang, M., & Mao, W. (2019). Data-driven ship energy efficiency analysis and optimization model for route planning in ice-covered Arctic waters. *Ocean Engineering*, *186*(June), 106071. https://doi.org/10.1016/j.oceaneng.2019.05.053