Application of machine learning to design low noise propellers

Doijode, P.S.

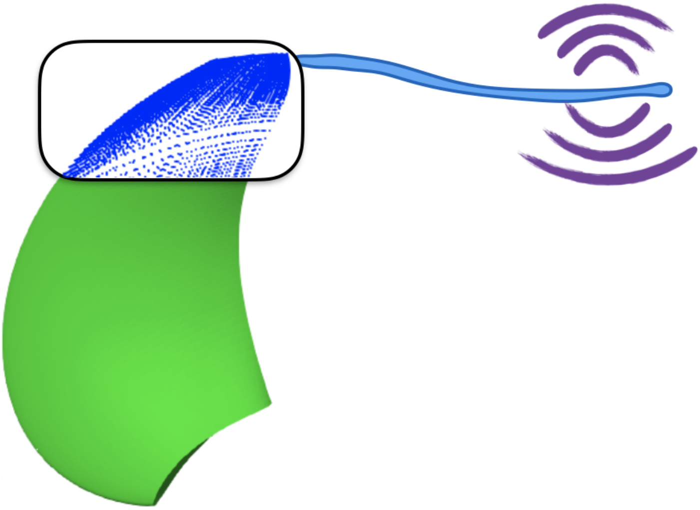
DOI
10.4233/uuid:ed855d71-19c5-45fb-b678-581ea1cecfaf

Publication date
2022

Document Version
Final published version

Citation (APA)
Doijode, P. S. (2022). *Application of machine learning to design low noise propellers*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:ed855d71-19c5-45fb-b678-581ea1cecfaf

APPLICATION OF MACHINE LEARNING TO DESIGN LOW NOISE PROPELLERS



APPLICATION OF MACHINE LEARNING TO DESIGN LOW NOISE PROPELLERS

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. T.H.J.J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op Dinsdag 6 September 2022 om 12:30 uur

door

Pranav Sumanth DOIJODE

Ingenieur Luchtvaart & Ruimtevaart,
Technische Universiteit Delft, Nederland,
geboren te Bangalore, India.

This dissertation has been approved by the promotors.

Rector Magnificus,	chairperson
prof. dr. ir. T.J.C. van Terwisga,	Delft University of Technology, Promotor
prof. dr. -ing. S Hickel,	Delft University of Technology, Promotor
ir. K. Visser,	Delft University of Technology, Copromotor
Independent members:	
Dr. A. Corradu	3mE, Delft University of Technology
Dr. L. Pahlavan	3mE, Delft University of Technology
Prof.dr. R. Bensow	Chalmers U., Sweden
Prof.dr.ir. L.L.M. Veldhuis	LR, Delft University of Technology
Prof.ir. J.J. Hopman	3mE, TU Delft

Keywords: Explainable Machine Learning, Propeller Design, ILES, Barotropic model

Printed by: Technishe Universitiet Delft

Front : A picture of a propeller blade with geometry probes and a noisy excited cavitating tip vortex

ISBN 978-94-6419-576-7

High uncertainty can cost us just a battle, but if we remain indecisive in uncertainty it may very well cost us everything and much more.

CONTENTS

SUMMARY

Over 90% of international trade is carried out over seas. Shipping is currently the cheapest mode of transoceanic transport. The traffic density of shipping lanes on seas, oceans, and also rivers is likely to increase. Consequently, the GHG, NOx, SOx and noise emissions from shipping will rise making it more difficult to meet stricter emission regulations which the IMO aims to achieve. One opportunity to reduce emissions is by designing more efficient and quieter propellers.

To design quieter and more efficient propellers an optimal blade loading solution is required. For a rigid propeller, the blade loading distribution is optimized by modifying the geometry. The propeller geometry must be modified to achieve optimal loading that maximizes efficiency and minimizes acoustic emissions. In addition to efficiency and noise considerations, propeller optimization must consider thrust, ship speed, fairing constraints as well as unsteady wake of the vessel.

Most modern propeller geometries are optimized to the bounds of the capabilities of low, high fidelity simulations and scaled experiments. Even in the preliminary design phase, optimizing a propeller is resource intensive. State-of-the-art optimization methods have used surrogate based methods and machine learning to improve the efficiency of optimization. Typically, optimization methods use (a) classical design variables to define the propeller geometry, (b) gradient free search algorithms to explore the design space and, (c) statistical models to learn the objectives and constraints. While they have been demonstrated to find well defined Pareto fronts for multi-objective constrained propeller optimization problems, they face three challenges. The *first* challenge is that most optimization methods tend to approach the optimization problem with classical design variables however, they lead to the problem of multicollinearity. The performance prediction of a propeller geometry in Boundary Element Method (BEM) or Computational Fluid Dynamics (CFD) depends on the discretized geometry i.e. mesh. Changing a single classical variable results in a change of multiple nodes on the mesh. Thus, the variables are not orthogonally independent w.r.t performance predictions. This orthogonal dependence is multicollinearity. It obscures the correlation between the mesh and predicted efficiency or cavitation behaviour. As a result, more than required evaluations may be necessary to train statistical models particularly when sensitive cavitation constraints are considered. This makes optimization inefficient. The *second* challenge is that preliminary design tools tend to have high uncertainties. These uncertainties result from lower-fidelity physical models, use of semi-empirical relations to predict skin-friction, numerical methods and variance in operational parameters. Consequently, comparable or similar efficiency and cavitation behaviour may be predicted for two designs even though their operational performances may be significantly different. This could lead optimization methods to converge to sub-optimal solutions. The *third* challenge is that computational cost of optimization for realistic cases is relatively high even in the preliminary design phase. Consequently simpler methods such as charting from open-water curves of legacy propeller families such as Wageningen series are preferred. However, as shape adaptive composite propellers become more prevalent, propeller design and optimization cannot resort to charting. In such circumstances improving efficiency of optimization process is crucial. In this body of work, each of the aforementioned three challenges are addressed.

In order to solve the first challenge of multicollinearity, a new parametric model is required whose features are orthogonally independent. Thus, the orthogonal parametric model is proposed. It is constructed by projecting the mesh of a geometry in a hyperspace defined by several orthonormal vectors. The orthonormal vectors are derived from Singular Value Decomposition (SVD) of all possible geometric variations of propeller meshes. The projections of a mesh in the hyperspace is demonstrated to be a viable method to accurately quantify geometric variations and also establish sensitives of performance w.r.t design variables. If performance correlates linearly with geometric variation, SVD can be used to identify sensitivities of design parameters a-priori. In situations when performance correlates non-linearly, SVD can be used to selectively sample the design space reducing objective function evaluations by almost 50%. This is demonstrated on the aeroacoustic optimization of a 2D airfoil. The trade-off with completeness is also found to be reasonable.

In order to solve the second challenge of preliminary design tools leading optimization to sub-optimal solutions, it is proposed that uncertainties be modelled and optimization methods favour designs with a better performance and no overlap in the 95 % confidence interval of performance prediction. This is possible when the variance in performance prediction is known and can be modelled. Soft regression and classification can be used to map the orthogonal parameters to mean and variance in predictions of efficiency and cavitation behaviour. Furthermore, the mean and variance predictions made by soft regression and classification can be linked to standard distributions resulting in confidence intervals. Thus, designs with better performance and no overlap in 95 % confidence intervals can be considered to dominate during search iterations.

For BEM, uncertainty resulting from physical modelling, use of semi-empirical relations and numerical methods are well investigated and documented. Thus, based on data from literature the variance in efficiency prediction of BEM is modelled. However, sources of uncertainties in Empirical Tip Vortex (ETV) model are not fully investigated. Consequently, it is difficult to model variance in predictions of broadband acoustics made by ETV. A detailed insight into sources of uncertainty and extant of variance in prediction of tip vortex size and broadband noise requires investigating the governing physics. Therefore, the dynamics of an isolated cavitating vortex without forced excitation is investigated with scale resolved compressible flow simulations. Observations indicate the presence of Kelvin-Helmholtz instabilities due to vapour-liquid interface. These instabilities grow resulting in helical and superimposed response modes. On the relevance of compressibility, it is observed that the vapour core tends to be compressible.

Density variations in the free-stream are negligible and predominantly due to acoustics. The free-stream may influence the dynamics within the cavity however, the momentum within the core is most likely too low to influence free-stream dynamics. This indicates that the ETV model's incompressibility assumption is most consistent with our observations. As compressibility is not expected to play a dominant role, the ETV is hypothesized to be most reliable in predicting the size of the vortex cavity which depends mostly on blade tip loading. Thus, in this research the ETV model is used to predict only the radius of the tip vortex. Due to the direct correlation with blade tip loading and for simplicity, it is assumed that variance in tip vortex core size prediction is the same as that of efficiency prediction documented in literature. With soft regression and classification, the mean and variance in efficiency and radius of the tip vortex core are modelled.

In order to address the third challenge of high computational cost of optimization, it is proposed that explainable supervised soft classifiers be trained to identify the location of the Pareto front. The classifiers advise search strategies to focus on lucrative regions of the design space. Furthermore, explainable machine learning models also provide insight into *why* a certain region is lucrative. The classifiers are trained with orthogonal features which capture geometric variation in radial distribution of pitch, skew, camber and chordlength. The method is demonstrated on a cavitating, unsteady flow case of Wageningen B-4 70 propeller with P/D=1.0 operating in the Seiun-Maru wake. Compared to the classical Non-dominated Sorting Genetic Algorithm - III (NSGA-III) the optimization method is able to reduce 30 % of evaluations per generation while reproducing a comparable Pareto front. Trade-offs between suction side, pressure side, tip-vortex cavitation and efficiency are identified. The non-elitist NSGA-III search algorithm in conjunction with the supervised classifiers are able to identify a Pareto front with very diverse solutions. Among the solutions, a design with no pressure side cavitation, low suction side cavitation and reasonable tip-vortex cavitation is found.

By addressing the three challenges, the current body of work contributes to new propeller optimization methods which are expected to be more efficient.

1

INTRODUCTION

1.1. BACKGROUND

Over 90% of international trade is carried out over seas [55]. The shipping routes shows that traffic cuts across the worlds seas and oceans predominantly transporting containers, bulk cargo, gas, oil, liquids and vehicles [1]. Fundamentally, this demand for commercial shipping is created by Global Value Chains (GVCs) where different stages of production processes are located in different countries [42] based on the respective country's comparative advantages. For the economy to benefit from these comparative advantages, it is a critical requirement to transport intermediate products to centres of manufacturing and finished products to markets at the lowest possible cost. Shipping can deliver on this requirement being the cheapest mode of transport.

As the economy recovers from impacts due to Covid-19, global trade is expected to increase and GVCs are likely to be more robust. Thus, the traffic density of shipping lanes on seas, oceans, and also rivers is likely to increase. Consequently, the Green House Gas (GHG) and noise emissions from shipping will rise making it more difficult to meet stricter emission regulations [106] and noise reduction guidelines [105]. The International Maritime Organization (IMO) aims to achieve a 50% reduction in GHG emissions by 2050 compared to 2008 levels [70] and ports have also started to incentivize silent ships with harbour due rate discounts [102].

There are several sources of airborne and underwater noise from ships [4]. For special purpose vessels mounted equipment can be a significant source of airborne noise. Compared to air, noise travels faster and farther in water. Underwater noise is generated by water flow on vessels, auxiliary machinery and equipment, diesel generators, prime movers, electric motors and propellers. Most modern propellers are typically well designed and tip-vortex cavitation, see Figure 1.1 is often the observed mode of cavitation. Averson et. al [8] report the underwater radiated noise for M/V Overseas Harriette at different ship speeds (propeller rpm). The ship has a cavitation inception speed of

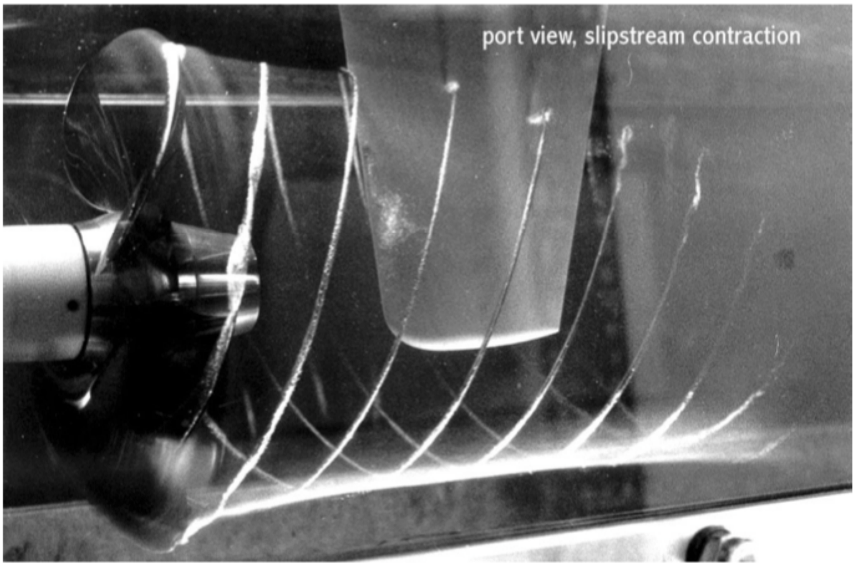


Figure 1.1: Cavitating tip vortex with slip stream contraction [56].

10 knots which corresponds to 86 RPM for the propeller. Figure 1.2 illustrates the one-third octave broadband spectrum for the vessel. It is observed at cavitation inception, the noise hump between 50-100 Hz gets more pronounced with increasing speed. This noise hump is typically associated with the broad band spectrum of tip vortex noise [108] which is often the dominant source of underwater radiated noise. Against this background, there is strong emphasis on making ship propellers more quiet and more efficient.

To design quieter and more efficient propellers an optimal blade loading is required. Blade loading plays an important role in propeller design as blade sections are continuously exposed to different angles of attack in the non-uniform wake-field they operate in. For a rigid propeller, the blade loading distribution is optimized by modifying the geometry. Traditionally, in addition to efficiency objectives indirect noise constraints as a limit on cavitation volume or area [64, 93] have been implemented. Acoustic objectives have been considered [101] but the focus has been limited to pressure fluctuations at blade passage frequencies. Most modern propellers are typically optimised to limiting bounds of performance and are more likely to operate with vortex cavitation as the dominant source of radiated pressure fluctuations and broadband noise [8]. In addition to noise and efficiency design objectives, propeller design must also account for constraints on hull excitation [88], cavitation erosion [111], ship speed, thrust and strength. State of the art optimization methods are able to tackle this multi-objective constrained problem. However, there are opportunities to improve current methods to make them more computationally efficient while still identifying lucrative designs in a vast design space.

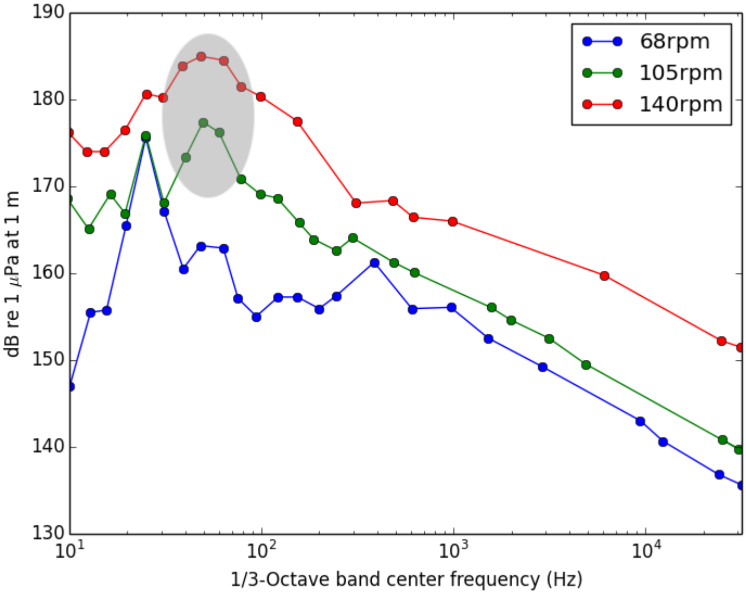


Figure 1.2: Third-octave broadband spectrum of M/V Overseas Harriette at different ship speeds [8]. The hump due to tip-vortex cavitation is highlighted.

1.2. STATE OF THE ART

The hydroacoustic design and optimisation problem of propellers is of immense strategic importance and is most likely well researched in defence. However, there is limited knowledge and reporting in the public domain. Access to specialized simulation software, models based on propriety data and intellectual property restrictions (due to sensitivity or embargo) result in hydroacoutic optimization of propellers being tractable to very few academic research groups. In North America, Yin Lu Young's group in Michigan university focus on flexible composite propellers [113] as a way towards designing quieter and more efficient propellers. One of the focus areas at the group is to design shape adaptive propellers which have a bend twist coupling to optimize propeller loading [97, 100, 109] and improve efficiency. Given the challenging structural modelling for composites [98], cavitation and noise constraints are secondary considerations. In Europe, MARIN has focused on developing and demonstrating tools for hydroacoustic design and optimization of propellers [27, 95]. Florian Vesting in Rickard Bensow's group has investigated algorithms, strategies and methods for hydrodynamic optimization of propellers with cavitation considerations [26, 84, 112]. In Asia, Nakashima propellers have focused on achieving better cavitation behaviour and improved efficiency for large vessels [110]. Better cavitation behaviour is expected to result in lower cavity volume fluctuations and thus lower pressure fluctuations and noise.

Up to approximately 2015, the focus in the aforementioned studies was predominantly on optimizing for efficiency and controlling sheet cavitation. In parallel, efforts were underway to better predict tip-vortex cavitation behaviour with simulations [29, 104]. These efforts continue even today as challenges related to physical modelling and numerical dissipation remain [96]. Thus, while propellers with better sheet cavitation behaviour were designed and introduced, the dominant source of noise was tending to be the cavitating tip-vortex (see Figure 1.2). However, predicting tip-vortex cavitation on propellers continues to be quite resource intensive for optimization. Furthermore, developing acoustic models for cavitating tip vortices based on simulation data requires a few challenges to be addressed (see Chapter 3). Recently, Bosschers .J [15] proposed the emperical tip vortex model which offers one approach to hydroacoutic optimization of propellers with tip-vortex considerations.

Most optimization methods that handle a maximise-efficiency objective are typically constrained with a minimum thrust requirements and maximum structural stress limits. Where an acoustic objective is also considered [101], the focus was limited to pressure fluctuations at blade passage frequencies. These pressure fluctuations tend to have low frequencies. Other examples have generally implemented indirect noise constraints as a limit on cavitation volume or area [64, 93]. As the only example, Huisman et. al [95] found a trade-off between tip-vortex cavitation noise and efficiency.

Summarizing most of the methods in literature, the author finds that Figure 1.3 illustrates the optimization method that is predominantly used in propeller optimization problems. At the start, objectives and constraints are specified (block 1). After objectives and constraints, the design variables are determined (block 2). Most parametric models rely on (a) classical design variables - pitch, diameter, hydrofoil camber, thickness distribution, skew and rake [6, 66], (b) control points of splines or coefficients of equations that define propeller families [62, 67, 85] and, (c) free-form control points [68]. However, they could lead to multicollinearity [2, 47]. Multicollinearity makes it difficult to isolate the impact of changing individual design variables on objectives and constraints. Therefore, the computational cost of optimization could increase as it could take more evaluations to try and isolate the individual impact of design variables. Consider a function $\phi = f(\vec{M}(\vec{s}))$, where $\vec{M}$ represents the mesh and $\vec{s}$ is the design vector. By chain rule, $\frac{\partial f}{\partial s} = \frac{\partial f}{\partial \vec{M}}.\frac{\partial \vec{M}}{\partial \vec{s}} = J_1.J_2$ where J_1 is the Jacobian of the function w.r.t the mesh and J_2 is the Jacobian of the mesh w.r.t the design variables. The optimization problem is free of multicollinearity if both J_1 and J_2 are orthonormal matrices. The orthonormality of J_1 is determined by the flow physics and J_2 by the chosen design variables. Any design variable which affects multiple nodes on the mesh will lead to J_2 which is not orthonormal. This is because the blade surface is often defined with splines thus, change in mesh nodes are strongly correlated and also predictable. Different classical design variables also tend to change the same mesh nodes. Thus, it is difficult to isolate the impact of individual design variables on objectives and constraints.

After defining the design variables, a search strategy is chosen (block 3). For the multi-objective constrained optimization problem Non-dominated Sorting Genetic Algorithm - II (NSGA-II) and Particle Swarm Optimisation (PSO) are mostly used as they are able to navigate a complex and discontinuous design space [57]. Among search strategies, NSGA-III is also a promising development which has been demonstrated on 3 to 15 objective optimization problems with convex, concave, disjointed and differently scaled Pareto fronts [90].

After the search strategy is chosen, the queried geometries are generated (block 4). The geometries are then evaluated, usually using panel methods or CFD (block 5). In most cases, to reduce the cost of several tens of thousand evaluations, Response Surface (RS) methods or statistical machine learning approaches are employed to learn the relation between objectives, constraints and design variables (block 5a). Most commonly Artificial Neural Networks, Krigging, iKrigging, Cascading Neural Networks [112] have been used. More recently Deep Learning [94] has also gained traction in propeller design and optimization [99].

In addition to deep learning, there is also a push towards explainable machine learning approaches [3]. This is particularly the case in critical applications such as health care where regulation requires the use of explainable models. Within the maritime domain, explainable machine learning presents very interesting opportunities to provide performance grantees thus presenting an interesting business advantage for propeller manufacturers. Furthermore, explainable machine learning can also meaningfully reduce the computational cost of optimization. Similar to RS and deep learning, they can be trained to learn the relation between objectives and constraints. Thus, they can be used to predict where lucrative designs may be found in the design space *and* also provide insight into *why* the regions are lucrative. However, explainable machine learning approaches have not yet been fully investigated or considered. The reason for this is not clearly known. Nonetheless, RS and deep learning models are trained till they are accurate and used to find the optimum (block 6). When the convergence criteria are met, the design method stops and reports the optimum. Along with the optimum, the necessary trade-off between objectives, constraints and their relation to design parameters is also often reported.

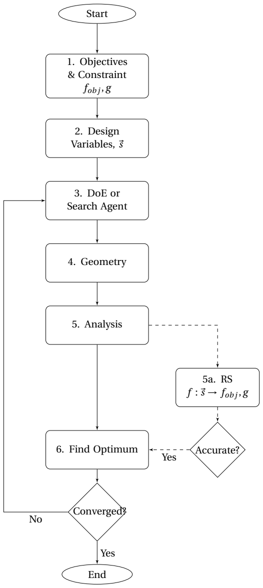


Figure 1.3: Traditional Design and Optimisation (D&O) workflow

A few design and optimization methods also account for uncertainty in operational parameters. These methods are identified as Robust Design [73]. The goal is to drive optimization towards designs with gradual performance degradation. To find such designs, Robust Design converts single point optimization problems into multi-point optimization problems where the mean and variance in operational parameters determine the range of operation over which a design is to be optimized. While uncertainties in operational parameters are accounted for, current optimization methods in the preliminary design phase are yet to account for uncertainties in performance prediction of preliminary design tools such as BEM and ETV [15, 83] used in hydroacoutic optimization problems. Preliminary design tools tend to have high uncertainties. These uncertainties result from lower-fidelity physical models, use of semi-empirical relations, numerical methods and variance in operational parameters. Consequently, comparable or similar efficiency and cavitation behaviour may be predicted for two designs even though their actual performances may be significantly different. This is a problem because it could lead optimization methods to converge to sub-optimal solutions resulting in expensive design reworks or performance surprises in the detailed design phases.

In summary, current multi-objective constrained propeller optimization problem need to address three challenges in order to improve the efficiency of optimization with a reasonable trade-off with completeness. The *first* challenge is that multicollinearity [91] resulting from classical design variables can potentially make optimization expensive. The *second* challenge is that in the preliminary design phase, BEM and ETV model may lead to high uncertainties which could lead hydroacoustic optimization to sub-optimal solutions. The *third* challenge is that even in the preliminary design phase optimization is computationally expensive. While RS and deep learning have been applied, there is an opportunity to use explainable machine learning models to reduce cost of optimization by predicting where a lucrative region in the design space exists and also also provide insight into *why* the region is lucrative.

The aforementioned challenges lead to three specific research objectives which are pursued in this body of work.

1.3. RESEARCH OBJECTIVES AND METHODS

The primary objective of this research is to improve the efficiency of optimization while achieving a reasonable trade-off with completeness and robustness. The specific objectives to address each of the three challenges - multicollinearity, uncertainty and computational cost - of state of the art optimization method are detailed below.

1.3.1. PARAMETRIZATION

OBJECTIVE

The *first* objective is to solve the problem of multicollinearity which can potentially make optimization inefficient and computationally expensive.

METHOD

The problem is solved with an orthogonal parametric model (see Section 4.2.4) which is derived from the propeller mesh. At the heart of the proposed model is SVD which is shown to accurately capture geometric variations [52]. As there is a direct correlation between the propeller mesh and predicted efficiency or cavitation performance, machine learning models can be trained quicker and better to identify lucrative regions in the design space. This quicker and better learning is expected to improve the efficiency of optimization.

1.3.2. UNCERTAINTY

OBJECTIVE

The *second* objective is to solve the problem of optimization methods converging to sub-optimal designs due to high uncertainties in prediction of efficiency and acoustic radiation in the preliminary design phase.

METHOD

To account for uncertainties in predictions of efficiency and cavitation behaviour, it is proposed that their mean and variance be modelled. The uncertainty in efficiency prediction of BEM is known and well documented [83]. However, sources of uncertainty in acoustic predictions of preliminary methods such as ETV-2 [16] are yet to be investigated.

Firstly, sources of uncertainty in the predicting acoustic behaviour of isolated tip vortices are investigated with scale resolved simulations to gain an insight into the governing physics which influences vortex dynamics. The insight gained in the investigation is used to model uncertainties in the predictions of cavitation behaviour made by preliminary design tools for propellers.

When uncertainties for both efficiency and cavitation metrics are known, soft regression and classification are applied to learn the mean and variance in performance estimations of preliminary design tools. The mean and variance predictions made by soft regression and classification models are used to identify 95 % confidence interval. Designs with a better mean performance *AND* no overlap in the 95 % confidence interval with compared designs are favoured in optimization iterations. This is expected to mitigate the risk of optimization methods converging to sub-optimal designs when uncertainty is not accounted for. However within the scope of this research, the expectation is not demonstrated.

1.3.3. EFFICIENCY

OBJECTIVE

The *third* objective is to address the problem of computational cost of optimization with a reasonable trade-off with completeness.

METHOD

Hydro-acoustic design and optimization of propellers is computationally expensive. Two methods are proposed to improve efficiency of optimization. The *first* is by the use of explainable machine learning methods to learn the location of lucrative designs in the design space. The *second* is by the use of semi-empirical models such as ETV-2 [16] to predict acoustic performance or tip-vortex cavitation behaviour.

1.4. OUTLINE

The first research objective is pursued in Chapter 2 where the application of SVD to capture geometric variation reliably is demonstrated for both a propeller and a 2D airfoil. Based on the finding that geometric variation can be accurately captured, the orthogonal parametric model is defined. The model is detailed in Section 4.2.4.

The second research objective is pursued and the method is demonstrated in Chapters 4-5. In Chapter 4, supervised and unsupervised learning strategies are synergised and applied to a single objective constrained propeller optimization problem with the goal to maximize efficiency. In the chapter, soft regression and classification are used to model mean and variance in efficiency predictions of BEM. In Chapter 3, the dynamics which influence the broad-band acoustic spectrum of tip vortices is investigated. The investigation provides insights into possible sources of uncertainty for empirical tip vortex models. In Chapter 5, mean and variance in efficiency and acoustic performance prediction is used to define the Pareto band where solutions dominate when there is no overlap in the 95 % confidence intervals of performance with compared solutions. This reduces the risk of false positives on whether a solution dominates or not. Consequently, Design and Optimisation (D&O) strategies are expected to yield a range of solutions whose predicted performance does not differ significantly from operational performance.

The third research objective is pursued in Chapter 4 and 5. Machine learning strategies are applied during optimization to identify lucrative regions in the design space with very few objective function evaluations with a reasonable trade-off with completeness.

2

GEOMETRY DEFINITION AND ANALYSIS

When optimizing propeller geometries it is very important to know sensitivities of performance w.r.t the chosen design parameters and also to make sure that the chosen parameters do not suffer from multicollinearity. If a suitable parameter set is chosen, it becomes possible to correlate geometric variation and performance. Establishing the correlation is also an important step towards facilitating machine learning approaches in optimization. This chapter introduces Model Order Reduction (MOD) as a possible solution to measure geometric variation accurately. The method to quantify variation is demonstrated on the skewed Seiun-Maru blade geometry and a 2D airfoil. For the latter, results of geometric variance study is used to selectively include a diverse set of geometries with both major and minor variations. The selective inclusion of design points results in focused evaluations thus the efficiency of multi-objective optimization improves significantly. The trade-off with completeness is also reasonable[1]*.*

2.1. INTRODUCTION

D&O of propellers can be cost effective if lucrative regions in the design space are explored efficiently. This requires insight into the sensitivities of performance w.r.t design variables in all regions of the design space. Using the sensitivities it is possible to reduce the number of design variables and/or number of samples required to explore the design space thus reducing the cost of D&O. These sensitivities are estimated via analysis (computational or experimental) which can be the most expensive steps of D&O routines. Thus, it would be a good advantage to gain insight into sensitivities with very few design point evaluations or in an a-priori manner i.e. without analysis. As the variations in performance correlate with variations in geometry of the propeller, such an a-priori insight into the magnitude of sensitivities could be gained by investigating variations in

---

[1] The research in this chapter is published in NuTTS conference 2018 [52].

propeller geometry. As a pre-requisite to such an investigation, the variations in propeller geometry need to be quantified. This chapter explores the applicability of MOD to quantify geometric sensitivities in an a-priori manner to enable the design space to be searched efficiently.

At the heart of MOD is Singular Value Decomposition where a real matrix $A_{m\times n}$ is factorized as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}V^T_{n\times n}, \tag{1}$$

where U and V are left and right singular vectors of A respectively. U is a collection of orthonormal vectors (*basis*) and can be interpreted as the principal dimensions of the data in A. Σ is a diagonal matrix which stores the singular values σ_i. They can be interpreted as the magnitude of each dimension. V can be interpreted as the rotation matrix for the data set.

The high-dimensional data in A can be approximated in a lower-dimensional subspace [60]. This dimension reduction is called Proper Orthogonal Decomposition (POD) when the leading *basis* of A are chosen [31]. SVD and POD have been successfully applied to quantify geometric variation and optimize geometries. Diez et. al [20] apply Karhunen Loéve expansion (KLE) (also referred to as POD) to reduce dimensions of the design space for D&O of a high speed catamaran hull. The design space with the largest variance is considered for optimisation. Subsequent dimensions are iteratively introduced to preserve 95% of the geometric variance. Oyama et. al [44] use POD to extract design information from shape data of Pareto-optimal transonic airfoils. The principal geometric modes are then used to classify geometries into low-drag designs, high-lift-to-drag designs, and high-lift design families. [59] perform SVD on the Jacobian matrix of a transonic airfoil's aerodynamic performance w.r.t highly dimensional geometric uncertainties. Dominant modes of geometric perturbation which influence performance are identified. These modes are then used to select a small number of inspection points on the airfoil surface to reduce uncertainty in airfoil performance by measuring manufacturing errors at the points.

In the above applications, the primary focus is to identify dominant modes or dimensions of the data. Thus, the smaller modes are ignored. This limits the applicability of dimension reduction techniques in detailed design phases where the small geometric feature variations may result in flow features such as transition or flow separation. Thus, it is important to preserve even the smaller modes. In this chapter, a sampling technique which also preserves the smaller modes is detailed. Furthermore, for MOD to be successful, the data matrix must be populated with data that accurately represents geometric variations.

In Section 2.2 the construction of the data matrix is detailed. In Sections 2.3 the Seiun-Maru propeller blade [54] and 2D airfoil test cases are detailed. Section 2.4 reports the results of the sensitivity study for the propeller and 2D airfoil. For the latter, the result of the sensitivity study is used to include designs with the largest and smallest variations allowing the design space to be explored efficiently during aeroacoutic optimization.

2.2. DATA MATRIX

To populate the data matrix, a spline surface is built from classical design parameters. For propellers examples of classical design parameters can be Diameter (D) and radial distributions of Pitch (p), Skew (s), Rake (r), Chord length (c)[7]. For 2D airfoil classical design parameters such as chord, thickness distribution and camber can be used as design parameters.

The built surface can be probed in the spline's parametric space where the coordinates $u, v \in [0.0, 1.0]$ represent the same point on the surface for all designs. In the Cartesian space, probed points $\vec{P}_i(u,v)$ and the local gradients $\nabla\vec{P}_i(u,v)$ are extracted from spline parameters. The difference between probed points and local gradients of the i-th design and baseline design are used to quantifying geometric variation. Thus, our data matrix (also referred to later as *geometry probe data matrix*) becomes $X_{N\times n} = \{\vec{X}_1, \vec{X}_2 ... \vec{X}_N\}$, where

$$\begin{aligned}\vec{X}_i = \{&\vec{P}_i(u_0,v_0) - \vec{P}_{ref}(u_0,v_0), \nabla\vec{P}_i(u_0,v_0) - \nabla\vec{P}_{ref}(u_0,v_0)...\\ &\vec{P}_i(u_{n'},v_{n'}) - \vec{P}_{ref}(u_{n'},v_{n'}), \nabla\vec{P}_i(u_{n'},v_{n'}) - \nabla\vec{P}_{ref}(u_{n'},v_{n'})\}\end{aligned} \tag{2}$$

The sum of singular values $\lambda = \sum_{i=1}^{n} \sigma_i$ can be a measure of the total variance of the data. When data corresponding to a particular design is removed from X, the new sum of singular values λ' is lower than λ. The magnitude of difference $|\lambda - \lambda'|$ indicates the reduction in total variance. $\vec{X}_i$ can also be defined as the design vector of classical variables making X the *design matrix*.

For the geometry probe data matrix, $\Delta\lambda_i$ indicates the magnitude of similarity between the i-th design (G_i) and the reference design (G_0). A low value of $\Delta\lambda$ indicates that G_i is very similar to the reference design and a high value of $\Delta\lambda$ indicates that G_i varies significantly from G_0. This information can then be used to identify design variables which significantly influence design as well as choose geometries which represent high, mid and low variance design clusters.

2.3. TEST CASES

Two cases are considered to test the viability of using SVD. The first is the Seiun-Maru propeller blade geometry and the second is a NACA4412 airfoil. The goal of this exercise is to investigate whether it is possible to classify classical design variables based on the magnitude of influence on the geometry and/or reduce the number of sample points required to explore the design space.

2.3.1. PROPELLER BLADE

The baseline geometry is initialized from coordinates of a reference propeller. Six hydrofoil cross-sections are defined and the blade is then lofted through them. Figure 2.1

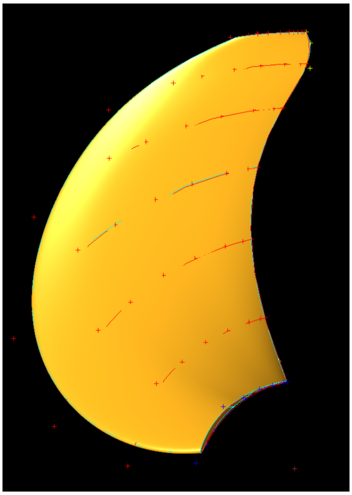


Figure 2.1: Generated Blade Geometry

illustrates the generated blade geometry. To study the geometric variation, design variables listed in Table 2.1 are varied within the specified bounds. The schematics to modify the geometry are detailed in Chapters 4 and 5. The suffix i indicates the section which is modified. The suffix sf indicates that the design variable is a scaling factor. For this case, it is important to note that variation in classical design variables are not proportional to variations in the geometry. For example, both $\Delta s_i, \Delta p_i \in [-0.056, 0.056]$ however, their influence on the geometry may be quite different. The test for the proposed method is to distinguish between the influence of these design variables.

2.3.2. NACA 4412

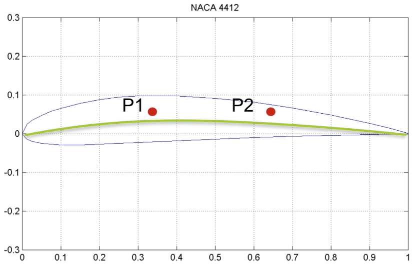


Figure 2.2: NACA4412 parametrisation

The NACA4412 airfoil's camberline is parametrized with two control points $\vec{P}_1, \vec{P}_2$ as illustrated in Figure 2.2. The design variables are the ordinates of the control points. In this case, both design variables y_{cp1}, y_{cp2} are varied within the bounds $[-4.5e-2, 1.0e-$

Table 2.1 Design Variables and Bounds

Design Variables	**Units**	**Lower Bounds**	**Upper Bounds**
ΔD	mm	-300.00	+400.00
Δs_4	rad	-0.056	+0.56
Δs_4	rad	-0.056	+0.56
$c_{sf,2}$	-	0.8	1.2
$c_{sf,3}$	-	0.8	1.2
$c_{sf,4}$	-	0.8	1.2
$c_{sf,5}$	-	0.8	1.2
Δp_1	rad	-0.056	+0.56
Δp_2	rad	-0.056	+0.56
Δp_3	rad	-0.056	+0.56
Δp_4	rad	-0.056	+0.56
Δp_5	rad	-0.056	+0.56
thk_dist$_{sf,1}$	-	-0.04	0.04
thk_dist$_{sf,2}$	-	-0.04	0.04
thik_dist$_{sf,5}$	-	-0.04	0.04

5]. The setup is such that it is not possible to reduce the number of design variables as both influence the geometry in comparable magnitudes. The goal thus is to investigate whether the number of samples required to map the design space can be reduced. To achieve this goal, trends in Coefficient of Lift (CL), ω_{SPLmax} and Sound Pressure Level (SPL)(ω) predicted from the *reduced sample set* must lead to minima and maxima which are in close proximity to that predicted by the unreduced sample set.

CL is predicted with XFOIL [23] and broadband turbulence noise is estimated with NAFNoise [40]. NAFNoise is a semi-empirical method developed to estimate the broadband frequency spectrum due to turbulent boundary layer of a 2D airfoil [41]. The underlying models correlate the thickness of the boundary layer, shear stresses within the boundary layer and freestream velocity to estimate amplitude and frequency of surface pressure fluctuations. While the absolute uncertainties in SPL prediction for broadband frequencies is high, NAFNoise claims to have low relative uncertainty in prediction for NACA0012 at moderate angles of attack with a laminar boundary layer. The authors recognize that in our case, high uncertainties are expected in the prediction of SPL(ω).

Both CL and SPL(ω) are predicted for a flow assumed to be at Mach number (M) = 0.2, $Re =$ 5.4e5 with an Angle of Attack (α) of 0^o.

2.4. RESULTS

2.4.1. PROPELLER BLADE

Figure 2.3 compares $\Delta\lambda_i/\lambda$ when data for the i-th design is removed from the design matrix and geometry probe data matrix. The histogram shows clustered bars and each

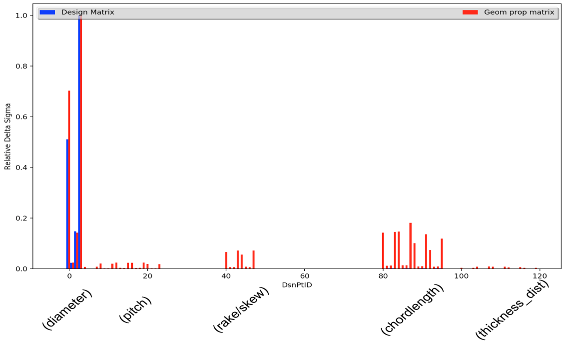


Figure 2.3: Propeller Blade $\Delta\lambda/\lambda$ vs design variables

cluster corresponds to a design variable. The leftmost bar in the cluster corresponds to the design when $x_i = lb_i$ and the right most bar in the cluster corresponds to $x_i = ub_i$.

SVD of the *design matrix* indicates that D has the highest influence on the geometry. This is because the magnitude of ΔD is highest among design variables. Given that the bounds for Δs_i and Δp_i are the same, corresponding $\Delta\lambda/\lambda$ are also the same. Thus, no inference can be drawn about the influence of each design variable on the geometry. However, SVD of the *geometry probe data matrix* makes a distinction between the influence of the two design variables. Furthermore, it also gives insight into sensitivities of the geometry w.r.t design variables allowing them to be classified. This classification can be beneficial in multi-stage D&O routines where variables which influence geometry significantly can be investigated in the preliminary design stages and finer modifications in the detailed design phases.

2.4.2. NACA 4412

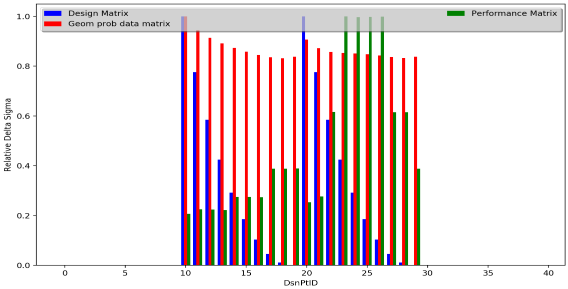
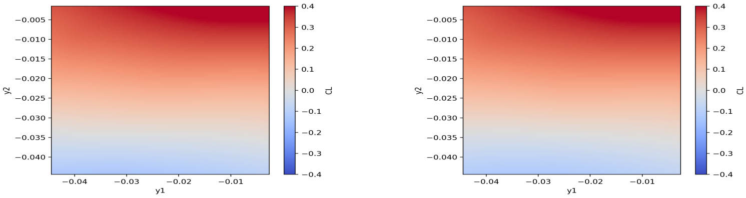


Figure 2.4: NACA 4412 $\Delta\lambda/\lambda$ vs design variables

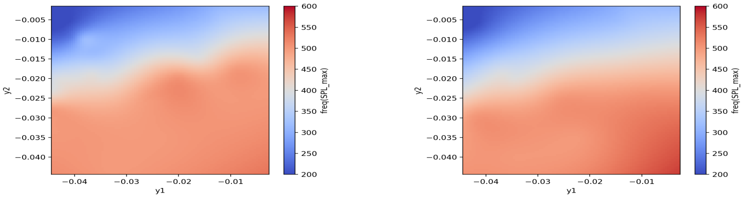
Figure 2.4 compares $\Delta\lambda_i/\lambda$ when data for the i-th design is removed from the design matrix and geometry probe data matrix. The performance matrix $X_{N\times3}$ has CL in the first column, ω_{SPLmax} in the second column and the corresponding SPL(ω_{maxSPL}) in the third column. The histogram shows clustered bars and each cluster corresponds to a design variable. As in the case of the propeller, the design variables are systematically changed one at a time. The leftmost bar in the cluster corresponds to the design vector when $x_i = lb_i$ and the right most bar in the cluster corresponds to $x_i = ub_i$.



(a) RS for CL
generated from all samples

(a) RS for CL
generated from 50% of original samples

Figure 2.5: CL Response Surfaces



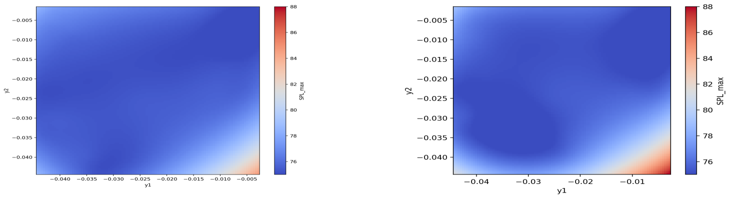
(a) RS for ω_{SPLmax}
generated from all samples

(a) RS for ω_{SPLmax}
generated from 50% of original samples

Figure 2.6: ω_{maxSPL} Response Surfaces

From Figure 2.4 it can be observed that the geometry is comparably sensitive to both the design variables. Thus, it is not possible to reduce the number of design variables in the current stage of design. While design variables cannot be classified, the designs can be clustered to significantly reduce the number of designs required to map the design space.

To make a-priori design selections for analysis, $\Delta\lambda_i$ from SVD analysis of the geometry probe data matrix is used to classify designs into high, mid and low variance categories. From each category, half the samples are selected. For the selected subset the performance predictions are made with XFOIL and NAFNoise. A Radial Basis Function (RBF) RS is built to correlate the design vector with performance predictions. Figure 2.5 - 2.7 illustrate the response surfaces built from the original design matrix ($N = 38$) and



(a) RS for SPL(ω_{SPLmax}) generated from all samples

(a) RS for SPL(ω_{SPLmax}) generated from 50% of original samples

Figure 2.7: SPL(ω_{SPLmax})Response Surfaces

Table 2.2 Minima

	RS1	**RS2**
Num Training Samples	38	19
CL_{min}	-0.14	-0.12
$\vec{X}_{sol,CL}$	[-0.033, -0.045]	[-0.032, -0.045]
$min(\omega_{SPLmax})$	179.95	174.36
$\vec{X}_{sol,min(\omega_{SPLmax})}$	[-4.50e-02, 1.00e-05]	[-4.50e-02, 1.00e-05]
$min(\text{SPL}_{max})$	74.48	74.29
$\vec{X}_{sol,SPL}$	[-0.0058, -0.0067]	[-0.033, -0.033]

the reduced design matrix ($N = 19$) for CL, ω_{SPLmax} and $SPL(\omega_{SPLmax})$. It is observed that there is reasonable agreement between the RS built with the unreduced sample set (RS1) and the reduced sample set (RS2). The RSs are searched with a constrained Sequential Least SQuares Programming (SLSQP) algorithm to find the minima. The results are listed in Table 2.2. It is observed that for CL and ω_{SPLmax}, the minima predicted by RS1 and RS2 are close. However, for $SPL(\omega_{SPLmax})$ the minima are quite far apart. This is because, there are two possible minima as shown in Figure 2.7 and one of them was not included in the reduced sample set. This issue can be addressed with adaptive sampling where ever a lucrative trough is located on the RS.

2.5. CONCLUSION

SVD seems to be a promising technique to (a) reduce the number of samples in Design of Experiments (DoE) and, (b) classify design variables based on their magnitude of influence on the geometry. For the method to be applicable, data populating the factorized matrix must represent geometric variations reliably. Thus, data such as probed points and local gradients are more useful to quantify geometric variation as opposed to design vectors with classical design variables. This is because, the magnitude of change in geometry is not always proportional to the magnitude of change in classical design variables all the time.

In the propeller blade case, SVD of the geometry probe data matrix enabled the classification of variables based on the magnitude of influence on the geometry. To gain insight into the sensitivities of performance w.r.t design variables, analysis of specific design points is required. The jacobians of performance data w.r.t design variables may be appended to the data matrix for analysis. The hypothesis is that the method may require fewer design point evaluations compared to classical methods to provide necessary insight into global sensitivities of performance parameters.

In the case of the airfoil, it was possible to identify from SVD of the geom probe data matrix that classification of variables is not viable. Furthermore, it was also possible to reduce the number of sample points by 50% while still capturing the performance trends globally in the design space. However, it is noted that the reducing the initial sample set may lead to exclusion of one or more local minima. This can be addressed by adaptively refining the RS in regions where a trough is observed and a local minima is expected.

Also, for the 2D airfoil case, the use of lower-fidelity methods leads to monotonic or convex trends in performance. But the use of higher fidelity methods could lead to more complex performance trends. Under such circumstances, SVD's applicability is yet to be investigated. This concern mainly arises from studies which drop all smaller modes/dimensions of the data. However, these smaller modes may be necessary to preserve dynamics within the data. Our methods preserve some of the smaller dimensions/modes, but they need to be demonstrated for cases with complex performance trends.

2.6. RECOMMENDATIONS

- In the current study, classical design variables - pitch, diameter, hydrofoil camber, thickness distribution, skew and rake are used to parametrize the geometries. Thus, it is recommended that the sensitivity studies also be repeated for other parametric models such as coefficients of polynomials that define propeller families and control points of splines and free-form deformation boxes.

- Model Order Reduction: For dimension reduction, it is recommended that Gappy POD be further investigated. Gappy PoD may preserver the smaller singular values which contain information regarding smaller geometric variations resulting from modifications to thickness distribution. Consequently, better relations between modified thickness distribution and cavitation behaviour can be established.

ACKNOWLEDGMENT

The authors would like to whole heartedly thank Evert-Jan Foeth for thoughts and comments on the research. We also thank Erik Rotteveel, Thomas Scholcz and Bart Schuiling at MARIN for interesting exchange of ideas on design and optimisation.

3

TIP VORTEX CAVITATION NOISE

In this research, the authors simulate a cavitating vortex without forced excitation. In the simulation, (a) the influence of Kelvin-Helmholtz instabilities on the vortex cavity and, (b) the relevance of compressibility is investigated. Due to the presence of the water-vapour interface, KH instabilities develop and excite the vortex cavity in an axis asymmetric manner. This results in a broadband acoustic spectrum. In the free-stream the acoustic waves result in relatively small density perturbations however, within the core of the vortex the gradient of density is significantly higher when compared to the free-stream. This suggests that water is largely incompressible and in contrast the vapour core is compressible. While dynamics in the free-stream can influence the flow within the core the momentum within the vapour core is expected to be too low to influence free-stream dynamics [1].

3.1. INTRODUCTION

The maritime industry has embraced its responsibility towards the environment and aims to significantly reduce emissions of GHG and Underwater Radiated Noise (URN). Modern propellers are typically optimised to bounds of performance thus, they are more likely to operate with vortex cavitation as the dominant source of radiated pressure fluctuations and broadband noise [8].

The acoustic spectrum of the propeller tip vortex cavity results from excitation by (a) the wake-field [108], (b) turbulence, (c) flow separation on the lifting surface [37] and, (d) sheet cavity feeding into the tip-vortex [15]. Furthermore, for a cavitating vortex, the presence of an interface is likely to result in Kelvin-Helmholtz instabilities [12, 28]. The relation between these instabilities and resulting broadband spectrum is investigated in this study.

---

[1] The research in this chapter will be submitted as a paper to International Journal of Fluid Mechanics

Chandrasekhar [17] and Drazin and Reid [22] studied the stability of perturbed vortex cavities. For stable vortex filaments which are excited, three primary modes of oscillation illustrated in Figure 3.1 are expected. Axisymmetric disturbances result in the $n = 0$ breathing mode. The bending mode $n = 1$ occurs when the axis of the vortex is deformed and velocity perturbations remain at $r = 0$. The helical mode $n = 2$ occurs when the vortex is deformed into an ellipse that rotates about the axis [49]. The first mode is expected to behave as a monopole, the second as a dipole and the third as a quadrupole [45].

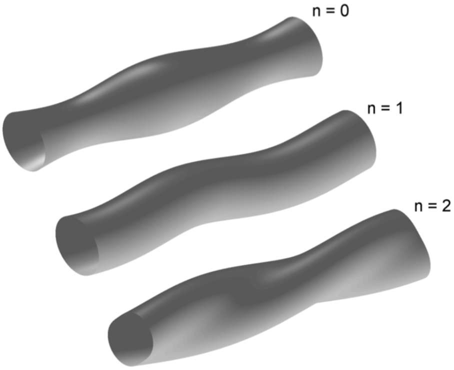


Figure 3.1: The primary vortex cavity oscillation modes. (Figure reproduced from [13])

To date, simulations of vortex cavitation [10, 37, 51] have been performed with an incompressibility assumption and on relatively coarse domains. In this research an isolated vortex cavity is simulated. It is initialised in a finer domain without explicit external perturbations to observe the resulting response of vortex cavity. Water is modelled as a compressible barotropic fluid. The governing equations and thermodynamic model are detailed in Sections 3.2.1-3.2.2. The turbulence model and the numerical scheme is discussed in Section 3.2.3. The geometry, boundary conditions and mesh for the computational domain is detailed in Section 3.3.

3.2. PHYSICAL AND NUMERICAL MODEL

3.2.1. GOVERNING EQUATIONS

The governing equations of flow are barotropic three-dimensional Navier-Stokes equations:

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho u_i)}{\partial x_i} = 0, \tag{1}$$

$$\frac{\partial (\rho u_i)}{\partial t} + \frac{\partial (\rho u_i u_j)}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{\partial \tau_{ij}}{\partial x_j}, \tag{2}$$

where t is time, u_i is the component of velocity, ρ is density, p is pressure and τ is the viscous stress tensor. The thermodynamic model correlates pressure and density to complete the system of equations.

3.2.2. THERMODYNAMIC MODEL

In the cavitating flow, phase change is modelled by assuming thermodynamic and thermodynamic equilibrium [34]. For water-vapour mixtures the pressure and vapour volume fraction are functions of mean density:

$$\bar{\rho} = \alpha \rho_{vap} + (1-\alpha)\rho_{liq}, \tag{3}$$

where the vapour volume fraction is

$$\alpha = \frac{V_{vap}}{V} \tag{4}$$

It is assumed that phase change is instantaneous, isentropic and in mechanical equilibrium. Thus, the vapour fraction can be computed as:

$$\alpha = \begin{cases} 0 & \bar{\rho} \geq \rho_{sat,liq} \\ \frac{\rho_{sat,liq}-\rho_{liq}}{\rho_{sat,liq}-\rho_{sat,vap}} & \bar{\rho} < \rho_{sat,liq} \end{cases} \tag{5}$$

Liquid water ($\alpha = 0$), is modelled by the Tait equation of state [34]:

$$\bar{p} = (p_{sat} + B).\left(\frac{\bar{\rho}}{\rho_{sat,liq}}\right)^N - B, \tag{6}$$

where N = 7.1 and B = 3.06e8 Pa. When $1 > \alpha > 0$, the equilibrium pressure is:

$$\bar{p} = p_{sat} + C\left(\frac{1}{\rho_{sat,liq}} - \frac{1}{\bar{\rho}}\right). \tag{7}$$

The parameter C is:

$$C = L^2 \frac{\rho^2_{sat,vap}}{(Cp_{liq}T)}, \tag{8}$$

where L is the latent heat of vaporisation and Cp is the specific heat of water at constant pressure and T is the temperature. The thermodynamic properties comply with International Association for the Properties of Water and Steam (IAWPS) Formulation [58].

Based on Beattie et. al [9] volume averaged viscosity is defined as:

$$\bar{\mu} = (1-\alpha)(1+\frac{5}{2}\alpha)\mu_{liq} + \alpha\mu_{vap}. \tag{9}$$

The above model is derived by merging the viscosity definitions

$$\bar{\mu} = (1+\frac{5}{2}\alpha)\mu_{liq} \qquad \text{(bubble flow)} \tag{10a}$$

$$\bar{\mu} = (1-\beta)\mu_{liq} + \beta\mu_{vap} \qquad \text{(annular flow)} \tag{10b}$$

The bubble and annular flow definitions are merged because Beattie et. al [9] find that at higher void fractions the coefficient 2.5 is too low and the annular flow definition is based on the concept of interfacial waves switching μ_{liq} and μ_{vap} intermittently. As both higher void fractions and interfacial waves are expected to play a role in the simulation, the hybrid viscosity definition is adopted.

3.2.3. LARGE EDDY SIMULATIONS

For Large Eddy Simulations a Sub-Grid Scale (SGS) closure model is required. For this an Implicit LES (ILES) turbulence model [32, 33] based on the Adaptive Local Deconvolution Method (ALDM) is used. ILES is an approach where the nonlinear truncation error resulting from discretisation of the convective terms is used as the SGS model [5]. ALDM is a non-linear discretisation scheme designed for ILES where the discretisation is based on a solution-adaptive deconvolution operator that allows control over the truncation error. The parameters of deconvolution were tuned to obtain an optimum spectral match for numerical viscosity. This approach has also been demonstrated to work well in simulating complex multi-phase fully compressible flows [24, 39, 43]. The ALDM reconstruction is applied to velocity and pressure. The density and free-gas mass fraction are reconstructed by first-order upwind biased method. The viscous flux is discretized by a second-order centred scheme. Time integration is performed with an explicit, third-order Runge-Kutta method.

3.3. COMPUTATIONAL DOMAIN

The computational domain is illustrated in Figure 3.2. The domain has a length of 0.02 m with a square cross-section of width 3.2 m. The principal axis of the domain is parallel to positive x-axis. The longitudinal faces are treated as periodic boundaries. The lateral faces are modelled as free-slip walls.

The cross section width is chosen to minimise the influence of a square cross section on an annular flow. The domain length of 0.02 m is chosen to specify the wavelength of resulting modes with the periodic boundary condition.

3.3.1. DOMAIN INITIALISATION

For the flow, $Re = 9e5$ and $\sigma \in \{1.2, 1.7, 2.6\}$, the system of Equations (11)-(13) are solved for free-stream density (ρ_∞), free-stream velocity (u_∞) and free-stream temperature (T_∞). The solutions are listed in Table 3.1. As the computational domain has periodic boundary conditions, u_∞=6.0193ms^{-1} is maintained for all three cases. Turbulence intensity of the flow is also initially set to zero. Thus, in the situation that the imposed boundary condition itself becomes a numerical source of excitation, a strong response would be expected at $\nu = \frac{\nu_\infty}{0.02} \approx 301 Hz$.

$$Re = \frac{\rho_{sat,liq}(T_\infty) \times u_\infty \times d}{\mu(T_\infty)} \tag{11}$$

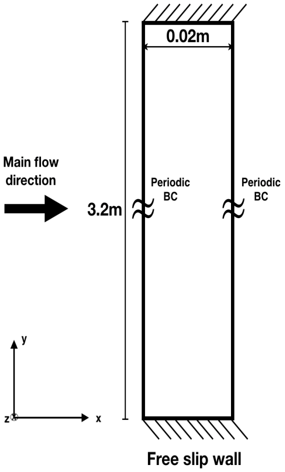


Figure 3.2: The computational domain to simulate the cavitation vortex.

Table 3.1 Free-stream temperature, velocity and density for the three cavitation numbers

Parameters(Units)	Cavitation numbers		
	1.2	1.7	2.6
T_∞ (K)	2.9830e2	3.009e2	3.007e2
u_∞ (ms^{-1})	6.3759	6.0193	6.0491
ρ_∞ (kgm^{-3})	9.9698e2	9.9626e2	9.9635e2

$$\sigma = \frac{p_\infty - p_{sat}(T_\infty)}{\frac{1}{2}\rho_{sat,liq}(T_\infty) \times u_\infty^2} \tag{12}$$

$$p_{rc} = p_{sat}(T_\infty) = -\int_{rc}^{\infty} \frac{\rho_{sat,liq}(T_\infty) \times u_\theta^2}{r} + p_\infty \tag{13}$$

The azimuthal velocity (u_θ) is initialised with the Proctor-Winkelmans kinematic model. Pennings et. al [46] study the flow field around cavitating vortices generated by an elliptical hydrofoil at different angles of attack and cavitation numbers. The goal of the study was to establish a correlation between the vapour cavity size and the circulation strength of the vortex. Towards this end, the kinematics of a wetted and cavitating vortex were studied. It was observed that the kinematics were best captured by the modified 2D Proctor-Winkelmans vortex model as it performed well even in the vortex roll-up region. When $r < 1.15r_{vc}$, the u_θ is specified by the Winckelmans model described in Equation (14) and when $r >= 1.15r_{vc}$, u_θ is specified by the modified Proctor model described in Equation (15). In Equation (14), B is the span of the lifting surface which generates the vortex and Vortex Circulation (Γ). β_0 accounts for the vortex roll-up and inclusion of the wing boundary layer wake in the outer part of the vortex flow. β_i sets

Table 3.2 Proctor-Winckelmans vortex model parameters for cavitating vortex cases

Parameter(Units)	Cavitation numbers		
	1.2	1.7	2.6
$r_{vc}(mm)$	2.8152	2.1627	1.7401
$r_c(mm)$	2.4042	1.8980	1.0995
$\Gamma(rad.m^2.s^{-1})$	0.2159	0.2163	0.2186
β_i	5.6195e3	7.6927e3	1.0470e4
β_0	1.6588e1	1.6375e1	1.6748e1
β	1.6590e1	1.6374e1	1.6731e1
B (m)	0.3	0.3	0.3
p	1.6200e2	3.6843e1	2.1417e1

the relation for viscous core as $r_v/B \approx (\beta_0/\beta_i)^{4/5}$. p is tuned to match the peak u_θ. In Equation (15), β is optimised to ensure u_θ is continuous.

$$u_\theta = \frac{\Gamma}{2\pi r}\left(1 - exp\left(\frac{-\beta_i\left(\frac{r}{B}\right)^2}{\left(1+\left(\frac{\beta_i}{\beta_0}\left(\frac{r}{B}\right)^{\frac{5}{4}}\right)^p\right)^{\frac{1}{p}}}\right)\right) \tag{14}$$

$$u_\theta = \frac{\Gamma}{2\pi r}\left(1 - e^{-\beta\left(\frac{r}{B}\right)^{0.75}}\right) \tag{15}$$

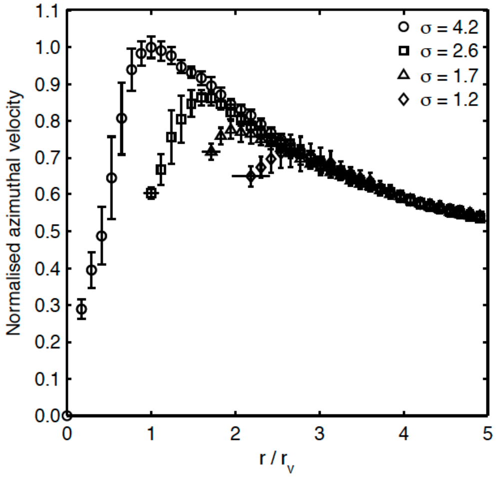


Figure 3.3: *(Reproduced from [45])*. Uncertainty in experimental measurements of azimuthal velocity for an elliptical hydrofoil at $C_L = 0.66$ and $Re = 9e5$ for various cavitation numbers. $u_{\theta,max} = 6.7ms^{-1}$ and $r_{vc} = 1.1mm$.

Figure 3.3 illustrates the azimuthal velocity and corresponding uncertainties in Tomographic Particle Image Velocimetry (PIV) measurements used to train the Proctor-Winckelmans kinematic model. It is observed that the uncertainty is relatively higher near the interface compared to the viscous roll-up region. Furthermore, as these measurements are made with Tomographic PIV, no measurements could be made within the core of the vortex. For the available data, Table 3.2 lists parameters of the trained kinematic model for the three cavitating vortex cases.

From the Proctor-Winckelmans kinematic model, the pressure and density field are initialised. The pressure at radius r from the principal axis for the 2D vortex is,

$$p(r) = p_\infty - \int_r^\infty \frac{\rho(p) u_\theta^2(r, \Gamma, \beta_i, \beta_0, B, \mathrm{p})}{r} \partial r. \tag{16}$$

As pressure and density are correlated, a simple pressure density iteration algorithm solves for pressure. The radial pressure at the j+1 iteration is,

$$p_{j+1}(r) = p_\infty - \int_r^\infty \frac{\rho(p_j) u_\theta^2(r, \Gamma, \beta_i, \beta_0, B, \mathrm{p})}{r} \partial r, \tag{17}$$

and the density based convergence criteria is,

$$\mathrm{abs}(\rho_{j+1} - \rho_j) < 10^{-14}. \tag{18}$$

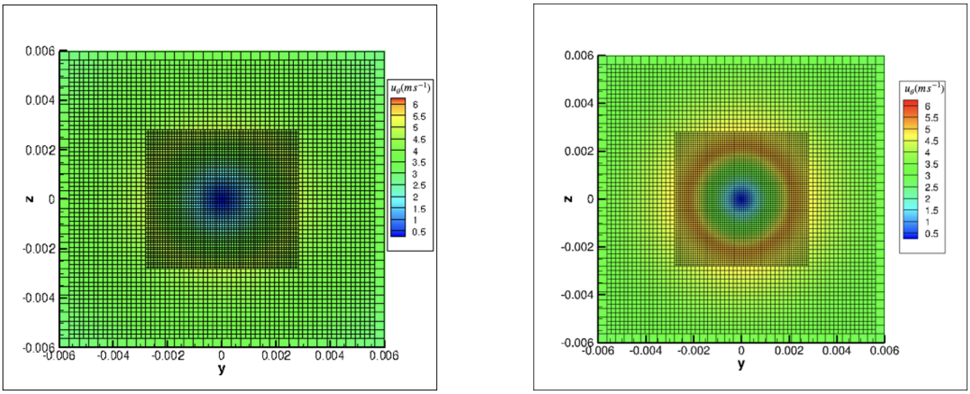
The domain is thus initialised with $u_\theta(r), \rho(p_{j+1})$.

3.3.2. ADAPTIVE MESH REFINEMENT

Cartesian grids for the conservative finite volume discretisation of the integral form of the Navier-Stokes equations is used. The Cartesian grid is iteratively refined to resolve the cavitating vortex. For a non-cavitating flow there are several vortex identification methods [25] available and the λ_2 criteria is employed in conjunction with spatial bounds to restrict mesh refinement to the free-stream region away from walls. Within the spatial bounds, the grid is refined in regions where λ_2 is below a specified threshold. Since iterative mesh refinement locally increases vorticity, λ_2 favourably reduces over each iteration respecting the threshold. This localised refinement tends to over-resolve the vortex core, thus, the grid is also uniformly refined in the viscous core region of the vortex. Most Eulerian local vortex identification methods fail when the resolved vortex begins to cavitate as they are derived for incompressible flows without a vortex jump. Thus, with the onset of cavitation λ_2 criteria is switched to a density threshold criteria.

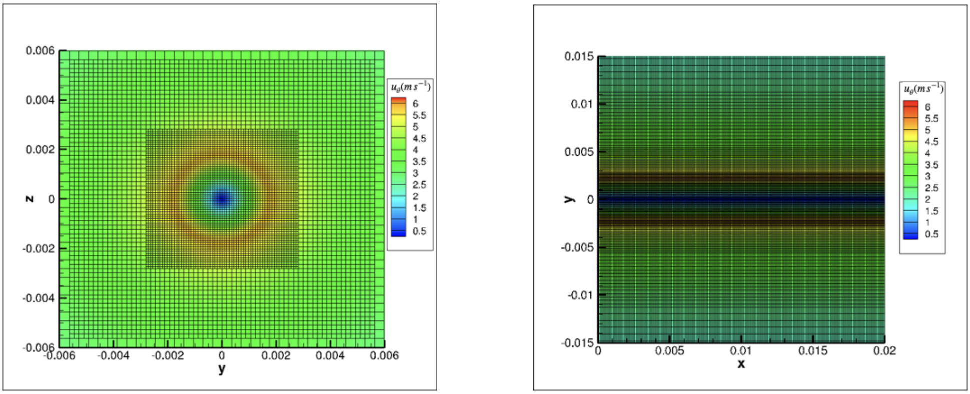
3.3.3. MESH

Figures 3.4a-3.4c illustrates the mesh in the X=0 plane and Figure 3.4d illustrates the mesh in the Z=0 plane for the three cases. The viscous core of the vortex is resolved with $\delta y = \delta z = 0.8789 \mu m$ and $\delta x = 0.7812 \mu m$. The vortex roll-up region is resolved with mesh coarsened by a factor of 2 w.r.t the mesh in the viscous core region.



(a) $\sigma = 1.2$

(b) $\sigma = 1.7$



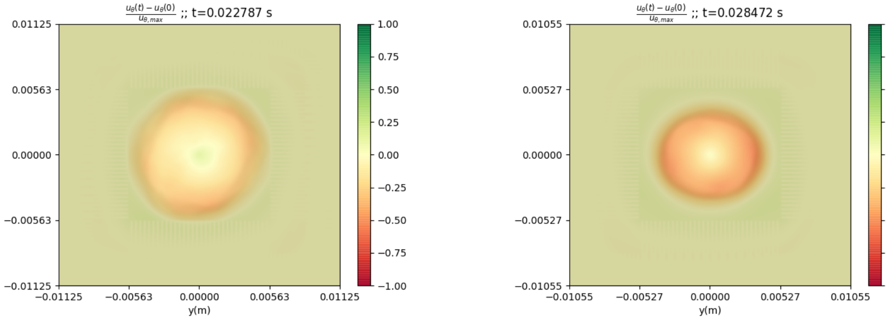
(c) $\sigma = 2.6$

(d) Longitudinal cross-section of the mesh

Figure 3.4: The mesh generated for the simulation domain visualised at X=0 Plane.

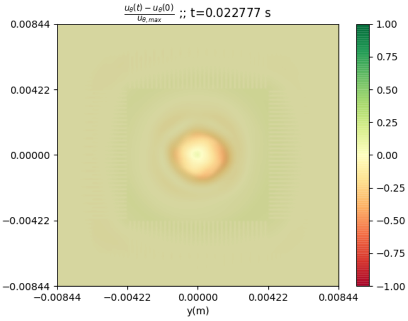
3.4. RESULTS

The following sections present the results of the simulation. Subsections 3.4.1-3.4.4 detail the observed dynamics in the flow and 3.4.5 details the resulting response modes and acoustics of the simulated cavitating vortices.



(a) $\sigma = 1.2$

(b) $\sigma = 1.7$



(c) $\sigma = 2.6$

Figure 3.5: The deviation of normalised tangential velocity u_θ of cavitating vortex cases for simulations initialised with Proctor-Winkelmans kinematic model. $u_{\theta,max} = 6.7ms^{-1}$. The plots are overlaid with $\alpha(t)$ filtered along the flow axis.

3.4.1. KINEMATICS

The simulation provides insight into the kinematics at the cavitating core. Figure 3.5 illustrates the difference in tangential velocity from the initialised Proctor-Winkelmans kinematic model for all three cases. The difference is computed from the tangential velocity filtered along the flow axis at $t_0 = 0.0000ms$. The velocity profile deviates significantly from the Proctor-Winkelmans kinematic model within the vapour cavity. Towards the centre of the cavity and up to the interface the kinematic model under-predicts u_θ. Near the viscous core the model over-predicts u_θ. The magnitude of difference reduces beyond the viscous core where the grid is also relatively coarser.

The observed deviations are expected to result from (a) the Proctor-Winkelmans model not being trained with measurements in the vortex core, (b) relatively high experimental uncertainties in azimuthal velocity measurements near the interface *(see Fig. 3.3)*, (c) 3D dynamic deformation modes which are not factored in the 2D kinematic model and, (d) uncertainties in simulation predictions. The observed deviation from the kinematic model indicates that it is important to account for uncertainties when making sound pressure level predictions with Empirical Tip Vortex [16] models where the size of the vortex cavity plays an important role in predicting mean frequencies.

3.4.2. KELVIN-HELMHOLTZ INSTABILITY

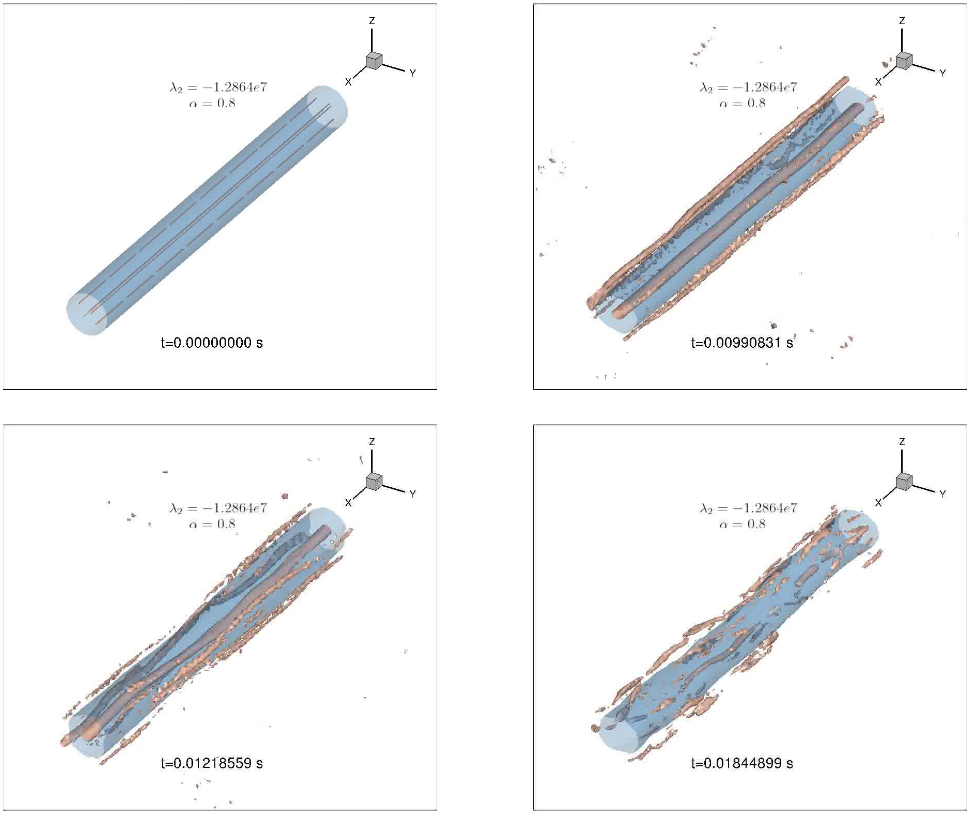
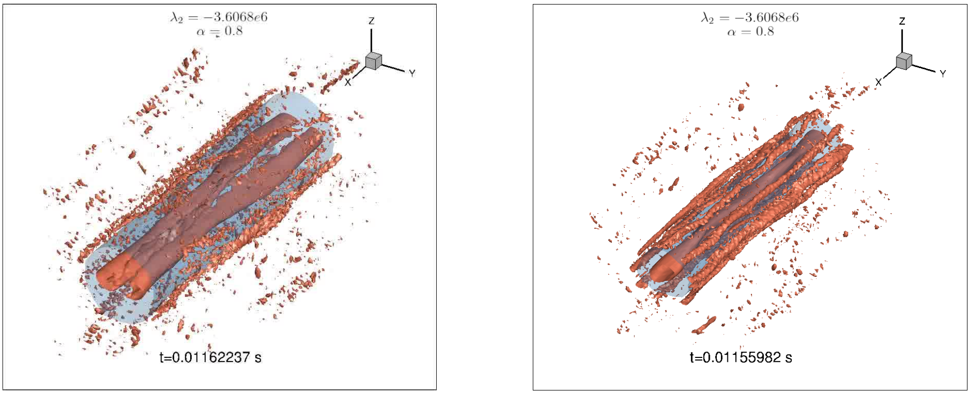


Figure 3.6: Iso-contours of λ_2 for $\sigma = 2.6$ at different time-steps showing developing bending modes. The liquid-vapour interface is illustrated as an iso-contour of $\alpha = 0.8$.



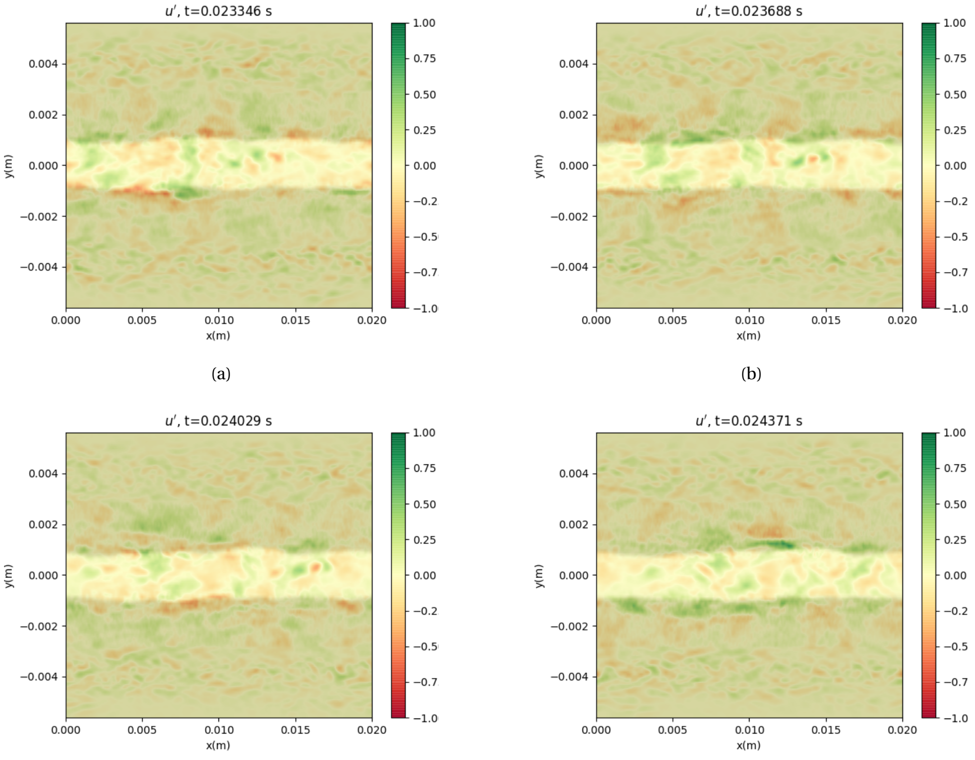
(a) $\sigma = 1.2$

(b) $\sigma = 1.7$

Figure 3.7: Iso-contours of λ_2 for $\sigma \in \{1.2, 2.6\}$ showing a dominant helical deformation mode. The liquid-vapour interface is illustrated as an iso-contour of $\alpha = 0.8$.

The presence of tangential velocity inflection near the interface results in Kelvin-Helmholtz instabilities which grow and influence the dynamics of the cavitating vortex in all three cases. Figure 3.6 illustrates the vortex structures due to KH instabilities surrounding the cavity for $\sigma = 2.6$ and the resulting dynamic response of the cavity. Figure 3.7 illustrates the dynamic response of the vortex cavity for $\sigma \in \{1.2, 1.7\}$. It is observed that the responses varies with cavity size. For $\sigma = 2.6$ the bending mode is visually discerned and for $\sigma \in \{1.2, 1.7\}$ the helical mode is visually discerned. The dynamic modes in addition to the above are detailed in Subsection 3.4.5.

3.4.3. TURBULENCE



(a)

(b)

(c)

(d)

Figure 3.8: Plots of turbulent velocity fluctuations (u') for different instances in time for case $\sigma = 2.6$. The plots are overlaid with $\alpha(t)$.

The turbulent velocity fluctuations

$$u'(t_n) = \frac{1}{t_{n+3} - t_n} \int_{t_n}^{t_n+3} u(t)\,dt - u(t_n), \tag{19}$$

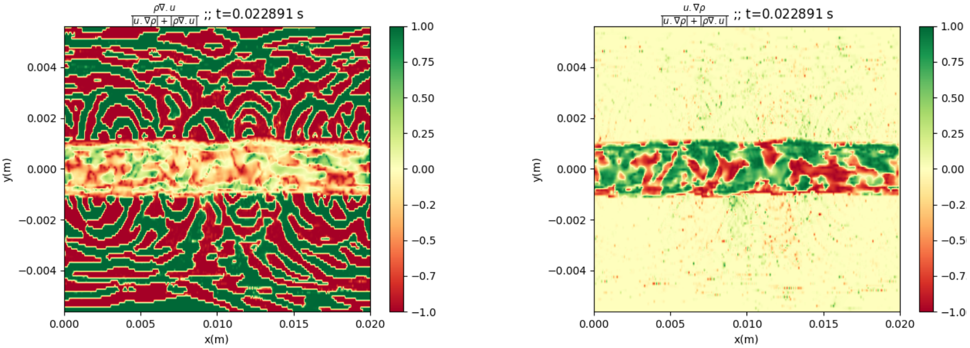
are shown in Figure 3.8. Since the initial condition is a laminar axisymmetric vortex flow without perturbations, the turbulent fluctuation are expected to develop in the flow. The turbulence in the viscous core and rollup region is expected to be a result of the growing Kelvin-Helmholtz instabilities. Turbulence also develops within the cavitating core of the vortex. The vapour cavity is relatively more turbulent than the free-stream liquid flow. However, the energy and momentum contained in these fluctuations is very small compared to that in the free-stream as water has a significantly higher density than vapour.

3.4.4. COMPRESSIBILITY

From the continuity equation, the rate of change of density is

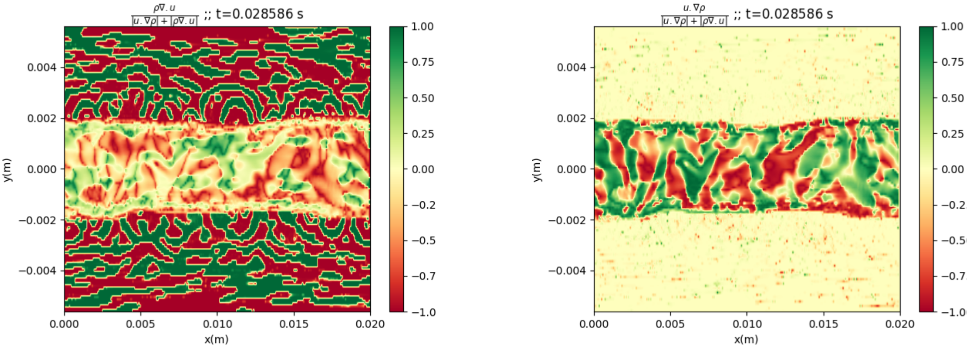
$$\frac{\partial \rho}{\partial t} = -\{u \cdot \nabla \rho + \rho \nabla \cdot u\}. \tag{20}$$

When water is treated as an incompressible fluid, $\nabla \rho = 0$ and $\nabla . u$ becomes a source term for the cavitation model. When water is treated as a compressible fluid, both the gradient of density and divergence of velocity are expected to be non-zero thus, they contribute to the rate of change of density.



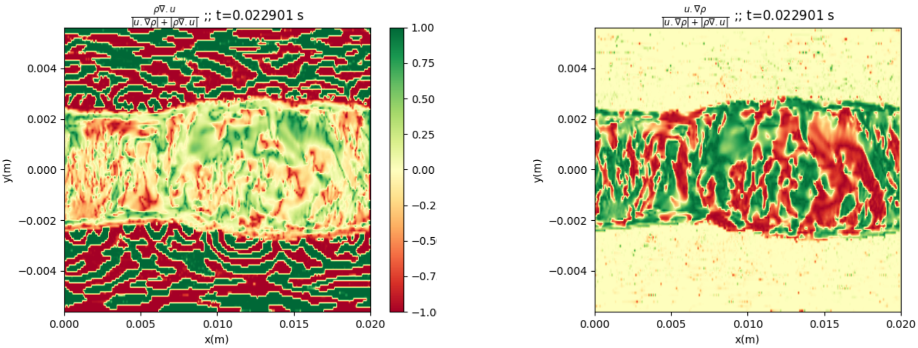
(a) $\frac{\rho\nabla.u}{|u.\nabla\rho|+|\rho\nabla.u|}$ (b) $\frac{u.\nabla\rho}{|u.\nabla\rho|+|\rho\nabla.u|}$

Figure 3.9: Relative contributions of density gradient and velocity divergence to rate of change of density for $\sigma = 2.6$



(a) $\frac{\rho\nabla.u}{|u.\nabla\rho|+|\rho\nabla.u|}$ (b) $\frac{u.\nabla\rho}{|u.\nabla\rho|+|\rho\nabla.u|}$

Figure 3.10: Relative contributions of density gradient and velocity divergence to rate of change of density for $\sigma = 1.7$



(a) $\frac{\rho\nabla.u}{|u.\nabla\rho|+|\rho\nabla.u|}$ (b) $\frac{u.\nabla\rho}{|u.\nabla\rho|+|\rho\nabla.u|}$

Figure 3.11: Relative contributions of density gradient and velocity divergence to rate of change of density for $\sigma = 1.2$

Figure 3.9 shows the relative contributions of density gradient and velocity divergence towards rate of change of density for $\sigma = 2.6$. Figure 3.9a illustrates the relative importance of velocity divergence. Acoustic waves in the free stream appear as arcs of positive and negative divergence bands. They are predominantly responsible for change in density of water as observed in Figure 3.9b. From Figure 3.9b it can be observed that density gradient is significantly higher in the vortex cavity compared to the free-stream. Similar observations are made on the relevance of compressibility for larger vortex cavities at $\sigma \in \{1.2, 1.7\}$ as seen in Figures 3.10-3.11.

It can thus be concluded that KH instabilities excite the vortex resulting in a broadband acoustic spectrum. These acoustic waves propagate into the free-stream and also into the vortex cavity. It is observed that in the free-stream, velocity divergence - primarily related to acoustics — leads to a change in density of water. Since the gradient of density is very low in the free-stream, water tends to be largely incompressible. Incompressibility rapidly decreases however when vapour is mixed in water. In the vapour core, the gradient of density ($\nabla\rho$) is much higher compared to the free-stream and the divergence of velocity is also a source for change in density. Thus, the vapour core behaves as a compressible gas.

While the core tends to behave as a compressible gas, the density of water vapour is several orders of magnitude lower than that of water. Thus, while the free-stream may influence the dynamics within the cavity, the momentum within the core is most likely too low to influence interface dynamics and thereby the free-stream dynamics.

3.4.5. ACOUSTICS AND RESPONSE MODES

The acoustics of the cavitating vortices are measured with probes in the free-stream situated close to the viscous core. The spectrum of the measured signal is extracted with Fast Fourier Transforms (FFT). The corresponding response modes of the cavity are extracted with Dynamic Mode Decomposition (DMD) [48, 50].

DMD of the scalar vapour fraction field in 3D is performed. For m snapshots, the snapshot matrix (X)

$$X = \begin{bmatrix} | & & | \\ \alpha_{ijk}(0) & \cdots & \alpha_{ijk}(t_{m-1}) \\ | & & | \end{bmatrix}, \tag{21}$$

and time-shifted snapshot matrix (X')

$$X' = \begin{bmatrix} | & & | \\ \alpha_{ijk}(1) & \cdots & \alpha_{ijk}(t_m) \\ | & & | \end{bmatrix}, \tag{22}$$

are assembled. A linear mapping A connects the i^{th} flow filed x_i to the subsequent flow field x_{i+1}. Thus, the mapping is defined as

$$A = X'X^{-1}, \tag{23}$$

where X^{-1} is the Moore-Penrose pseudo-inverse [53]. With thin SVD the inverse can be efficiently and accurately computed as

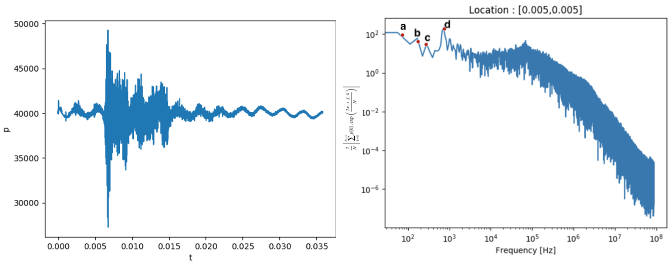
$$X^{-1} \approx U\Sigma V^T. \tag{24}$$

Thus, the system is approximated as

$$\tilde{A} = X\ V\Sigma^{-1}U^T \tag{25}$$

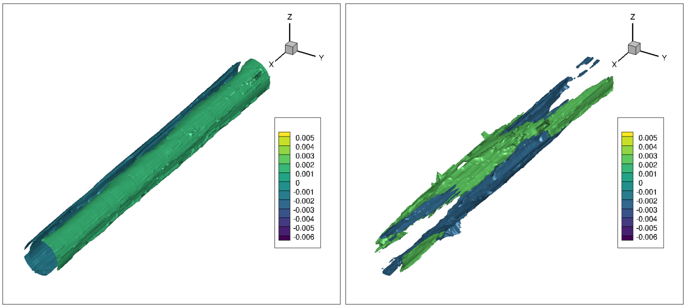
The Eigen decomposition of $\tilde{A}.W = W\Lambda$ yields the different Eigen modes and their frequencies $\nu_i = \frac{angle(w_i)}{2\pi\Delta t}$.

The Fourier analysis of the pressure spectrum, the corresponding Eigen modes and their frequencies for two cavitating vortices are detailed below.

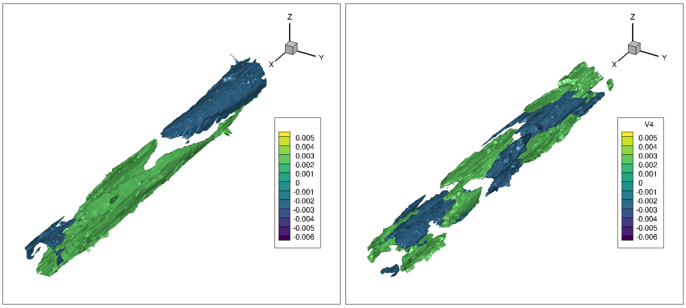


(a) Pressure signal (b) FFT of the pressure signal.

Figure 3.12: Probe signal and FFT at Location [5 mm, 5 mm, 0 mm] for $\sigma = 2.6$. The response modes corresponding to tonal frequencies (a-d) in Figure 3.12b are illustrated in Figure 3.13.



(a) $\nu = 69.2519$ Hz (b) $\nu = 164.9226$ Hz

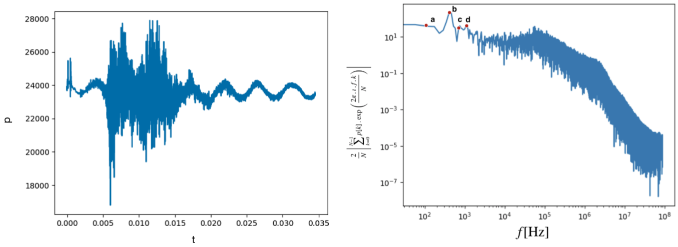


(c) $\nu = 256.6732$ Hz (d) $\nu = 703.8970$ Hz

Figure 3.13: Response modes of vortex cavity for $\sigma = 2.6$ corresponding to tonal frequencies *(a-d)* in Figure 3.12b.

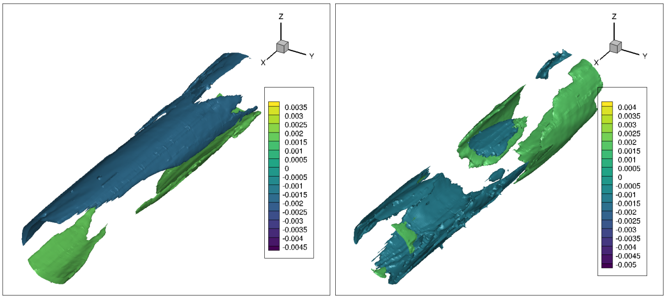
Case 1 - $\sigma = 2.6$

Figures 3.12a-3.12b illustrates the pressure signal and its FFT outside the cavitating core for $\sigma = 2.6$. The response modes corresponding to the labelled spikes (a-d) in FFT of the pressure signal are shown in Figure 3.13. The First mode illustrated in Figure 3.13a shows that the cross section of the vortex develops an eccentricity. The response modes illustrated in Figure 3.13b — 3.13c are relatively complex and suggest changes in both cavity size and eccentricity. The response mode illustrated in Figure 3.13d shows a helical mode of oscillation which corresponds to the peak frequency.

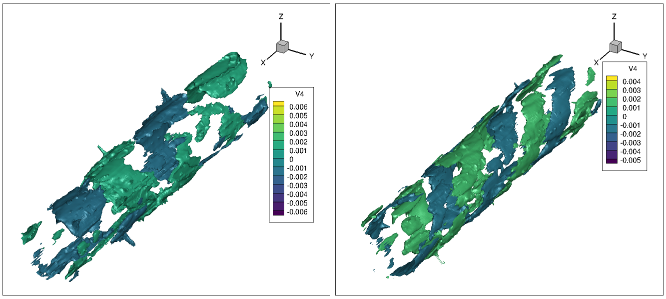


(a) Pressure signal (b) FFT of the pressure signal

Figure 3.14: Probe signal and FFT at Location [5 mm, 5 mm, 0 mm] for $\sigma = 1.7$. The response modes corresponding to tonal frequencies (a-d) in Figure 3.14b are illustrated in Figure 3.15.



(a) $\nu = 105.6182$ Hz (b) $\nu = 402.7549$ Hz



(c) $\nu = 680.5261$ Hz (d) $\nu = 1079.1321$ Hz

Figure 3.15: Response modes of vortex cavity for $\sigma = 1.7$ corresponding to tonal frequencies *(a-d)* in Figure 3.14b

Case 2 - $\sigma = 1.7$

Figures 3.14a-3.14b illustrates the pressure signal and its FFT outside the cavitating core for $\sigma = 1.7$. The response modes corresponding to the labelled spikes (a-d) in FFT of the pressure signal are shown in Figure 3.15. The mode illustrated in Figure 3.15a shows that the cross section of the vortex tends to develop an eccentricity similar to the previous case. The response mode illustrated in Figure 3.15b suggests a breathing mode of the cavity and this corresponds to the peak acoustic emission in the FFT of the pressure signal. Figure 3.15c suggests a response mode which changing cavity size as well as eccentricity. Finally, Figure 3.15d shows a higher harmonic of the helical mode of oscillation.

The domain is initialized with a laminar axisymmetric vortex flow without perturbations thus, the observed dynamics are expected to have developed in the well resolved simulation. One possible source of artificial excitation could be the periodic boundary condition itself. However, the peak responses observed vary for cavities of different sizes and differ meaningfully from $v = \frac{v_\infty}{0.02} \approx 301$ Hz. Thus, it is expected that the observed cavity response and the modes result from excitation due to growing Kelvin-Helmholtz instabilities at the interface. In all our cases, the excitation from KH instabilities appears to be axis-asymmetric and the helical mode of vortex oscillation is observed as a strong tonal deformation.

3.5. CONCLUSION

In this research, the dynamics of an isolated cavitating vortex without forced excitation is simulated. The vortex is initialised with the Proctor-Winkelmans kinematic model fit to experimental data [46]. Observations have been made on the kinematics, dynamics and the relevance of compressibility in simulating acoustics of cavitating vortices.

The tangential velocity of the vortex is observed to differ significantly from that predicted by the Proctor-Winkelmans model at the cavitating core and up to the outer edge of the viscous core. The model over-predicts the tangential velocity at the core and under-predicts the velocity gradient up to the viscous core. The observed deviation implies relatively higher uncertainties in tangential velocity predictions. Higher uncertainties in predicted tangential velocity result in higher uncertainties for predicted core radius of the cavitating vortex. Consequently, any semi-empirical models which predict acoustic radiation of the tip-vortex based on its core radius are also likely to have higher uncertainties.

The presence of a tangential velocity inflection leads to Kelvin-Helmholtz instabilities. These instabilities grow into axis asymmetric excitation sources for the vortex cavity resulting in helical and superimposed response modes. Thus, KH instabilities are expected to influence the acoustic spectrum at higher frequencies and with lower amplitudes compared to axis-symmetric sources which result in breathing modes for the cavitating vortex.

On the relevance of compressibility, it is observed that the vapour core tends to behave as a compressible gas. In contrast density variations in the free-stream are predominantly small. This observation leads to the conclusion that the free-stream may be regarded as an incompressible liquid. The free-stream may influence the dynamics within the cavity however, the momentum within the core is most likely too low to influence free-stream dynamics.

3.6. RECOMMENDATIONS

- Since the current simulation assumes the vortex to be infinite, no collapse is simulated. It is recommended that a finite cavitating vortex be simulated in a much longer domain with a high grid resolution where collapse is not triggered by imposed boundary conditions or pressure gradients.

- In the current simulation, Volume of Fluids leads to a smeared interface, thus, it is recommended that the interface also be modelled and tracked.

- The current simulation imposes no axis-symmetric excitations which are expected on a propeller operating in the wake of a hull. Therefore, it is recommended that a wakefield be specified at the inflow boundary while the vortex center is rotated about the propeller shaft axis.

- The current research simulates a rotating flow on a Cartesian grid. Thus, it is also recommended to repeat the simulation with an O-grid topology.

- The current research simulated two cavitating vortex cases. It is recommended that a larger number of cavitating vortices be simulated to generate data which correlates vortex features and acoustic emissions. The dataset can be used to train machine learning models that predict acoustic spectrum of infinite vortices. This could be a step towards a physics driven tip-vortex acoustic model which can be used for optimization of propellers.

- The current research provides a qualitative insight into sources of uncertainties in predicting the vortex core radius. It is recommended that more cavitating vortices be simulated to quantify the uncertainties as well.

ACKNOWLEDGEMENTS

The authors extend their gratitude and acknowledge the valuable exchange of thoughts and ideas with Ir. Maarten Klapwijk, PhD Candidate, 3mE, TU Delft, The Netherlands and Dr. ir. Johan Bosschers, Senior Researcher, R& D MARIN, The Netherlands.

4

DYNAMIC OPTIMIZATION

This chapter introduces a machine learning approach for optimizing propellers. The method aims to improve the computational cost of optimization by reducing the number of evaluations required to find solutions. This is achieved by focusing search towards design clusters with good performance, i.e. high propulsive efficiency and low cavitation. Three types of clusters are expected. The first cluster constitutes designs with performance of interest, i.e. high efficiency and low cavitation. The second cluster constitutes designs with performance not of interest, i.e. low efficiency and high cavitation. The third cluster constitutes designs whose performance cannot be estimated with the Boundary Element Methods (BEM) that is used in this study. In simple cases with single objective optimization to maximize efficiency, these clusters can be identified a-priori with unsupervised classifiers provided that orthogonally independent parameters are used as demonstrated in this chapter. For multi-objective constrained optimization to maximize efficiency and minimize cavitation, supervised classifiers may be required to learn the clusters. Classical design variables such as chordlength, pitch, skew, rake, thickness distribution and camber of hydrofoils cannot be used to identify these clusters because of multicollinearity. Thus, a new orthogonal parametric model is proposed where the parameters are directly derived from the propeller blade mesh. As the blade surface mesh is used as boundary conditions to solve the governing equations, the orthogonal parameters are expected to have a strong correlations with performance predictions of BEM or Computational Fluid Dynamics (CFD) compared to classical design variables. It is demonstrated that design clusters with good performance can be identified with few BEM evaluations. Furthermore, the method synergises explainable supervised and unsupervised learning to advice search algorithms and quickly guide them to lucrative regions in the design space. However, reducing the cost of optimization results in a trade-off with completeness of the search; this is also investigated in this chapter. The method is demonstrated on a simple fully wetted flow case of the benchmark Wageningen B-4 70 propeller with P/D=1.0, as the geometry and open-water curves are readily accessible allowing back of the envelope verification *and validation of our results* [1].

4.1. INTRODUCTION

The maritime industry has embraced its responsibilities towards the environment and is making active efforts to reduce GHG emissions. The IMO aims to achieve a 50% reduction in GHG emissions by 2050 compared to 2008 levels [70]. In addition to improving thermodynamic efficiency, transitioning to cleaner fuels, and adopting better propulsion architectures, realising this goal requires significantly improving the propulsive efficiency. For modern vessels, propulsive efficiency is typically between 29-35% [80]. Improving the hydrodynamic efficiency of propellers is an important area of focus as it plays a critical role in determining the total propulsive efficiency. Optimizing propellers to improve efficiency and cavitation behaviour is computationally expensive even in the preliminary design phase. This is especially the case for clean-slate designs which do not use legacy propeller series. Traditional preliminary design of propellers involves two steps [84]: In the first step, systematic open-water studies of legacy propeller series are performed to identify the main design parameters. In the second step, a D&O problem is solved to numerically determine specific design parameters. Currently, the maritime industry is transitioning to shape adaptive propellers [97] and also aims for hydroacoustic optimization [64]. Designers may need to start from fresh propeller geometries whose open water curves are not available. Consequently, methods to numerically determine design parameters, i.e. D&O are starting to become the first step in preliminary design. Therefore, it is very important to the improve the reliability of D&O and also reduce its computational cost.

Optimization methods primarily focus on improving efficiency and cavitation behaviour. In addition to efficiency and cavitation objectives, propeller optimization must also account for constraints on hull excitation [88], cavitation erosion [111], ship speed, thrust, and strength. Typically, the lower computational cost of Boundary Element Methods (BEM) motivates their use in the preliminary design phase to predict the efficiency and cavitation behaviour of a blade geometry. To generate the blade geometry, most methods rely on (a) classical design variables - pitch, diameter, hydrofoil camber, thickness distribution, skew and rake [6, 66], (b) control points of splines or coefficients of equations that define propeller families [62, 67, 85] and, (c) free-form control points [68]. Preliminary design methods typically use classical design variables as their values also communicate design specifications. However, classical design variables do not always lead to a good formulation of the optimization problem. An optimization problem is well formulated if design variables can be considered orthogonally independent. In the case of classical design variables, changing a single design variable perturbs multiple nodes on the mesh. Since, the surface is often defined with splines, the translation in mesh nodes is strongly correlated and also predictable. Furthermore, in most cases, multiple classical design variables tend to change the same mesh nodes. This results in *multicollinearity*, i.e. two or more variables being strongly correlated [2]. Thus, it is difficult to

---

[1] The research in this chapter has been approved by reviewers of the journal of Applied Ocean Research. Editor's decision is expected soon.

isolate the impact of individual design variables on objectives and constraints.

Consider the objective

$$\max f = \eta\left(\vec{M}(\vec{s})\right), \tag{1}$$

where the blade geometry is defined by classical design variables $\vec{s}$, η is the hydrodynamic efficiency and $\vec{M}$ is the generated mesh for a blade geometry (see Eq. 11). The optimization problem is well formulated if $\vec{s}$ can be considered orthogonally independent. The Jacobian of f by chain rule is

$$J_{\mathbf{f}} = J_1 \cdot J_2 = \begin{bmatrix} \frac{\partial f}{\partial m_0} & \cdots & \frac{\partial f}{\partial m_q} \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial m_0}{\partial s_0} & \cdots & \frac{\partial m_0}{\partial s_r} \\ \vdots & \ddots & \vdots \\ \frac{\partial m_q}{\partial s_0} & \cdots & \frac{\partial m_q}{\partial s_r} \end{bmatrix}, \tag{2}$$

where J_1 is the Jacobian of f with respect to to the mesh $\vec{M} \in \mathfrak{R}^q$ and J_2 is the Jacobian of the mesh with respect to the design vector $\vec{s} \in \mathfrak{R}^r$. $\vec{s}$ can be considered orthogonally independent if J_2 is an orthonormal matrix, i.e. $\frac{\partial \vec{M}}{\partial s_i} \cdot \frac{\partial \vec{M}}{\partial s_j} = 0$. Because changing a single classical variable perturbs multiple nodes of the mesh and multiple classical design variables tend to change the same mesh nodes, $\frac{\partial \vec{M}}{\partial s_i} \cdot \frac{\partial \vec{M}}{\partial s_j} \neq 0$. Thus, $\vec{s}$ cannot be considered orthogonally independent. This is also the case for other objectives/constraints such as Suction side cavity volume ($V_{cav,b}$), Pressure side cavity volume ($V_{cav,f}$) and tip-vortex cavity radius for a propeller blade. Hence, the relation between η, $V_{cav,b}$, $V_{cav,f}$, r_c and classical design variables is obscured.

Often, in the preliminary design phases, the use of BEM results in relatively low computational cost but comes with high uncertainties [83]. Thus, the relation between objectives, constraints and mesh (J_1 in Eq. 2) is also obscured to an extent. Hence, optimization problems depending on classical design variables and BEM may require more evaluations than necessary to identify lucrative regions in the design space. This is irrespective of whether a gradient descent or gradient free search algorithm is used.

In this chapter, the computational cost of optimization is demonstrated to be reduced by (a) addressing the challenge of multicollinearity which affects J_2 in Equation (2), with a set of orthogonal parameters $\vec{x}$ such that $\frac{\partial \vec{M}}{\partial x_i} \cdot \frac{\partial \vec{M}}{\partial x_j} = 0$ and, (b) applying explainable machine learning to model the uncertainties in predictions made by BEM, which affects J_1 (Equation (2)) in the preliminary design phase.

The orthogonal parameters are derived from SVD of a data matrix populated with variations of the blade geometry. There are two comparable applications of SVD. The first is for the symbolic reformulation of design and optimization problems [78] and the second is for the dimension reduction in single- and multi- disciplinary hull form optimizations [21]. In both cases, relatively small data sets with large dimensionality are reduced to very few dimensions. In [78], optimization problems are even clustered based on symbolic similarities.

BEM typically tends to predict similar performance, i.e. efficiency and cavitation behaviour, for comparable blade geometries. Thus, three clusters are expected during design and optimization. The first cluster constitutes of designs with performance of interest, i.e. high efficiency and low cavitation. The second cluster constitutes of designs with performance not of interest, i.e. low efficiency and high cavitation. The third cluster consists of designs whose performance cannot be estimated accurately due to limitations of the BEM used in this study. The latter two clusters as designs not of interest. Thus, designs of interest are assigned the label $y_c = 1$ and designs not of interest are assigned the label $y_c = 0$. With the use of the proposed orthogonal parameters, it is expected that these clusters are discernable as the relation between Hydrodynamic efficiency (η), $V_{cav,b}$, $V_{cav,f}$ and r_c are not obscured. In simple cases with single objective optimization, these clusters can be identified a-priori with unsupervised classifiers as demonstrated in this chapter. Thus, fewer evaluations are required to identify lucrative regions in the design space. For multi-objective constrained optimization, supervised classifiers are required to learn the clusters and focus search in lucrative regions. The multi-objective constrained optimization with supervised classifiers is demonstrated in a separate chapter 5. In both chapters, explainable classifiers are used [3] as they provide insight into *why* a prediction is made.

The resulting optimization method - Dynamic Optimization - is elaborated in Section 5.2. The method is called Dynamic Optimization because the parameters are derived from the blade meshes that become boundary conditions for the equations governing the *dynamics* of the flow. Subsections 4.2.1-4.2.5 detail the building blocks of the method. Section 4.3 details the demonstration setup and modelling of uncertainty. Section 5.3 presents the results on reliability and performance of the orthogonal parameters, supervised and unsupervised classifiers. The trade-offs between efficiency, completeness and optimality for Dynamic Optimization is demonstrated in Subsection 4.4.3. The conclusions are provided in Section 4.5.

4.2. DYNAMIC OPTIMIZATION

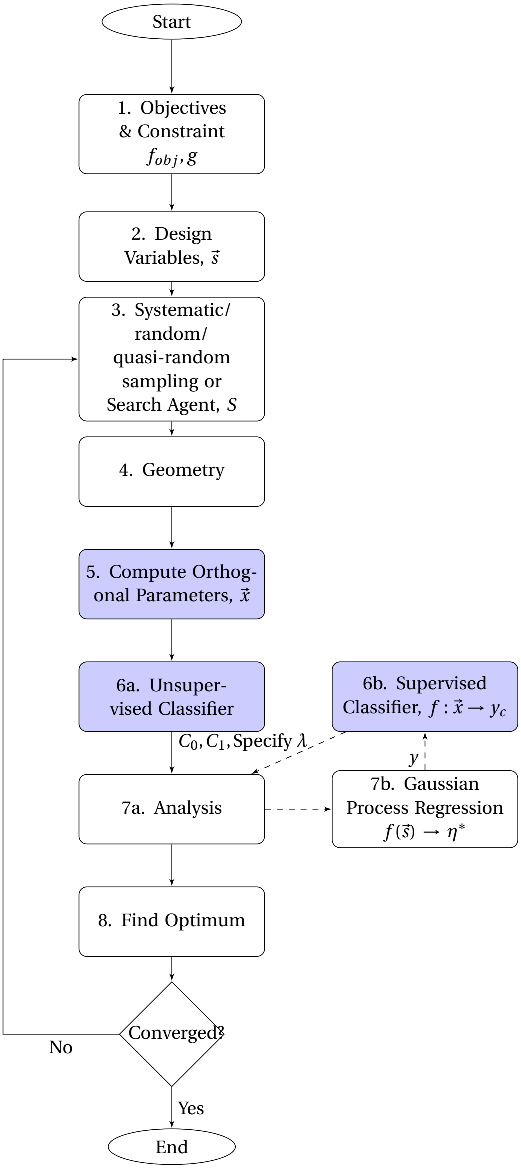


Figure 4.1: Dynamic Optimization workflow

Algorithm 1 Dynamic Optimization - Single Objective

```
Input: $f_1, g_1, g_2, ..., g_m, \lambda, N, \sigma^2$
Output: $\eta_{max}, \vec{s}_{sol}$
procedure DYNAMIC OPTIMIZATION - SINGLE OBJECTIVE
    $S, X \leftarrow \emptyset$
    $\eta_{max} \leftarrow 0$
    $S \leftarrow$ SEARCH(*params)
    for $i \leftarrow (0 \rightarrow N)$ do
        $\vec{s} \leftarrow S_i$
        Generate geometry and mesh ($\vec{M}^*$) for $\vec{s}$
        $\vec{x}^* \leftarrow$ PROJECTION($\mathfrak{R}_{ortho}, k', \vec{M}^*$)    ▷ See Algorithm 5
        $X \leftarrow X \cup \vec{x}^*$
    $X \leftarrow$ NORMALIZE($X$)
    $C0, C1 \leftarrow GET\_CLUSTERS(X, \lambda)$    ▷ See Algorithm 7 in Appendix, $\lambda \in (0, 100.0]$
    $C_{first}, C_{second}, \eta_{max}, \vec{s}_{sol} \leftarrow BEST\_CLUSTER(C0, C1)$    ▷ See Algorithm 8 in Appendix
    #At this point $\lambda/100 \times N$ designs have been evaluated and the better cluster is known.
    $X_{nb} \leftarrow \{x_i | i \in C0 \cup C1\}$
    GNB $\leftarrow$ GAUSSIAN-NB($X_{nb}, y_c$)    ▷ $y_c = 1$ for $C_{first}$ and $y_c = 1$ for $C_{second}$
    $X' \leftarrow X - X_{nb}$
    if $S' \neq \emptyset$ then
        for $i \leftarrow (0 \rightarrow len(X'))$ do
            if GNB($x_i$) == 1 then
                $\eta* \leftarrow$ EVAL($s_i$)    ▷ BEM or CFD.
                if $\eta* > \eta_{max}$ then
                    $\eta_{max}, \vec{s}_{sol} \leftarrow \eta*, s_i$
            else
                Pass    ▷ Skipped an evaluation and reduced cost of optimization.
    return $\eta_{max}, \vec{s}_{sol}$
```

Dynamic Optimization (DO) is a constraint respecting design method, where search agents are advised by classifiers. The workflow is described in Figure 4.1 and the implementation of DO is detailed in Algorithm 1. In the figure each element of the method is labelled. At the start, objectives, constraints and design variables are specified (Block 1-2, Fig. 4.1). Then a search/sampling strategy based on systematic, random or pseudo-random sampling or another search agent is chosen (Block 3). Based on the sampling, the geometry is generated (Block 4, Fig. 4.1). After modelling, a set of orthogonal parameters ($\vec{x}^*$) that define the design are computed. These orthogonal parameters are projections of the propeller mesh $\vec{M}^*$ in a k' dimensional hyperspace ($\mathfrak{R}_{ortho}$). Section 4.2.4 details the derivation of orthogonal parameters. These parameters are defined to address the limitations of traditional parametric models. In the first step, based on the normalized orthogonal parameters (X), the unsupervised classifier finds two clusters ($C0, C1$) within which features are comparable (Block 6a, Fig. 4.1). It is also possible to increase the number of clusters; the number of clusters can be determined by the number of objectives, constraints and also the expected variance in the propeller geometry. Next, the designer needs to specify the percentage of designs (λ) to sample from each each cluster C_i for which the detailed analysis is performed with BEM or CFD (Block 7a, Fig. 4.1). The results of analysis indicates the cluster with best designs (C_{first}). C_{first} contains the solution ($\vec{s}_{sol}$) with the best efficiency (η_{max}). At this point λ percentage of designs have been evaluated and the best design is known. All the evaluated designs X_{nb} and corresponding cluster labels y_c are used to train the supervised explainable Naive-Bayes classifiers (GNB, Algorithm 1:L15 and Block 6b, Fig. 4.1). Based on learnt labels, the classifiers guide search agents to lucrative clusters in the design space while passing on designs that are not in the better cluster (Block 8, Fig. 4.1). False positives are caught with evaluations. The effect of false negatives can be controlled by increasing λ. Thus, the trade-off with completeness of optimization is controlled by the designer via the parameter λ. Finally, if all convergence criteria are satisfied, the design method reports the optimum else it is proposed to explore new regions of the design space. The subsequent sections detail the critical blocks of DO. For the method to be successful, the reliability of Blocks 5, 6a and 6b (emphasised in Figure 4.1) are of critical importance.

4.2.1. OBJECTIVE, CONSTRAINTS AND DESIGN VARIABLES

Propeller design and optimization is a complex problem which requires numerous considerations. One of the primary objectives is to maximise efficiency (η). The problem is constrained by target thrust T_0 at a specified ship speed (v_{ship}). Other constraints to consider are cavitation, noise, hull induced pressure fluctuations and structural integrity.

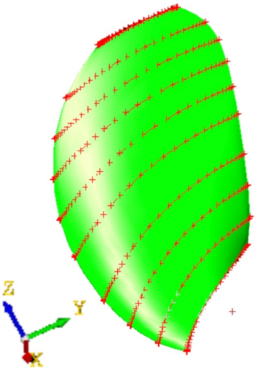
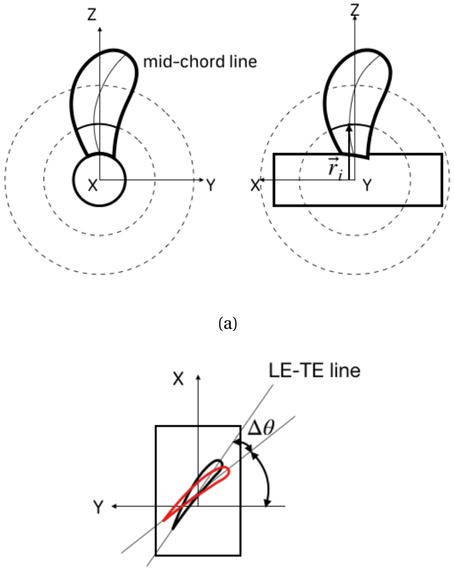


Figure 4.2: Hydrofoil cross-sections and their locations

In this demonstration, the single objective is to maximise the open-water hydrodynamic efficiency (η) of a Wageningen B4-70 propeller blade. Our starting geometry is a blade with P/D=1.0. The propeller is modelled with 10 hydrofoil cross-sections as illustrated in Figure 4.2. The constraints are (a) target thrust $T_0 = 27kN$ and (b) constant $v_{ship} = 1.72ms^{-1}$. From the B4-70 series charts, B4-70 P/D=0.8 is a better starting point for the chosen operating point. However, the baseline is chosen to be B4-70 P/D=1.0 as it is further away from a known local optimum or known optimal search direction. This increases the chances of finding new search directions. In our optimization, the thrust is maintained by modifying the rotation rate of the propeller. The geometry is varied by modifying the pitch distribution, hence, the mean pitch also changes. Therefore, the pitch distribution is not restricted to that of the Wageningen B4 series. The baseline performance is listed in Table 4.3.



(a)

(b)

Figure 4.3: (a) Mid-chord vector ($\vec{r}_i$) and, (b) schematic illustrating pitch modification

There are four primary reasons for defining a simplified optimization problem. Firstly, the simplicity of the problem enables a clear demonstration and better testing of each component of DO. In order to be applied to realistic problems (Part-II), DO must first work for simple propeller optimization problems. Secondly, the open-water curves of the propeller are thoroughly investigated and also publicly available [61]. Thus, the results of the current optimization can be quickly verified and validated with existing open-water curves. Thirdly, the geometry has been published and is readily available without knowledge embargo. Finally, the analysis tool PROCAL [82] has been validated for the open-water fully wetted flow case of this propeller.

Classical design variables such as diameter (D), pitch distribution and camber correlate with η. A larger propeller tends to be more efficient and the upper-bound for diameter is often determined by hull-integration, tip-speed, hull induced pressure fluctuation and cavitation constraints. At the maximum possible diameter, it is possible to improve efficiency by modifying the pitch distribution. Cavitation behaviour is improved by modifying camber. As the primary objective of the demonstration case is efficiency, the aim is to find the optimal pitch distribution.

To modify the blade pitch, the parameter θ is used. Figure 5.4b illustrates the change in geometry when θ is varied. θ is modified from the 4th-10th cross-sections i.e. mid section to blade tip. The target cross-section is rotated by $\Delta\theta$ about the Mid-chord vector ($\vec{r}_i$). $\vec{r}_i$ is the vector parallel to z-axis which points to the midpoint of the LE-TE line as illustrated in Figure 5.4a. Pitch is expected to be sensitive to $\Delta\theta$. An increase in θ results in a higher pitch exposing the blade to higher inflow angles thus also higher loading. Note that $\Delta\theta < 0$ implies that the cross-section was depitched. Otherwise the cross-section's pitch either stayed the same or it increased. While θ may influence skew (γ) and rake (τ), they are not expected to be very sensitive to θ.

4.2.2. HALTON SAMPLING

To demonstrate the existence of clusters, the Halton sequence [69] is used to generate samples. The sequence is chosen as it is a systematic, repeatable and quasi random sampling method. A possible strategy to improve hydrodynamic efficiency is to reduce rotational losses by off-loading the blade [79]. Thus, pitches from the 4th-10th cross-sections are chosen as design variables. They are perturbed from the baseline design by $\Delta\theta_{p,i} = \pm 0.150$.

Within the stated bounds, 200 points are sampled with the Halton sequence to generate the sample data matrix

$$S = \begin{bmatrix} \Delta\theta_{1,4} & \cdots & \Delta\theta_{1,10} \\ & & \vdots \\ \Delta\theta_{N,4} & \cdots & \Delta\theta_{N,10} \end{bmatrix}. \tag{3}$$

The sequence is characterised by low-discrepancy and low-dispersion compared to random sampling methods when $k < 10$, where k is dimensionality of the sampling space [72]. The modified geometries are constructed with B-Splines as detailed in the following section.

4.2.3. GEOMETRY

The blade geometry is defined by a single B-Spline surface [75] defined as

$$\vec{p}(u,v) = \sum_{i=0}^{m}\sum_{j=0}^{n} N_{i,d_1}(u)N_{j,d_2}(v)\vec{P}_{i,j}, \tag{4}$$

where $u, v \in [0.0, 1.0]$ are knots in the u and v directions; $\vec{P}_{i,j}$ is the control point of the i-th row and j-th column; $m+1$ and $n+1$ are the number of rows and columns the control points are organized in; d_1, d_2 are the degree of the B-Spline in u and v directions respectively and; N_{i,d_1} and N_{j,d_2} are the basis function in u and v directions respectively. The surface is fitted to hydrofoil cross-sections in Cartesian coordinates at different radial locations. In our case $d_1 = 3$ as the spline is closed in u direction and $d_2 = 8$ to find an optimal fit for a hydrofoil with camber and thickness distribution with fewer control points. The basis functions $(N(t))$ for each directions are defined as

$$N_{i,0}(t) = \begin{cases} 1; t_i \leq t < t_{i+1} \\ 0; otherwise \end{cases}, \tag{5}$$

$$N_{i,d} = \frac{t - t_i}{t_{i+d} - t_i} N_{i,d-1}(t) + \frac{t_{i+d+1} - t}{t_{i+d+1} - t_{i+1}} N_{i+1,d-1}(t). \tag{6}$$

where t is the knot value and d is the degree of the spline. Based on the geometry, the orthogonal parameters are derived.

4.2.4. ORTHOGONAL PARAMETRIC MODEL

The primary objective of the proposed parametric model is to alleviate the limitations of classical design variables. Thus, the parametric model must (a) accurately quantify geometric variations, (b) yield orthogonally independent parameters, i.e. $\frac{\partial \vec{M}}{\partial x_i} \cdot \frac{\partial \vec{M}}{\partial x_j} = 0$ and, (c) be generalized for all possible geometries. To satisfy the above requirements, the parametrisation is based on the coordinates of the blades' surface mesh such that all possible variations of all possible families are represented and the relation between geometry and design objectives is not obscured. The orthogonal parameters $x_i \in \vec{x}$ of the proposed parametric model are derived by projecting the surface mesh of propellers in an orthonormal hyperspace $\mathfrak{R}_{orth}$ to ensure that the parameters are orthogonally independent. Defining $\mathfrak{R}_{orth}$ involves four steps.

The *first step* is to ensure that the mesh used for the analysis of each design is comparable. Thus, the surface mesh is generated by uniformly discretising the B-Spline surface in the B-Spline parametric space, i.e. along the u and v directions of the spline surface. The mesh M^* consisting of discrete nodes in the B-Spline parametric space is defined in Equation (10), where m is the number of nodes on the hydrofoil cross-section and n is the number of cross-sections along the diameter. Thus, the mesh for the i-th propeller $\vec{M}_i \in \mathfrak{R}^k_{cart}$ with dimensionality $k = 3 \times m \times n$ is

$$M^*_{m\times n} = \begin{bmatrix} (0.0, 0.0) & \cdots & \left(0.0, \frac{(j-1)}{(n-1)}\right) & \cdots & (0.0, 1.0) \\ \vdots & \ddots & & & \vdots \\ \left(\frac{(i-1)}{(m-1)}, 0.0\right) & & \left(\frac{(i-1)}{(m-1)}, \frac{(j-1)}{(n-1)}\right) & & \vdots \\ \vdots & & & \ddots & \vdots \\ (1.0, 0.0) & \cdots & \cdots & \cdots & (1.0, 1.0) \end{bmatrix}, \tag{7}$$

$$\vec{M}_i = \{p(m^*_{ij}) | m^*_{ij} \in M^*_{m\times n}\}. \tag{8}$$

The *second step* is to build the data matrix M by stacking meshes of all N sampled propellers and then transforming them into cylindrical coordinates:

$$\tilde{M} = \begin{bmatrix} \vec{M}_1 \\ \vdots \\ \vec{M}_N \end{bmatrix}, \tag{9}$$

$$M = T(\tilde{M}), T : \mathfrak{R}^{N\times k}_{cart} \to \mathfrak{R}^{N\times k}_{cyl}. \tag{10}$$

This pre-processing step of stacking and transforming the meshes into cylindrical coordinates reduces the variance in $M_{(:,i)} \forall i \in \{0,3,6...\}$ as cross-sections are defined at a constant radius from the propeller axis. Thus, variance in form and shape of the blade is emphasised in the data matrix.

The *third step* is to define the origin of $\mathfrak{R}_{orth}$. The origin is chosen to be the mean

$$\vec{\mu} = \frac{\sum_{i=1}^{N} \vec{M}_i}{N}, \tag{11}$$

of all N propeller meshes, leading to the translated data matrix

$$M' = \begin{bmatrix} \vec{M}_1 - \vec{\mu} \\ \vdots \\ \vec{M}_N - \vec{\mu} \end{bmatrix}. \tag{12}$$

The *fourth step* is to factorise M' using thin SVD [81], that is,

$$M'_{k\times N} = U_{k\times N} \cdot \Sigma_{N\times N} \cdot V'_{N\times N}. \tag{13}$$

This yields the left singular vectors U, the singular values Σ and the right singular vectors V. The product of the k' left and k' right singular vectors yields the orthonormal reference frame

$$\mathfrak{R}_{orth} = U_{k\times k'} \cdot V'_{k'\times k'}. \tag{14}$$

$\mathfrak{R}_{orth}$ is defined by k' orthonormal vectors and any surface mesh $\vec{M}$ can now be projected in $\mathfrak{R}_{orth}$. Thus the surface mesh can be represented as

$$\vec{M} = \sum_{i=0}^{k'} Mcos(x_i)\hat{\mathfrak{R}}_{orth:,i} \tag{15}$$

where $M = |\vec{M}|$, $\hat{\mathfrak{R}}_{orth:,i}$ is the i-th unit basis vector of the reference frame $\mathfrak{R}_{orth}$ and x_i is the projection angle of the mesh on the i-th unit basis vector. Thus, the orthogonal parameters proposed are x_i. The orthogonal parameters can be represented by $\vec{x} \in \mathfrak{R}^{k'}$. Note that $\frac{\partial \vec{M}}{\partial x_i} \cdot \frac{\partial \vec{M}}{\partial x_j} = 0$ for all values of k'.

The procedure to project the mesh in $\mathfrak{R}_{ortho}$ and compute the orthogonal parameters is detailed in Algorithm 5. This algorithm then yields the Parameter Matrix (X) with values for each vector $\vec{x}$. As $k' << k$, it is possible to significantly reduce the number of parameters while not obscuring the relation between the geometry and the mesh, thus also the solution. Hence, it is ensured that this parametrisation also opens doors for dimension reduction.

One limitation of the proposed parametric model is that perturbing orthogonal parameters results in non-conformal projections in the orthonormal spaces. These projection are non-conformal because they may not result in smooth splines which are required for propeller blade geometries. Thus, models similar to conformal deformation proposed for Free Form Deformation (FFD) [74] need to be formulated for orthogonal parameters as well. Another possibility is to establish a one-to-one map between orthogonal parameters and classical design variables to allow for conformal modifications of the geometry. In the current demonstration, classical design variables are used to modify the geometry and orthogonal parameters are used as features for classifiers and regressors.

Algorithm 2 Subspace Projection

```
procedure PROJECTION
    $M_{N\times k} \leftarrow$ Data matrix of designs to analyze
    Compute $\mathfrak{R}_{ortho}$
    $\mathfrak{R} \leftarrow \mathfrak{R}_{ortho}$
    $X \leftarrow 0_{N,k'+1}$
    for $i = 1$ to $N$ do
        for $j = 1$ to $N'$ do
            $X_{i,j} \leftarrow cos^{-1}\left(\frac{M_{i,:}\cdot\mathfrak{R}_{:,j}}{|M_{i,:}||\mathfrak{R}_{:,j}|}\right)$
     return $X_{N,k'+1}$
```

4.2.5. CLASSIFIERS

Both supervised and unsupervised classifiers are employed. Explainable machine learning models are opted for because they provide insights into why certain predictions are made. The goal of the unsupervised classifier is to identify existing clusters in the design space with comparable geometries and the goal of supervised classifiers is to learn which cluster of geometries are lucrative. Once trained, the supervised classifiers can guide search algorithms towards more lucrative regions in the design space thus reducing the number of required evaluations. Implementations of supervised and unsupervised classifiers in SciKit-Learn [107] are used.

The classifiers use orthogonal parameters (X) as features given that they capture geometric variance reliably. In principle, the orthogonal parametric model is expected to satisfy the *i.i.d* assumption; thus, classifiers based on both generative models $p(X|y_c)$ and discriminate models $p(y_c|X)$ are applicable. In the following sections, a concise description of the classifiers is provided. For more details, the readers are advised to read the references in the respective subsections.

UNSUPERVISED CLASSIFIER

In the zero-th design iteration, since the number of clusters is pre-determined and clusters are expected to differ in shape and share boundaries, Gaussian Mixture Models (GMM) with full covariance matrices [63] is employed. The GMM classifier predicts class labels y_c as

$$y_c = \text{argmax}_j\ \pi_j p(\vec{x}_i|\mu_j, \sigma_j^2), \tag{16}$$

where j maximizes the probability of $\vec{x}_i$ occurring. In Equation (16), $\pi_j = n_j/N$ is the ratio of the number of designs in the j-th cluster (n_j) to the total number of designs N, thus it is also the cluster probability. The means μ_j and variance σ_j^2 for the j-th cluster are found using the Expectation-Maximisation algorithm with a Euclidean distance based K-Means centroid initialisation.

SUPERVISED CLASSIFIERS

Naive-Bayes (NB) with kernel approximations of Support Vector Classifiers (SVC) [18] and Gaussian Process Classifiers (GPC) [76] are compared. The classifiers are as follows:

NAIVE-BAYES

The Naive-Bayes classifier predicts labels as

$$y_c = \text{argmax}_j\ \pi_j p(\vec{x}_i|\vec{\mu}_j, \vec{\sigma^2}_j), \tag{17}$$

where j maximizes the probability of $\vec{x}_i$ occurring. In Equation (18), the probability (p)s of $\vec{x}_i$ occurring given the mean and variance $\vec{\mu}_j, \vec{\sigma^2}_j$ for the j-th cluster is

$$p(\vec{x}_i|\mu_j, \sigma_j^2) = \prod_{d=1}^{k'} \frac{1}{\sqrt{2\pi\sigma_{j,d}^2}} . \exp\left(-\frac{|x_{i,d} - \mu_{j,d}|^2}{2\sigma_{j,d}^2}\right). \tag{18}$$

k' is the number of orthogonal parameters, $x_{i,d}$ is the d-th parameter in $\vec{x}_i$, $\mu_{j,d}$ and $\sigma^2{}_{j,d}$ are the mean and variance of $X_{:,d}$.

SUPPORT VECTOR CLASSIFIER

The decision function for SVC predicts the label y_c as

$$y_c = sgn\left(\sum_{n=1}^{N} y_i \alpha_i \mathrm{k}_* + \rho\right), \tag{19}$$

where $y_i\alpha_i$ are the dual coefficients, k_* is the correlation vector and ρ is the intercept. For our comparison study, the i-th components of the correlation vector are computed with the RBF kernel,

$$k_i(x_i, x_*) = \exp(-|x_i - x_*|^2). \tag{20}$$

Multi-label classifications are implemented with an one-against-one approach [71]. In Equation (19), the dual coefficients ($y_i\alpha_i$) are solutions to the dual problem of SVCs primal problem [18]. The dual problem defined as

$$\min_{\alpha} \frac{1}{2} \alpha^T Q \alpha - \mathbb{1}^T \alpha, \tag{21a}$$

$$\text{subject to } y^T \alpha = 0; 0 \leq \alpha_i \leq C; \tag{21b}$$

where $C > 0$ is the upper bound, $Q_{ij} = y_i y_j \phi(x_i)\phi(x_j)$ and $\phi(x_i), \phi(x_j)$ are the feature map of input x_i, x_j.

GAUSSIAN PROCESS CLASSIFIER

The Gaussian Process Classifier (GPC) is based on the Laplace approximation method described by Rasmussen et. al [76]. The posterior prediction ($\bar{\pi}_*$) for a query x_* is

$$\bar{\pi}_* = \int \sigma(f_*) p(f_* | X, y, x_*) d f_*. \tag{22}$$

Multi-label predictions can be made with both one-against-one and one-against-rest approaches.

In Equation (22), the *first* term, $\sigma(f_*) = \sigma(f(\vec{x}_*)) = \sigma(x^T w)$. GPC places a GP prior with a RBF kernel, Equation (20), on a latent function $f(\vec{x}) = x^T w$. The range of $f(\vec{x})$ is $(-\infty, +\infty)$. Mapping the range to $[0,1]$ allows it to be interpreted as probability of labels in soft classification between two classes $y_c \in \{0,1\}$. Thus the range of $f(\vec{x})$ is squashed through the logistic link function

$$\sigma(f(\vec{x})) = \frac{1}{1 + exp(-f(\vec{x}))}, \tag{23}$$

that maps the domain $(-\infty, +\infty)$ onto $[0,1]$.

In Equation (22), the *second* term $p(f_* | X, y, x_*)$ can be approximated with the Laplace approximation method. The term can be expanded as

$$p(f_* | X, y, x_*) = \int p(f_* | X, x_*, \mathbf{f}) p(\mathbf{f} | X, y) d\mathbf{f}, \tag{24}$$

where $p(\mathbf{f}|X, y) = p(y|\mathbf{f}) p(\mathbf{f}|X) / p(\mathbf{y}|X)$. In the Laplace approximation method, a Gaussian approximation $q(\mathbf{f}|X, y)$ to the posterior of the latent function $p(\mathbf{f}|X, y)$ is utilized. The Gaussian approximation is obtained by doing a second order Taylor expansion of $\log(p(\mathbf{f}|X, y))$ around the maximum of the posterior as defined below:

$$q(\mathbf{f}|X, y) = \mathcal{N}(\mathbf{f}|\hat{\mathbf{f}}, A^{-1}) \approx exp(-\frac{1}{2}(\mathbf{f} - \hat{\mathbf{f}})^T A (\mathbf{f} - \hat{\mathbf{f}})), \tag{25}$$

where $\hat{\mathbf{f}} = \text{argmax}_{\mathbf{f}} p(\mathbf{f}|X, y)$ and $A = -\nabla\nabla log\, p(\mathbf{f}|X, y|(\mathbf{f} = \hat{\mathbf{f}}))$ is the Hessian of the negative log posterior at the point.

CLASSIFICATION LABELS (y_c)

Designs of interest are labelled $y_c = 1$ and designs not of interest are labelled $y_c = 0$. y_c is determined as

$$y_c = \begin{cases} 0 & \bar{\eta}_*(\vec{s}, T_0) \in [\eta_{min}, \eta_{min} + \delta] \\ 1 & \bar{\eta}_*(\vec{s}, T_0) \in [\eta_{min} + \delta, \eta_{max}] \end{cases}. \tag{26}$$

Note that the labels are based on posterior efficiency prediction $\bar{\eta}_*$ and a parameter $\delta = (\eta_{max} - \eta_{min})(2/3)$, where η_{max}, η_{min} are the maximum and minimum efficiencies of all designs in the samples. $y_c = 1$ indicates a lucrative cluster where search is to be focused.

UNCERTAINTY IN EARLY DESIGN

The posterior efficiency prediction $\bar{\eta}_*$ is made by modelling the uncertainties in BEM predictions [83]. Typically, uncertainties are addressed in the detailed design stage. Those resulting from geometry and operational parameters are dealt with Robust Design methods [73], where the primary objective is to reduce the parameter dependent variance in performance. Uncertainties resulting from modelling and discretisation errors are dealt with higher-fidelity CFD. However, delayed consideration of uncertainties results in lower marginal gains in performance. Thus, there is a need to account for uncertainties already in preliminary stages of design. For this purpose, Gaussian Process Regression is employed as detailed in this sub-section.

There are two primary sources of uncertainty. The first is the use of BEM for which performance predictions could have uncertainties as high as 15% [83]. The second is uncertainties resulting from regression itself. Both uncertainties are accounted for in a linear regression model with Gaussian noise as described by Rasmussen et. al [76]:

$$y = f(s) + \epsilon, \tag{27a}$$

$$f(s) = \phi(s)^T w, \tag{27b}$$

$$\epsilon \sim N(0, \sigma_n^2). \tag{27c}$$

Equation (27a) describes the response y with noise ϵ. Note that ϵ models both analysis and regression uncertainties. The function $f(\vec{s}) : \mathbb{R}^N \rightarrow \mathbb{R}$ is assumed to be linear as described in Equation (27b). The inputs $\vec{s} \in \mathbb{R}^N$ are projected onto feature spaces by the function ϕ. The vector space of the feature vector is the feature space. The noise ϵ is assumed to be an *i.i.d* Gaussian distribution with $\mu = 0$ and variance σ_n^2 as described in Equation (27c) is used. The posterior prediction of mean and variance are

$$\bar{f}_* = \mathbf{k}_*^T (K + \sigma_n^2 I)^{-1} \mathbf{y}, \tag{28a}$$

$$\mathbb{V}[f_*] = k(s_*, s_*) - \mathbf{k}_*^T (K + \sigma_n^2 I)^{-1} \mathbf{k}_*. \tag{28b}$$

In Equations (28a) - (28b), K is the Gram matrix and $\mathbf{k}_*$ is the correlation vector. In this study, the components of the correlation vector and Gram matrix are computed with an RBF kernel, see Equation (20).

4.3. DEMONSTRATION SETUP

As a benchmark case, the Wageningen B4-70 propeller with P/D=1.0 is chosen. The geometry and mesh is illustrated in Figure 5.2a.

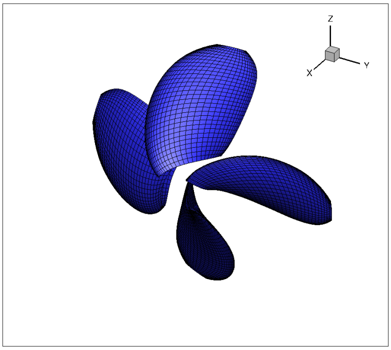


Figure 4.4: Wageningen B4-70 propeller

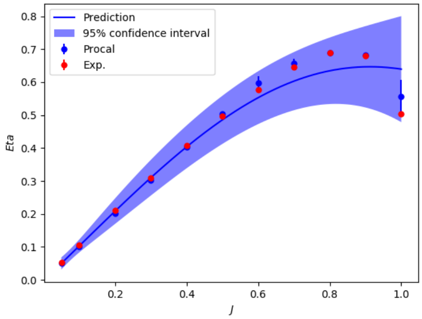


Figure 4.5: PROCAL validation and Gaussian Process Regression results of open-water curve with modelled uncertainties

Table 4.1 Operational Parameters

Parameter	Units	Value
Open-Water Speed	ms^{-1}	[0.18-3.6]
Free-stream density	$kg.m^{-3}$	1025
Absolute atmospheric pressure	Pa	102500
Propeller rotation speed	$rev \cdot s^{-1}$	15

PROCAL is validated for the operating parameters listed in Table 4.1. The validation results are illustrated in Figure 4.5. The experimental data (*Exp.*) in Figure 4.5 is bundled and shipped with PROCAL. It can be observed that the PROCAL efficiency predictions match well with experimental observations up-to $J \approx 0.9$ and thereafter uncertainties increase. In this demonstration, the hub is not modelled. However, PROCAL predictions for hydrodynamic efficiency are comparable to experimental observations. This could be because of weak blade-hub interactions. One limitation of the current approach is that root cavitation is not captured during the optimization and if blade-hub interactions are strong the uncertainties in performance predictions could also be higher.

In the current setup, the resulting uncertainties in performance prediction are accounted for with the GPR estimator. The posterior efficiency prediction ($\bar{\eta}_*$) (-) and confidence interval is estimated with the Gaussian Process Regression (GPR) model defined in Equations 28a-28b. The variance is estimated to be 20% of the response η. This yields a 95% confidence interval that encapsulates predictions, experimental observations as well as uncertainties resulting from regression. For the GPR estimator, the mean prediction (-) would coincide with the experimental observations if the variance is assumed to be 0%. However, assuming a variance of 0% implies that PROCAL predictions do not differ from experimental observations, which is not the case.

For varied geometries and operational conditions the residuals of the panel method can be different. Thus, the variance expected can be different from the estimated 20% in this demonstration. For the sake of simplicity, the expected variance is assumed to be a constant. However, the trends in variance w.r.t geometries, operational conditions and residuals of numerical methods, if available, can also be considered in the GPR estimator during training (Block 7b in Figure 4.1). This could lead to better cluster predictions made by the supervised classifier (Block 6b in Figure 4.1). Better cluster predictions could reduce the number of iterations required to convergence.

4.4. RESULTS

For DO to be successful, orthogonal parametric model (Block 5) and classifiers (Block 6a-6b in Figure 4.1) need to be effective and reliable. The orthogonal parametric model is deemed reliable if it (a) is able to identify spatially discernible clusters and, (b) preserves the geometric correlations. The unsupervised classifiers are deemed reliable when they identify existing clusters. The supervised classifiers are deemed reliable if they direct search to lucrative clusters while continuously learning. The following sections report the results of reliability tests for the critical blocks and DO itself.

4.4.1. ORTHOGONAL PARAMETRIC MODEL RELIABILITY

In this section, the results of the tests on reliability of the orthogonal parametric model are discussed. The ability of the orthogonal parametric model to yield spatially discernible clusters are determined with *silhouette scores*. The *silhouette score* [77] indicates whether there are spatially separated clusters in the data. The score for a design $\vec{x}_i$ is defined as

$$S(\vec{x}_i) = \frac{b(\vec{x}_i) - a(\vec{x}_i)}{max\{a(\vec{x}_i), b(\vec{x}_i)\}}, \tag{29a}$$

$$a(\vec{x}_i) = \frac{1}{|C_i| - 1} \sum_{j \in C_i, i \neq j} d(i,j), \tag{29b}$$

$$b(\vec{x}_i) = \min_{l \neq i} \frac{1}{|C_l|} \sum_{j \in C_l} d(i,j), \tag{29c}$$

where $a(\vec{x}_i)$ is the average dissimilarity of $\vec{x}_i$ with all other designs in the same cluster and $b(\vec{x}_i)$ is the minimum value of average dissimilarity with another cluster. This results in $S(\vec{x}_i) \in [-1,1]$, with -1 indicating wrong labels, 0 indicating similarities between designs from different clusters thus, no clearly discernible clusters, and 1 indicating clearly distinguished clusters. Discerning clusters in a high-dimensional hyperspace is challenging. However, if clusters do exists, they would be discernible in combinations of any two parameters, $s_i, s_j \in \vec{s}$ or $x_i, x_j \in \vec{x}$. Thus, *silhouette scores* are calculated for all possible combinations of two parameters in $\vec{s}$ and $\vec{x}$.

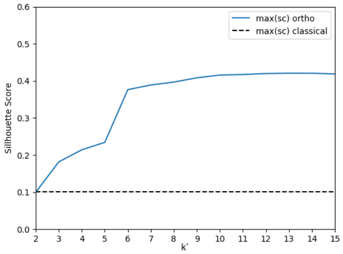


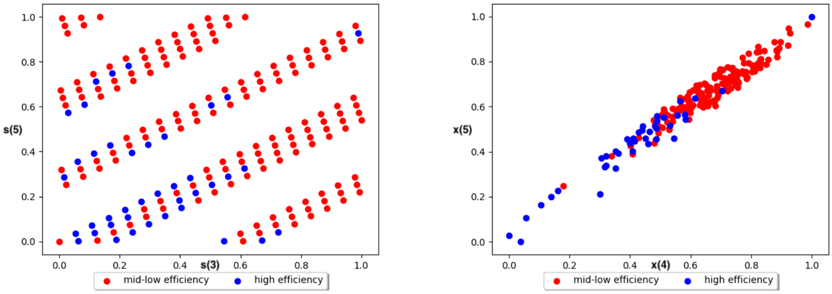
Figure 4.6: Silhouette Scores for 2 parameter combinations with $k' \in [2,15]$

The maximum score of all possible parameter pairs is a good indicator of performance and is plotted in Figure 4.6. The orthogonal parameters tend to perform significantly better than classical design variables with 4 times higher silhouette scores. It is also observed from Figure 4.6 that the score does not improve anymore for $k' > 11$. This indicates that most of the variance is captured in the first 10 orthonormal vectors defining $\mathfrak{R}_{orth}$. Thus, representing the designs in higher dimensions does not yield more distinguished clusters.

Figure 4.7 illustrates the clusters generated by classical and orthogonal parameters for thrust requirement $T_0 = 27kN$. Note that $s(1) = \Delta\theta_4$, thus in Figure 4.7a, $s(3), s(5)$ represent normalised $\Delta\theta$ for the 7th and 9th cross-section respectively. For classical design variables, it is observed in Figure 4.7a that there are no clearly distinguishable design clusters. However, it is observed in Figure 4.7b that orthogonal parameters result in marginally overlapped clusters with one of the two containing lucrative designs.

4.4.2. CLASSIFIER PERFORMANCE

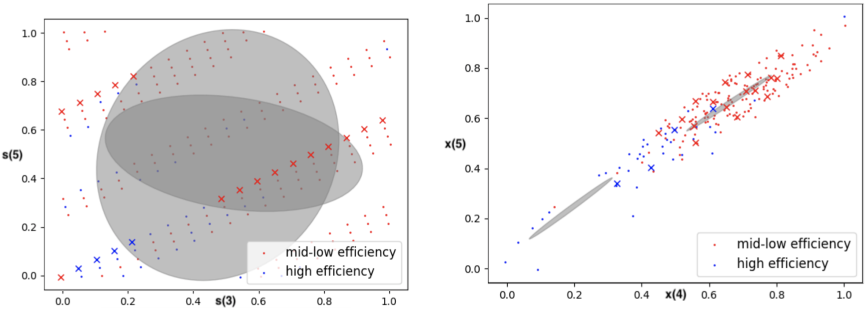
In this section, the results of reliability tests for the classifiers (Block 6a-6b) are presented and discussed. The ability of the unsupervised classifier - Gaussian Mixture Model Classifier - to identify existing clusters with orthogonal parameters and classical design pa-



(a)

(b)

Figure 4.7: Best spatial clustering parameter combinations at $T_0 = 27kN$ for (a) Classical design variables $\Delta\theta_7$ and $\Delta\theta_9$, $s(\{\Delta\theta_7, \Delta\theta_9\}) = 0.1007$. (b)Orthogonal design parameters x_4, x_5 when $k' = 13$, $s(\{x_4, x_5\}) = 0.4206$.



(a)

(b)

Figure 4.8: GMM classifiers confidence ellipsoides for (a) Classical design variables $\{\Delta\theta_7, \Delta\theta_9\}$ and (b) Orthogonal parameters $\{x_4, x_5\}$, when $k' = 13$.

rameters is tested. Thereafter, the supervised classifiers' reliability in predicting whether a design is efficient or otherwise is tested with precision scores.

GAUSSIAN MIXTURE MODEL (GMM)

The GMM classifier is used in the zero-th design iteration when no performance information is available yet. The classifier's ability to generalise depends strongly on the presence of discernible clusters. Testing the classifier's reliability via validation studies is not possible in principle as the ground truth is not known. Thus, the reliability is often implied from the posterior silhouette scores.

From the silhouette scores (Figure 4.6), it is observed that certain parameter combinations have high silhouette scores. Thus, for parameter combinations with high scores, confidence ellipsoids of the trained unsupervised GMM classifier are plotted in Figures 4.8a - 4.8b . It is observed that for classical parameters, the confidence ellipsoids of the classifier significantly overlap. This is clear indication that a-priori clustering of designs is not possible with classical parameters. On the contrary, there is no overlap in the confidence ellipsoids of orthogonal parameters indicating that a-priori clustering of designs is possible.

SUPERVISED CLASSIFIERS

The reliability of the NB classifier, SVC and GPC is evaluated with *precision scores*. The labels (y_c) for the training and testing samples are detailed in Section 4.2.5. The precision score for a classifier

$$A = \frac{tp}{tp + fp}, \tag{30}$$

is the ratio of true positive (tp) predictions to the sum of true positive and false positive (fp) predictions. This scoring is a reliable indicator when the number of members in each class for the testing set are disproportionate. Precision scores, $A > 0.5$ indicate that the classifier predictions had more true positives than false positives. The machine learning models used in DO aim to maximize true positives and the optimization method is designed to catch false positives thus improving reliability. Weighted average precision scores for all class labels are determined with 10 fold stratified cross validation studies. This reduces the risk of missing false negatives.

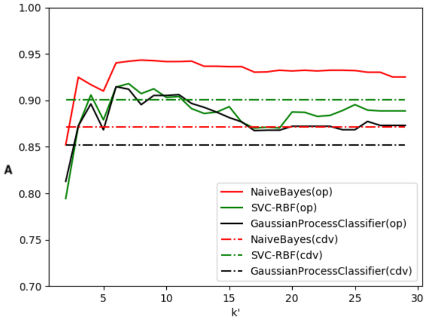


Figure 4.9: Mean precision scores for stratified 10 fold cross validation studies of supervised classifiers

Figure 4.9 illustrates the results of the cross validation study for the classifiers. The plots show the mean cross-validation precision score for predicting $y_c = 0$, i.e. for an inefficient design. It is observed that all classifiers have $p > 0.5$. Thus, true positive predictions are more likely than false positive predictions.

It is observed that precision scores for classifiers using orthogonal parameters are typically higher than those using classical design variables. However, SVC with orthogonal parameters tends to perform better than SVC with classical parameter only for $5 < k' < 11$. This could be because the RBF kernel for SVC tends to be a good feature map for classical design variables in the current demonstration case. Among the classifiers, NB has the highest precision score with $p = 0.9433$ for $k' = 8$ when using orthogonal parameters. For $6 \leq k' \leq 11$, all classifiers using orthogonal parameters outperform their counterparts that use classical design variables. For the domain $k' > 11$, it is inferred from silhouette scores that there is no additional geometric information, however, the dimensionality of the problem increases. This adversely affects classifier performance. Nonetheless, it is observed that NB with orthogonal parameters is the most reliable classifier for the identification of in-efficient designs. The reliability of NB with orthogonal parameters for multi-objective constrained optimization with supervised classifiers is demonstrated in a separate publication [86]

The precision scores indicate that orthogonal parameters are better able to capture geometric variations and not obscure the relation between geometry, mesh and predicted performance. This addresses the first problem of multicollinearity. Furthermore, it enables more training and deploying more reliable machine learning models.

4.4.3. DYNAMIC OPTIMIZATION PERFORMANCE

To test the effectiveness of orthogonal parameters in identifying clusters of designs, it is run 50 times on the same Halton sample set. The goal of the test is to quantify possible gains of using orthogonal parameters versus classical design variables. The tests are run for two setups. In *Setup 1*, GMM in DO uses classical design variables for a-priori cluster identification and in *Setup 2*, it uses orthogonal parameters. In both setups, the supervised classifier is chosen to be the Naive-Bayes classifier due its high reliability. The following sections describe the optimum and the results of the test.

OPTIMUM

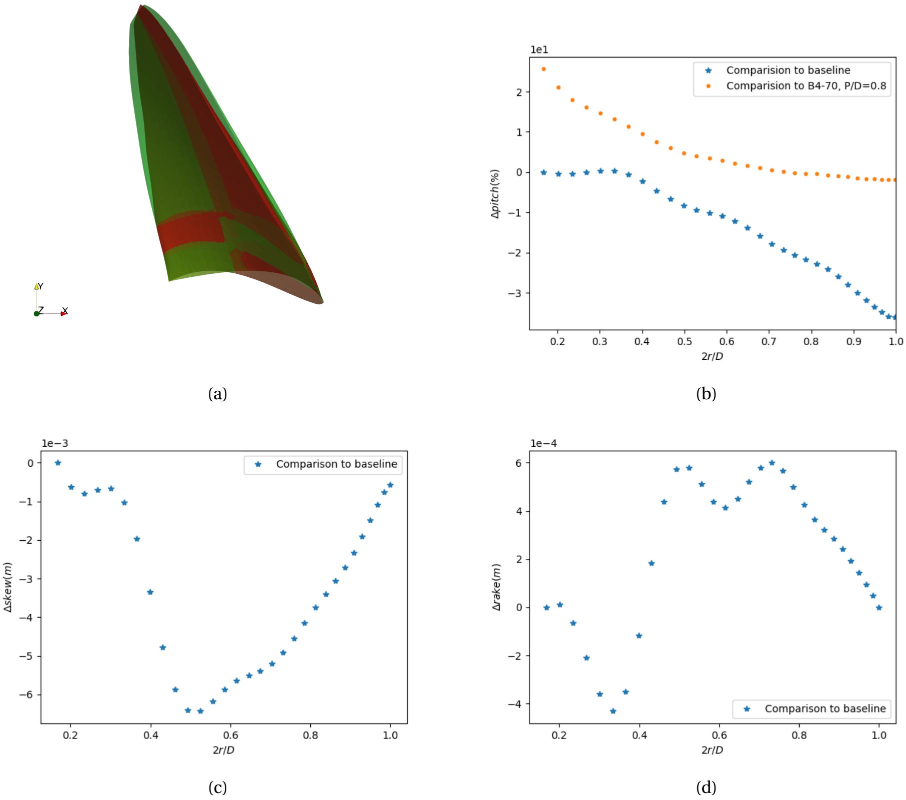
Table 4.2 Optimum in Halton samples - Parameters

	Optimum (rad)
$\Delta\theta_{p,4}$	-0.0409
$\Delta\theta_{p,5}$	-0.0577
$\Delta\theta_{p,6}$	-0.0794
$\Delta\theta_{p,7}$	-0.0868
$\Delta\theta_{p,8}$	-0.0978
$\Delta\theta_{p,9}$	-0.1086
$\Delta\theta_{p,10}$	-0.1113

The optimal solution in the Halton samples is shown in Figure 4.10a. The design parameters are listed in Table 4.2. Note that $\Delta\theta_{p,i} < 0.0$, this implies that the blade pitch has been reduced from the 4th-10th cross-section. $\Delta\theta$ is also the magnitude of depitching in radians. Thus, the optimum is found by depitching the blade.

Table 4.3 Optimum in Halton samples - Performance

	Units	B4-70, P/D=1.0	B4-70, P/D=0.8	Optimum
J	-	0.4035	0.3353	0.3553
n	rps	1.9800	2.3910	2.2500
K_t	-	0.3133	0.2374	0.2439
K_q	-	0.0493	0.0308	0.0315
T (eq. constraint)	kN	27.0000	27.0000	27.0000
η (objective)	-	40.77 %	41.18 %	43.74 %



(a)

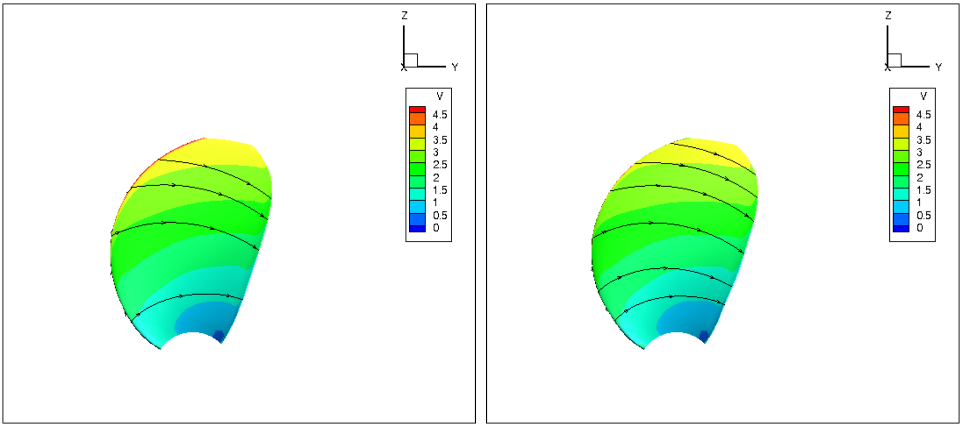
(b)

(c)

(d)

Figure 4.10: Comparison of the optimum and baseline (a) geometry overlaid with the baseline Wageningen B4-70 P/D=1.0 blade (green), (b) percentage change in pitch distributions of optimum, baseline and B4-70 P/D=0.8 blade (c) delta skew distribution for baseline and optimum, (d) delta rake distribution for baseline and optimum.

The performance details are listed in Table 4.3. It is observed that the optimum delivers the target thrust of $27kN$ for the specified $v_{ship} = 1.72ms^{-1}$ with higher efficiency



(a)

(b)

Figure 4.11: Super-velocities (a) baseline and, (b) optimum.

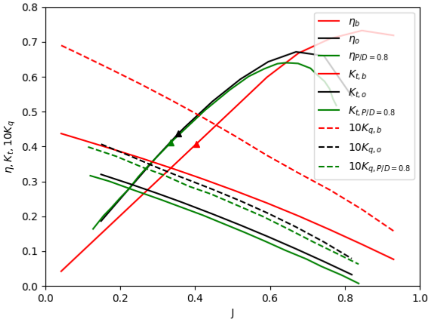


Figure 4.12: Open water curves of optimum in Halton samples and baseline design. The markers are the point of operation for $v_{ship} = 1.72 ms^{-1}, T_o = 27kN$.

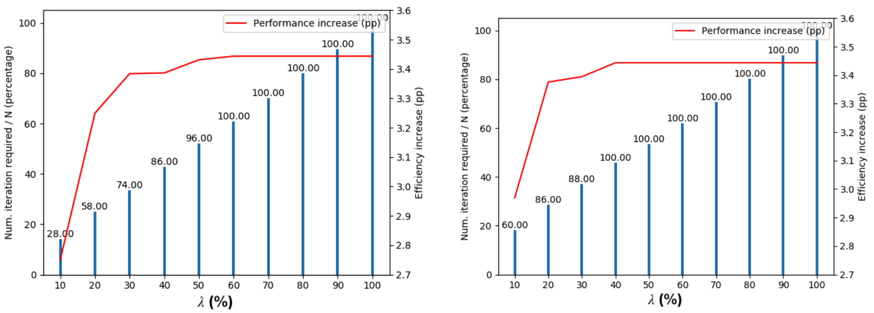
than the baseline. The gains in efficiency are possible when axial, rotational and viscous losses are reduced [79]. Figure 4.10a-5.17f compares the optimum to the baseline. The optimum's pitch distribution is also compared with a Wageningen B4-70 P/D=0.8 blade as its efficiency at the operating point is better than the baseline. Furthermore, its open-water efficiency curve is closest to that of the optimum in the B4-70 series as shown in Figure 4.12.

Figure 5.17c shows that the pitch for the optimum is reduced at the mid-section and tip compared to the baseline. However, the pitch for the optimum at the root is higher than the Wageningen B4-70 P/D=0.8 blade and similar at the tip. Due to the depitching, a reduction in rotational losses is expected. In order to generate the required thrust, the depitched propeller operates at a slightly higher rotation rate. As illustrated in Figure 4.11a-4.11b, despite the slightly higher rotation rate, the depitching results in lower super-velocities at the blade tip. This is particularly visible at the leading edge. The term super-velocity is used in aerospace design for regions where the flow is accelerated to higher velocities than incoming flow or moving geometries. Consequently, the skin friction is also lower reducing viscous losses. The resulting open-water curves and the operating point for the propellers are shown in Figure 4.12.

Figures 5.17d-5.17f show that skew and rake distribution for both propeller are very close. The differences are expected to result from the (a) B-Spline curve fit to the coordinates and, (b) surface lofting operation through the various fit curves. By having stricter convergence criteria for the B-Spline fit and lofting, it may be possible to reduce the difference even further.

Note, as indicated in Section 4.3, the results can quickly be verified and validated by using open-water curves of the Wageningen B4-70 propeller, which are readily available [61]. The baseline geometry corresponds to the Wagenginen B4-70 propeller with P/D=1.0. Thus, at the operating point propellers with a P/D < 1.0 can be found to be more efficient. Similarly, DO finds a better propeller by reducing the effective P/D of the baseline geometry. It is recognized that the optimum in the sample set may not be the global optimum and that it is possible to choose a better starting P/D for the current operating point. Nonetheless, this is not expected to affect the comparison between orthogonal parameters and classical design variables.

DO PERFORMANCE RESULTS



(a) (b)

Figure 4.13: Dynamic Optimization performance results for (a) Classical design variables - Setup 1 and (b) Orthogonal parameters - Setup 2. $\lambda = (s_i/s)/100.0$.

Figure 4.13 illustrates the results for the design and optimization case with both classical and orthogonal design parameters. The x-axis shows the percentage of representative designs selected by the unsupervised classifier from each design cluster. The bars represent the ratio of the number of required iterations to the number of samples $N = 200$. The number at the head of each bar indicates the percentage of times the optimum design reported in Section 4.4.3 is found. The line indicates the mean performance increase of found solutions compared to the baseline design in percentage points.

It is observed that DO always finds designs that have an efficiency greater than the baseline Wageningen B4-70 propeller for both classical and orthogonal design variables within the Halton samples. DO performs better with orthogonal parameters. It tends to find the optimum design more often than its classical counterpart. Furthermore by evaluating under 30% ($\lambda = 20\%$) of the samples, solutions with an increase in efficiency of 3 percentage points are found. When ($\lambda > 30\%$), DO with orthogonal parameters finds the optimum in the Halton samples with a probability of 100%. On the other hand, DO with classical counterparts needs more than 50% of representative designs to be selected in order to find the optimum with 100% probability.

Figure 4.13 also contains information regarding convergence of DO. After λ is specified, the designs to evaluate from each cluster are determined at random. This makes DO for single objective optimization stochastic. Therefore, the optimization is repeated 50 times for each specified λ. From Figure 4.13 it can be inferred that the number of iterations to convergence depends on λ. Low values of λ result in quicker convergence, however, the probability of missing the lucrative cluster increases. For sufficiently high values of λ the number of iterations to convergence increases and the probability of finding the lucrative cluster and design also increases.

The results indicate that DO with orthogonal parameters is able to identify design clusters and also able to guide search to lucrative clusters more efficiently than when using classical design variables. This is expected to be an outcome of orthogonal parameters addressing the problem of multicollinearity. Furthermore, the orthogonal parameters enable more reliable machine learning models to be trained and deployed.

4.5. CONCLUSION

A new design and optimisation method - *Dynamic Optimization* - is proposed that synergises supervised and unsupervised learning for the efficient design and optimisation of propellers. Compared to a standard DoE based search, DO is almost 50 % more efficient in the current demonstration case. Furthermore, DO with orthogonal parameters tends to be more efficient than DO with classical design parameters. This gain in efficiency is attributed to the reliability of its two core components - (a) an orthogonal parametric model and, (b) classifiers. The orthogonal parameters are demonstrated to address the challenge of multicollinearity thus preserving the correlation between geometry, mesh and performance. They also enable more reliable classifiers to be trained and deployed. Reliable and explainable classifiers are able to guide search algorithms to lucrative regions in the design space with fewer design evaluations thus reducing the computational cost of optimization. However, it is noted that improving efficiency of design and optimisation is accompanied by a trade-off with completeness and optimality of search as demonstrated on a D&O problem of the benchmark Wageningen B4-70 propeller. Nonetheless, for the current thrust and ship speed constrained open-water efficiency objective, DO finds designs that have a higher efficiency than the Wageningen B4-70 propeller at a much lower cost than standard DoE based optimisation methods. Furthermore, the results can be quickly verified and validated with the open-water curves of Wageningen B4-70 propeller. As a next step, the method is to be demonstrated on cavitating propellers in a wake-field as a representative design problem in the following chapter.

4.6. RECOMMENDATIONS

- It is recommended that the trade-off study between completeness, optimality and computational cost also be repeated for coefficients of polynomials that define pitch distribution. The coefficients are not as numerous as classical design variables which could reduce computational cost of optimization. Furthermore, more insight can be gained into trade-off of computational cost of optimization with completeness and optimality for different parametric models.

- The orthogonal parameters are derived from SVD. It is recommended that the applicability of Gappy POD also be investigated in deriving orthogonal parameters. Gappy POD preserves smaller variations in geometry potentially leading to better correlations between orthogonal parameters and cavitation, flow separation or adverse pressure gradients on the blade.

- The orthogonal parametric model is able to identify lucrative regions in the design space however, they cannot be directly used to generate new geometries. In this case, classical design variables are still required to generate geometries. Thus, further investigation is recommended into using orthogonal parameters to generate propeller geometries.

- The study assumes that uncertainty in performance prediction is about 20 % of response η. While this is true for operational regimes near the upper and lower bounds of advance ratios, 20 % uncertainty may be too high for other operational regimes. Thus, a better model for uncertainty is also required. The model developed must depend on the propeller geometry, operational parameters and the analysis tool used to predict performance.

- Given that the Wageningen B4-70 propeller is (a) in the public domain, (b) its open-water performance is well documented and, (c) PROCAL is validated for its open-water curves, the geometry was chosen for demonstration. It is recommended that the study be repeated with other propeller families such as the Wageningen C-series as well as ducted propellers.

- This study investigates the use of orthogonal parameters with explainable machine learning models in propeller design and optimization. Thus, it is recommended that deep learning with orthogonal parameters also be investigated in propeller design and optimization.

- The study utilizes BEM to predict propeller performance. Higher fidelity simulations such as Reynolds Averaged Navier Stokes (RANS), Large Eddy Simulations

(LES) may lead to different or more detailed trends in predicted performance and also different levels of uncertainties. It is recommended that the impact on computational cost of optimization resulting from use of orthogonal parameters be investigated in cases where high fidelity simulations predict propeller performance.

ACKNOWLEDGEMENTS

The author extends his gratitude and acknowledges the valuable exchange of thoughts and ideas with Ir. Evert -Jan Foeth and Dr. Thomas Scholcz of MARIN, The Netherlands. The author is also grateful for CRS (Cooperative Research Ships) for supporting this research with PROCAL.

5

MULTIOBJECTIVE DYNAMIC OPTIMIZATION

An optimization method that uses a machine learning approach to solve multi-objective, constrained propeller optimization problems is proposed and analyzed. The method uses an online learning strategy where explainable supervised classifiers learn the location of the Pareto front and advise search strategies. The classifiers are trained with orthogonal features that capture geometric variation in radial distribution of pitch, skew, camber and chordlength. Based on orthogonal features, the classifiers predict whether or not a design lies on the Pareto front. If the design is predicted to lie on the Pareto front, the method verifies this with an evaluation. If the design is predicted to not lie on the Pareto front with a high confidence level, then the design is ignored. This skipped evaluation reduces the computational effort of optimization. The method is demonstrated on a cavitating, unsteady flow case of the Wageningen B-4 70 propeller with P/D=1.0 operating in the Seiun-Maru wake. Compared to the classical NSGA-III the optimization method is able to reduce 30 % of evaluations per generation while reproducing a comparable Pareto front. Trade-offs between suction side, pressure side, tip-vortex cavitation and efficiency are identified from the Pareto front. The non-elitist NSGA-III search algorithm in conjunction with the explainable supervised classifiers also find very diverse solutions. Among the solutions, a design with no pressure side cavitation, low suction side cavitation and reasonable tip-vortex cavitation is found[1].

5.1. INTRODUCTION

Global trade is expected to increase the traffic density of shipping lanes on seas, oceans, and also rivers. Consequently, the GHG, NOx, SOx, and noise emissions from shipping will increase making it more difficult to meet stricter emission regulations [106] and

---

[1] The research in this chapter has been approved by two of four reviewers of Applied Ocean Research. Two reviewers had very minor comments which have been incorporated and the paper has been resubmitted. Editor's decision is expected soon.

noise reduction guidelines [105]. The IMO aims to achieve a 50% reduction in GHG emissions by 2050 compared to 2008 levels [70] and ports have also started to incentivize silent ships with harbour due rate discounts [102]. Against this background, there is strong emphasis on making ship propellers more quiet and more efficient.

For vessels there are various sources of airborne and underwater noise [4]. Compared to air, noise travels faster and farther in water. Underwater noise is generated by water flow on vessels, auxiliary machinery and equipment, diesel generators, prime movers, electric motors and propellers; however, cavitating propellers are expected to be the dominant source of underwater noise. Averson et. al [8] report the underwater radiated noise for M/V Overseas Harriette at different ship speeds (propeller rpm). It was found that after cavitation inception, the noise hump between 50-100 Hz gets more pronounced with increasing speed. This noise hump is typically associated with the broad band spectrum of tip vortex noise [108] which is often the dominant source of underwater radiated noise. Thus, one opportunity to make propellers quieter is by reducing noise radiation of tip-vortex cavities.

To improve efficiency and reduce tip-vortex cavitation noise, an optimal blade loading is required. Blade loading plays an important role in propeller design as blade sections are continuously exposed to different angles of attack in the non-uniform wakefield they operate in. For a rigid propeller, the blade loading distribution is optimized by modifying the geometry. Traditionally, in addition to efficiency objectives, indirect noise constraints as a limit on cavitation volume or area have been implemented [64, 93]. Acoustic objectives have been considered but the focus has been limited to radiated sound power at blade passage frequencies [101]. The hydroacoustic design and optimization problem of propellers is of immense strategic importance and is most likely well researched in defence. However, there is limited knowledge and reporting in the public domain. Access to specialized simulation software, models based on propriety data and intellectual property restrictions (due to sensitivity or embargo) result in hydroacoutic optimization of propellers being tractable to very few academic research groups. In North America, Yin Lu Young's group at Michigan University focuses on flexible composite propellers [113] as a way towards designing quieter and more efficient propellers. One of the focus areas at the group is to design shape adaptive propellers, which have a bend twist coupling to optimize propeller loading [97, 100, 109] and improve efficiency. Given the challenging structural modelling for composites [98], cavitation and noise constraints are secondary considerations. In Europe, MARIN has focused on developing and demonstrating tools for hydroacoustic design and optimization of propellers [27, 95]. Rickard Bensow's group at Chalmers University has investigated algorithms, strategies and methods for hydrodynamic optimization of propellers with cavitation considerations [26, 84, 112]. In Asia, Nakashima propellers, for example, have focused on achieving better cavitation behaviour and improved efficiency for large vessels [110]. Better cavitation behaviour is expected to result in lower cavity volume fluctuations and thus lower pressure fluctuations and noise.

Up to approximately 2015, the focus in the aforementioned studies was predominantly on optimizing for efficiency and controlling sheet cavitation. In parallel, efforts were underway to better predict tip-vortex cavitation behaviour with simulations [29, 96, 104]. Thus, while propellers with better sheet cavitation behaviour were designed and introduced, the dominant source of noise was tending to be the cavitating tip-vortex. However, predicting tip-vortex cavitation on propellers continues to be too resource intensive for optimization. Recently, Bosschers .J [15] proposed an Emperical Tip Vortex (ETV) model which offers one approach to hydroacoutic optimization of propellers with tip-vortex considerations. In addition to noise and efficiency design objectives, propeller design must also account for constraints on hull excitation [88], cavitation erosion [111], ship speed and thrust. State of the art optimization methods are able to tackle this multi-objective constrained problem; however, there are three challenges which make them computationally expensive [91]:

The *first* challenge is that most optimization methods tend to approach the optimization problem with classical design variables. However, they could lead to multicollinearity [2] which makes it difficult to isolate the impact of changing individual design variables on objectives and constraints. By solving multicollinearity, there is an opportunity to reduce computational cost of optimization by reducing the number of objective/constraint evaluations to isolate the sensitivities of individual classical design variables. Consider a function

$$\phi = f(\vec{M}(\vec{s})), \tag{1}$$

where $\vec{M}$ represents the mesh and $\vec{s}$ is the design vector. By chain rule,

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial \vec{M}} \cdot \frac{\partial \vec{M}}{\partial \vec{s}} = J_1 \cdot J_2, \tag{2}$$

where J_1 is the Jacobian of the function with respect to the mesh and J_2 is the Jacobian of the mesh with respect to the design variables. The optimization problem is free of multicollinearity if both J_1 and J_2 are orthonormal matrices. The orthonormality of J_1 is determined by the flow physics and J_2 by the chosen design variables. Any design variable which affects multiple nodes on the mesh will lead to J_2 which is not orthonormal. This is because the blade surface is often defined with splines and changes in mesh nodes are strongly correlated. Different classical design variables also tend to change the same mesh nodes. Thus, it is difficult to isolate the impact of individual design variables on objectives and constraints.

The *second* challenge is that in the preliminary design phase, BEM and ETV model tend to have uncertainties [83]. These uncertainties results from numerical methods, domain discretization and physical models. In specific regions in the design space, these uncertainties could lead to the similar efficiency or cavitation performance being predicted by BEM and ETV for very different designs. Consequently, optimization methods could converge to solutions whose operational performance is significantly different from predicted performance.

The *third* challenge is that the computational cost of optimization for realistic cases is relatively high even in the preliminary design phase. Most commonly Artificial Neural

Networks, Krigging, iKrigging, Cascading Neural Networks [112] have been used to reduce the cost of optimization. More recently Deep Learning [94] has also gained traction in propeller design and optimization [99]. In addition, explainable machine learning approaches [3] presents very interesting opportunities to provide or extrapolate performance guarantees, which is not possible with other approaches for propeller manufacturers. These developments are relatively new in the maritime industry. However, the models used in explainable machine learning such as Gaussian mixture models and Naive-Bayes are fundamental methods which are well known. Explainable machine learning can be used to predict where lucrative designs may be found in the design space *and* also provide insight into *why* the regions are lucrative.

To address the *first* challenge of multicollinearity, an orthogonal feature set is used. The orthogonal features have been demonstrated to capture geometric variations more accurately than classical design variables [91]. To address the *second* challenge and to account for uncertainties in preliminary design phases, performance predictions are assumed to have a mean and variance. Solutions dominate, i.e., lie on the Pareto front only when the mean is high and there is no overlap in the 95 % confidence intervals of performance with compared solutions. This reduces the risk of false positives on whether a solution dominates or not. Consequently, D&O strategies are expected to yield a range of solutions whose predicted performance does not differ significantly from operational performance. To address the *third* challenge of computational cost, soft explainable supervised classifiers with online training are used. The classifiers are taught the location of the Pareto front and, when sufficiently trained, they could potentially exclude designs which lie meaningfully away from the Pareto front, making the optimization method efficient.

In this paper, an optimization method is proposed to solve multi-objective constrained optimization problems based on the above machine learning strategies. The method referred to as DO is detailed in Section 5.2. It is demonstrated on a case with the starting point as a Wageningen B-4 70 propeller with P/D=1.0 operating in the Seiun-Maru wakefield. Subsections 5.2.1-5.2.6 detail the modules of DO. Section 5.3 compares the performance of classical NSGA-III optimization with DO. the results are summarized in Section 5.4.

5.2. DYNAMIC OPTIMIZATION

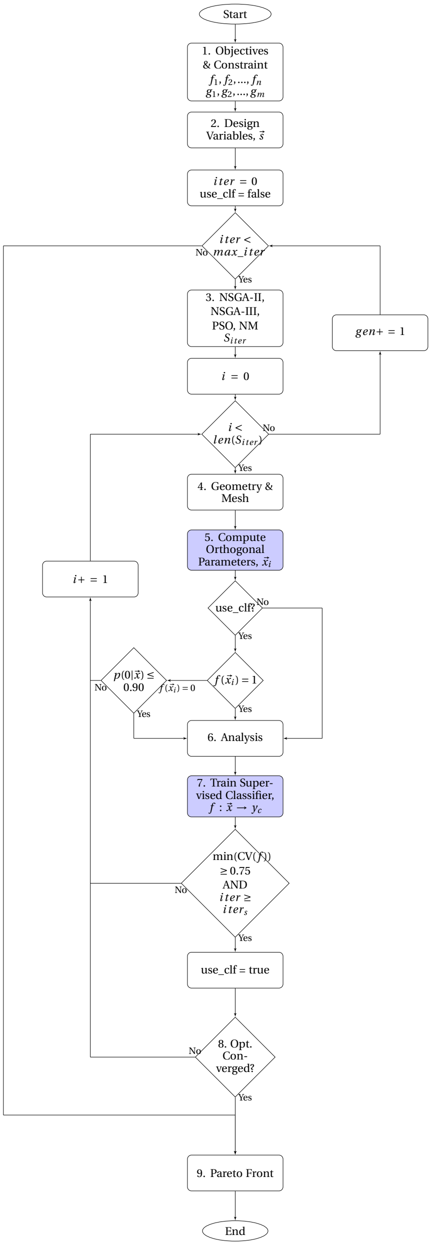


Figure 5.1: Dynamic Optimization flowchart.

Dynamic Optimization is a method for multiobjective, constrained propeller optimization problems where search agents are advised by supervised classifiers. The method is detailed in Algorithm 3 and the flowchart in Figure 5.1. As inputs, the method takes the objectives, constraints and iteration limits. When the optimization begins, for the i-th iteration, an objective driven search strategy such as NSGA-II, NSGA-III [90], PSO, Nelder-Mead (NM) or others generate samples S_i. For each sample $\vec{s} \in S_i$, the geometries and meshes are generated. A set of orthogonal features $\vec{x}^* \in \Re^{k'}$ which are projections of the i-th mesh in a k' dimensional hyperspace is computed. Until there are sufficient samples to train and validate, the classifier is not used. When the classifier is not used, designs $\vec{s} \in S_i$ are analysed with CFD or BEM. After each design is analysed they are assigned a label. Interesting designs are labelled as $y_c = 1$ and clearly inferior designs are labelled $y_c = 0$. If the design lies on the Pareto front and has an efficiency greater than $a \times \eta_{max}$, where $a = 0.85$ is an input parameter. With a, it is possible to control the explore/exploit bias. A higher value of a makes the search more exploitative. For details on assigning labels see Algorithm 6.

After the first two samples are evaluated, the supervised classifier is trained with orthogonal parameters ($\vec{x}$) as features and target labels ($y_c \in [0,1]$). After at least 10 samples are evaluated, it is possible to validate the classifier with 10-fold cross-validation studies. In such a case, if the classifier has a mean 10-fold cross-validation score of 75% or higher the classifier is used. Weighted average precision scores are used for cross-validation. In iterations where the classifier is used, design labelled as $y_c = 1$ are analysed to confirm the prediction. Else, the design is not analysed when $p(\vec{x}|\mu_0, \sigma_0^2) \geq p_{th}(\vec{x}|\mu_0, \sigma_0^2)$, where $p_{th}(\vec{x}|\mu_0, \sigma_0^2)$ is the threshold for the probability that $y_c = 0$. A high threshold indicates a high confidence that the design is much further away from the Pareto front. Thus, by skipping analysis of designs which lie away from the Pareto front, computational cost of optimization is reduced. In our demonstration $p_{th}(\vec{x}|\mu_0, \sigma_0^2) = 0.9$. The method catches most false positives. However, there is a risk that false negatives are missed. This risk can be mitigated by increasing $p_{th}(\vec{x}|\mu_0, \sigma_0^2)$. If after an iteration or generation the convergence criteria are met, the optimizer reports the Pareto front and exits. Else, it continues to the next iteration or generation.

5.2.1. OBJECTIVE AND CONSTRAINTS

For demonstration of DO, the Wageningen B-4 70 propeller with P/D=1.0 is chosen as the starting point. The blade is operating in the effective wake of the Seiun-Maru vessel with $v_{ship,0} = 10.80 m/s$ and $T_0 = 577.4 kN$ [38]. The full scale measurements for flow characteristics around the hull and resulting propeller inflow and turbulence are reported by ITTC [92]. Particulars of the Seiun-Maru vessel are also provided in [114].

The optimization problem has multiple objectives and constraints. Table 5.1 lists and describes the objectives and constraints. The goal is to maximize efficiency and minimize (a) cavity volume, (b) cavity volume acceleration, (c) vortex cavity radius and (d) peak pressure. The efficiency and cavitation behaviour are predicted with PROCAL [82]. The radius of the vortex cavity is predicted with the ETV-2 model [16]. In addition

Algorithm 3 Dynamic Optimization for multi objective applications

```
Input: $f_1, f_2, ..., f_n,\ g_1, g_2, ..., g_m$, max_iter, thresh_iter
Output: $S_{pareto}$
procedure DYNAMIC OPTIMIZATION
    $S_{pareto}, S, F \leftarrow \emptyset$
    $i \leftarrow 0$
    use_clf ← false
    while $i \le$ max_iter do
        $S_i \leftarrow$ SEARCH(*params)
        $S \leftarrow S \cup S_i$
        for $j \leftarrow (0 \rightarrow len(S_i))$ do
            $\vec{s} \leftarrow S_{i,j}$
            Generate geometry and mesh $(\vec{M}^*)$ for $\vec{s}$
            $\vec{x}^* \leftarrow$ PROJECTION$(\mathfrak{R}_{ortho}, k', \vec{M}^*)$        ▷ See Algorithm 5
            if use_clf then
                if $f(\vec{x}^*) = 1$ AND $p(\vec{x}^*|\mu_0, \sigma_0^2) \le 0.90$ then        ▷ The classifier $f : \mathfrak{R}^{k'}_{ortho} \rightarrow \mathfrak{R}$
                    Analyse design $\vec{s}$
                    $F \leftarrow F \cup \left[\left[\begin{matrix} f_1 \\ \vdots \\ f_n \end{matrix}\right]\right]$
                else
                    continue        ▷ Skipped an evaluation, thus reduced cost of optimization
            else
                Analyse design $\vec{s}$
                $F \leftarrow F \cup \left[\left[\begin{matrix} f_1 \\ \vdots \\ f_n \end{matrix}\right]\right]$
            Train Supervised Classifier, $f : \mathfrak{R}^{k'}_{ortho} \rightarrow \mathfrak{R}$
            if min(10-k Cross Validation) $\le 0.75$ AND $i \ge$ thresh_iter then
                use_clf ← true
            else
                use_clf ← false
            if Converged then
                break
    $\vec{y}_c \leftarrow LABELS(F, \epsilon)$        ▷ See Algorithm 6
    $S_{pareto} \leftarrow S_{pareto} \cup \{s_i | S[\vec{y}_c == 1]\}$
    return $S_{pareto}$
```

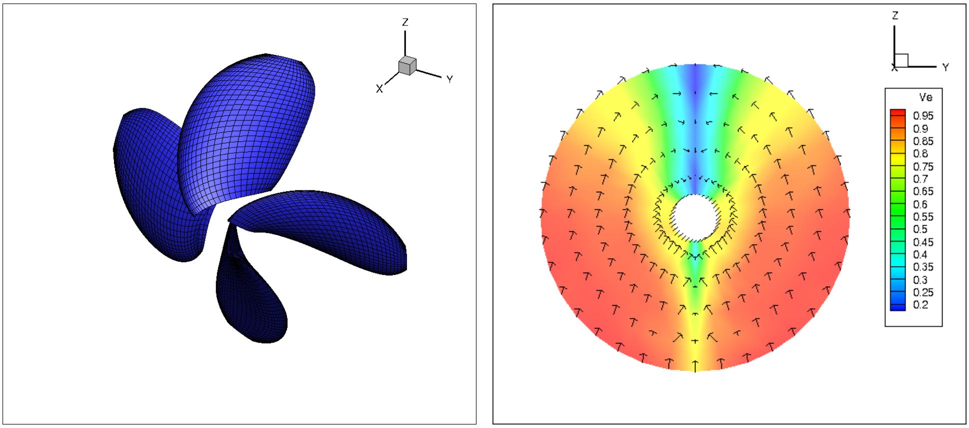
Table 5.1 Objectives and Constraints.

Obj./Cons.	Description
$\max f_1(\vec{s}) = \eta$	Maximize efficiency
$\min f_2(\vec{s}) = V_{cav,b}$	Minimize suction surface cavity
$\min f_2(\vec{s}) = V_{cav,f}$	Minimize pressure surface cavity
$\min f_4(\vec{s}) = \frac{\partial^2 V_{cav}}{\partial^2 t}$	Minimize cavity volume acceleration
$\min f_5(\vec{s}) = \max(r_c)$	Minimize tip vortex cavity radius
$\min f_6(\vec{s}) = \max(\lvert C_p \rvert) = \max\left(\left\lvert \frac{p - p_{r,\infty}}{\frac{1}{2}\rho_\infty\left(v_{ship}^2 + 2\pi\omega r^2\right)} \right\rvert\right)$	Minimise peak pressures. ω is the propeller rotation rate in revolutions per second. $p_\infty = 1.01325e5Pa$, $\rho_\infty = 1025kgm^{-3}$.
$g_1(\vec{s}) = T - T_0 = 0$	Thrust constraint. Note that this is not changed during optimization.
$g_2(\vec{s}) = v_{ship} - v_{ship,0} = 0$	Ship speed constraint. Note that this is not changed during optimization.
$g_3(\vec{s}) = \max\left(\left\lVert \frac{s_{\theta,i} - s_{\theta,i+1}}{s_{\theta,i+1}} \right\rVert\right) - 0.3 \leq 0$	Fairing constraint for pitch
$g_4(\vec{s}) = \max\left(\left\lVert \frac{s_{\gamma,i} - s_{\gamma,i+1}}{s_{\gamma,i+1}} \right\rVert\right) - 0.3 \leq 0$	Fairing constraint for skew
$g_5(\vec{s}) = \max\left(\left\lVert \frac{s_{chord,i} - s_{chord,i+1}}{s_{chord,i+1}} \right\rVert\right) - 0.3 \leq 0$	Fairing constraint for chord-length
$g_6(\vec{s}) = \max\left(\left\lVert \frac{s_{cam,i} - s_{cam,i+1}}{s_{cam,i+1}} \right\rVert\right) - 0.3 \leq 0$	Fairing constraint for camber

to thrust and ship-speed constraints, fairing constraints are specified to guide search towards solutions without high geometric variations along the radius. In order to match the thrust requirement, the propeller rotation rate is modified while the ship-speed is maintained a constant.

It is recognized that the skewed Seiun-Maru propeller is a good starting point for the demonstration. In the scope of this research, to demonstrate DO in an explore biased optimization, we prefer to start from a design much further away from a known local optimum. In the current demonstration, the baseline propeller is the Wageningen B-4

70 propeller with P/D=1.0. The baseline performance is detailed in Section 5.3.1. This starting point implies a more challenging task for the optimization algorithm. At the same time, it increases the chances of finding new design solutions and search directions, compared to starting from an already more optimal design. The starting propeller geometry, effective wake and operating conditions are illustrated in Figure 5.2.



(a)

(b)

Figure 5.2: (a) Wageningen B4-70 propeller base geometry and (b) Effective wakefield of Seiun-Maru vessel, $V_e = \sqrt{V_x^2 + V_y^2 + V_z^2}$. *The wakefield was provided by MARIN for this research and is available on request.*

5.2.2. Design Variables

Table 5.2 Bounds of design variables.

Variable	Lower Bound	Upper Bound
$\Delta\theta$ (rad)	-0.36	036
$\Delta\gamma$ (rad)	-0.063	0.063
s_{camber} (rad)	-0.045	0.045
s_{chord} (-)	0.76	1.24
n (rps)	0	inf

The blade is parametrised with four equispaced hydrofoil cross-sections at locations $[0, \frac{span}{3}, \frac{2span}{3}, span]$. Four parameter $[\theta, \gamma, s_{chord}, s_{camber}]$ are used as design variables to modify the cross-sections. They are intended to change pitch, skew, chordlength and camber respectively. While θ, γ are modified for all four cross-sections, s_{camber}, s_{chord} are modified for the last three cross-sections. The modifications are in the sequence θ, γ, s_{chord} and s_{camber}. The fifth design variable is the propeller rotation rate $n(rps)$.

The baseline geometry and the location of the hydrofoils is illustrated in Figure 5.3. The bounds of the 14 design variables are listed in Table 5.2. The bounds result from the

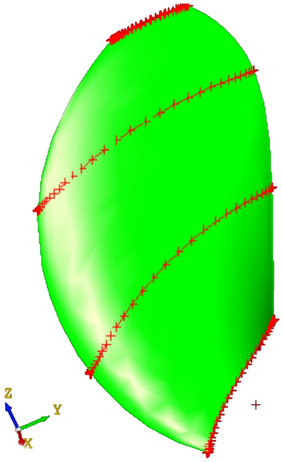
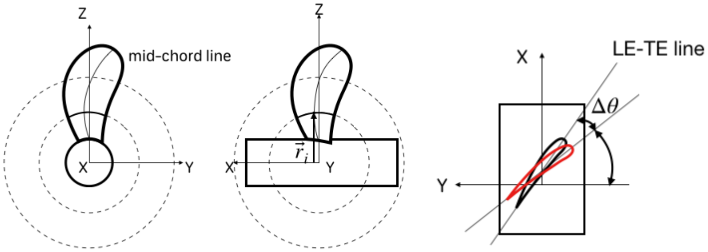


Figure 5.3: Blade hydrofoils

applicability range for PROCAL. The found solution lies away from the bounds (see Table 5.5) thus, it is demonstrated that the applicability bounds of PROCAL are not restrictive.

The schemes for geometry modification are detailed below.



(a) (b)

Figure 5.4: (a) $\vec{r}_i$, (b) change in geometry resulting from $\Delta\theta$.

Parameter θ: Figure 5.4b illustrates the change in geometry when θ is varied. The target cross-section is rotated about a vector parallel to z-axis passing through the midchord of the section. An increase in θ results in the blade experiencing a higher inflow angle thus also higher loading. Pitch is expected to be very sensitive to θ. It is also possible that θ influences the camber of the hydrofoil.

Parameter γ: Figure 5.5 illustrates the change in geometry when γ is varied. The target cross-section is rotated at the mid-chord vector about the x-axis. Skew is expected to be very sensitive to γ. The proposed parameter helps control the rate at which loading on

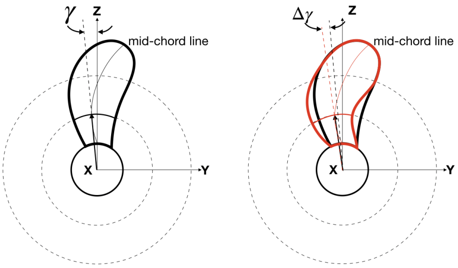


Figure 5.5: Change in geometry resulting from $\Delta\gamma$.

the blade increases in the wake-field. By doing so, it helps regulate cavity volume acceleration.

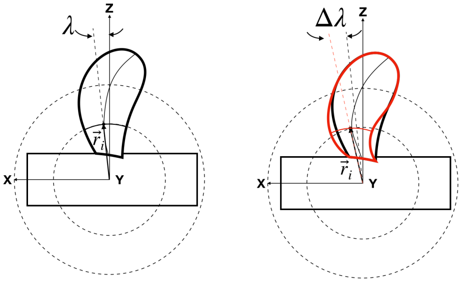


Figure 5.6: Change in geometry resulting from $\Delta\lambda$.

Parameter λ: Figure 5.6 illustrates the change in geometry when λ is varied. The target cross-section is modified by rotating $\vec{r}_i$ about the y-axis. Rake is expected to be very sensitive to λ.

Camber: The schematic for modifying hydrofoil camber is illustrated in Figure 5.7. Camber is modified by rotating each of the coordinates defining the hydrofoil about the x-axis. The angle of rotation is

$$\phi = \frac{\phi'^2 sin(\phi') s_{camber,i}}{c}, \tag{3}$$

where $c = 3.9453$ is a smoothing constant that can be defined by the designer and $\phi' = \frac{x - x_{te}}{x_{le} - x_{te}} \cdot \pi$. Note that the camber modification does not influence the leading edge or trailing edge thus, pitch is not expected to be impacted by changes in s_{camber}. The coordinates of the cross-section are transformed as

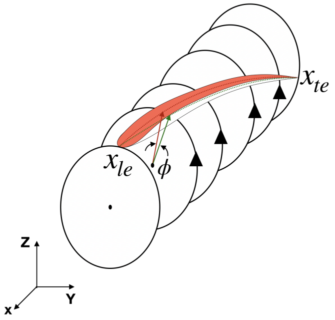
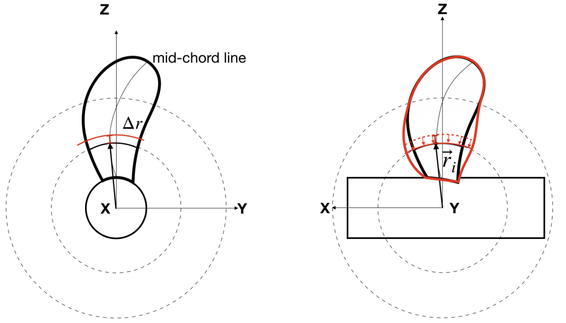


Figure 5.7: Schemetic illustrating camber modification.

$$\vec{p}' = \begin{bmatrix} p_x & p_y & p_z \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\phi & -sin\phi \\ 0 & sin\phi & cos\phi \end{bmatrix}. \tag{4}$$

Increasing camber for the cross-section is expected to increase loading while also shifting the centre of pressure for the blade. For flexible propellers, increased loading could introduce bending forces on the blade. Furthermore, the shift in centre of pressure could lead to a twisting moment about the inertial axis.



(a) Step 1 - Scaling hydrofoil along the mid-chord vector

(b) Step 2 - Translating back to initial radial location

Figure 5.8: Schemetic illustrating chordlength modification.

Chord-length: The schematic for modifying hydrofoil chord-length is illustrated in Figure 5.8. Chord-length is modified in two steps. The first step, illustrated in Figure 5.8a, involves scaling the hydrofoil along the mid-chord vector by s_{chord}. The coordinates of the cross-section are transformed as

$$\vec{p}' = \begin{bmatrix} p_x & p_y & p_z \end{bmatrix} \begin{bmatrix} s_{chord} & 0 & 0 \\ 0 & s_{chord} & 0 \\ 0 & 0 & s_{chord} \end{bmatrix}. \tag{5}$$

The second step, illustrated in Figure 5.8b, involves translating it back to the initial radial location as

$$\vec{p}' = \begin{bmatrix} 0 & 0 & 0 & \delta x \\ 0 & 0 & 0 & \delta y \\ 0 & 0 & 0 & \delta z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \\ 1 \end{bmatrix}, \tag{6}$$

where $[\delta x, \delta y, \delta z] = \vec{r} - \vec{r}_{scaled}$.

One limitation of the above proposed parameters is that the pitch, skew and rake are indirectly influenced. Furthermore, the sequence in which the parameters are varied could determine whether skew and rake are preserved or not. It is also possible that modifying θ could impact the camber. However, it is to be noted that even with very carefully considered and more robust alternatives, J_2 in Equation 2 will not be orthonormal if varying one parameter influences multiple mesh nodes on the discretized blade geometry. Consequently the parametrisation is not orthogonally independent. This problem will be tackled in Section 5.2.5, where a set of orthogonal parameters are derived.

5.2.3. NSGA-III

For the current multi-objective constrained optimization problem NSGA-III is chosen, as it is an established method on 3 to 15 objective optimization problems with convex, concave, disjointed and differently scaled Pareto fronts [90]. The framework of NSGA-III is similar to that of NSGA-II with modifications to mating selection and survival [87].

Algorithm 4 details the selection of the survivors set (S) for NSGA-III. At the t-th generation, the selection procedure takes as inputs (a) $R^t = P^t \cup Q^t$, where P^t is the parent population, Q^t is the offspring population, (b) the population size N, (c) the reference directions Z and, (d) $\hat{z}^*$ and $\hat{z}^{nadir}$ which are bounds to normalise individuals in the selection set S. The output is $P^{(t+1)} = S$. The first step involves non-dominated sorting of merged population R^t into hierarchical fronts $(F_1, F_2, ...)$. The fronts are sequentially appended to the initially null set S until $|S| + |F_i| \leq N$. When $|S| + |F_L| > N$, the splitting front F_L is identified. Individuals in S and F_L are then normalised by using $\hat{z}^*$ and the nadir point estimation $\hat{z}^{nadir}$ as lower and upper bounds respectively. The individuals in $\bar{S}, \bar{F_L}$ are assigned to a reference direction π_k with the shortest perpendicular distance d_k. The reference directions are generated based on the Riesz s-energy concept to obtain a well-spaced distribution over the optimal Pareto front [11]. The niching method selects $N - |S|$ individuals based on niche counts (ρ), π_k and d_k and appends them to S.

Algorithm 4 NSGA-III survival selection

```
  Input: $R^t, N, Z, \hat{z}^*, \hat{z}^{nadir}$
  Output: $P^{(t+1)}$
procedure SURVIVORS
    $(F_1, F_2, ...) \leftarrow non\_dominated\_sort(R^t)$
    $S = \emptyset, i = 1$
    while $|S| + |F_i| \leq N$ do
        $S \leftarrow S \cup F_i$
        $i = i + 1$
    $F_L \leftarrow F_i$
    if $|S| + |F_L| > N$ then
        $\bar{S}, \bar{F_L}, \hat{z}^*, \hat{z}^{nadir} \leftarrow normalize(S, F_L, \hat{z}^*, \hat{z}^{nadir})$
        $\rho, \pi, d \leftarrow 0$
        for $k \leftarrow 1$ to $|S|$ do
            $\pi_k, d_k \leftarrow associate(\bar{S}_k, Z)$
            $\rho_{\pi_k} \leftarrow \rho_{\pi_k} + 1$
        $S \leftarrow S \cup niching(\bar{F}_L, n - |S|, \rho, \pi, d)$
    $P^{(t+1)} \leftarrow S$
    return $P^{(t+1)}$
```

5.2.4. GEOMETRY & MESH

The blade geometry is defined by a single B-Spline surface defined as

$$\vec{p}(u,v) = \sum_{i=0}^{m} \sum_{j=0}^{n} N_{i,d_1}(u) N_{j,d_2}(v) \vec{P}_{i,j}, \tag{7}$$

where $u, v \in [0.0, 1.0]$ are knots in the u and v directions; $\vec{P}_{i,j}$ is the control point of the i-th row and j-th column; $m+1$ and $n+1$ are the number of rows and columns the control points are organized in; d_1, d_2 are the degree of the B-Splines in u and v direction respectively and; N_{i,d_1} and N_{j,d_2} are the basis functions in the u and v direction respectively. The surface is fitted to hydrofoil cross-sections in Cartesian coordinates at different radial locations. In our case $d_1 = 3$ as the spline is closed in u direction and $d_2 = 8$ to find an optimal fit for a hydrofoil with camber and thickness distribution with fewer control points. The basis functions for each directions are defined as

$$N_{i,0}(t) = \begin{cases} 1; t_i \leq t < t_{i+1} \\ 0; otherwise \end{cases}, \tag{8}$$

$$N_{i,d} = \frac{t - t_i}{t_{i+d} - t_i} N_{i,d-1}(t) + \frac{t_{i+d+1} - t}{t_{i+d+1} - t_{i+1}} N_{i+1,d-1}(t). \tag{9}$$

where t is the knot value and d is the degree of the spline. The surface mesh is generated by uniformly discretising the B-Spline surface in the B-Spline parametric space. The discrete nodes in the B-Spline parametric space are

$$M^*_{m\times n} = \begin{bmatrix} (0.0,0.0) & \cdots & \left(0.0,\dfrac{(j-1)}{(n-1)}\right) & \cdots & (0.0,1.0) \\ \vdots & \ddots & & & \vdots \\ \left(\dfrac{(i-1)}{(m-1)},0.0\right) & & \left(\dfrac{(i-1)}{(m-1)},\dfrac{(j-1)}{(n-1)}\right) & & \vdots \\ \vdots & & & \ddots & \vdots \\ (1.0,0.0) & \cdots & \cdots & \cdots & (1.0,1.0) \end{bmatrix}, \tag{10}$$

where m is the number of nodes on the hydrofoil cross-section and n is the number of cross-sections along the radius. The nodes correspond to the mesh for the i-th propeller $\vec{M}_i \in \mathfrak{R}^{3\times m\times n}_{cart}$

$$\vec{M}_i = \{p(m^*_{ij})|m^*_{ij} \in M^*_{m\times n}\}. \tag{11}$$

5.2.5. ORTHOGONAL PARAMETERS

The primary objective of the proposed parametric model is to alleviate the limitations of classical design variables. Thus, the parametric model must (a) accurately quantify geometric variations, (b) yield orthogonally independent parameters and, (c) be generalized for all possible geometries. To satisfy the above requirements, the parametrisation is based on the coordinates of the blades' surface mesh such that all possible variations of all possible families are represented and the relation between geometry and design objectives is not obscured. The parameters of the proposed model are derived by projecting the surface mesh of propellers in an orthonormal hyperspace $\mathfrak{R}_{orth}$ to ensure that the parameters are orthogonally independent. An orthonormal hyperspace $\mathfrak{R}_{orth}$ is defined with a geometry variance study where pitch, skew, rake, chord-length and camber are individually varied within the bounds listed in Table 5.3. Chord-length and camber are varied for the last three hydrofoils. 100 equispaced samples are taken within the bounds for each variable.

Table 5.3 Bounds for variance study.

Variable	Lower Bound	Upper Bound
$\Delta\theta$ (rad)	-0.6	0.6
$\Delta\gamma$ (rad)	-0.2	0.2
s_{camber} (rad)	-0.10	0.10
s_{chord} (-)	0.60	1.40
$\Delta\lambda$ (rad)	-0.10	0.10

Our data matrix M is stacked with meshes of all propellers in the variance study and then transformed into cylindrical coordinates:

$$\tilde{M} = \begin{bmatrix} \vec{M}_1 \\ \vdots \\ \vec{M}_N \end{bmatrix}, \tag{12}$$

$$M = T(\tilde{M}),\, T : \mathfrak{R}^{N\times k}_{cart} \rightarrow \mathfrak{R}^{N\times k}_{cyl}. \tag{13}$$

Algorithm 5 Subspace Projection

```
  Input: $\mathfrak{R}_{ortho}, k', \vec{M}^*$
  Output: $\vec{x}$
procedure PROJECTION
    $\mathfrak{R} \leftarrow \mathfrak{R}_{ortho}$
    $\vec{x}^* \leftarrow 0_{k'+1}$
    $\vec{x}^*_{k'+1} \leftarrow |\vec{M}^*|$
    for $j = 1$ to $k'$ do
        $\vec{x}^*_j \leftarrow cos^{-1}\left(\frac{\vec{M}^* \cdot \mathfrak{R}_{:,j}}{|\vec{M}^*||\mathfrak{R}_{:,j}|}\right)$
    return $\vec{x}^*$
```

The origin of $\mathfrak{R}_{orth}$ is chosen to be the mean of all propeller meshes,

$$\vec{\mu} = \frac{\sum_{i=1}^{N} \vec{M}_i}{N}, \tag{14}$$

thus, the translated data matrix is

$$M' = \begin{bmatrix} \vec{M}_1 - \vec{\mu} \\ \vdots \\ \vec{M}_N - \vec{\mu} \end{bmatrix}. \tag{15}$$

M' is factorized using thin SVD

$$M'_{k\times N} = U_{k\times N} \Sigma_{N\times N} V'_{N\times N}. \tag{16}$$

In this worked example, the orthonormal reference frame is defined as the product of $k' = 10$ right and left singular vectors.

$$\mathfrak{R}_{orth} = U_{k\times k'} \cdot V'_{k'\times k'}. \tag{17}$$

Algorithm 5 details the procedure to project $\vec{M}^*$ sampled by NSGA-III in $\mathfrak{R}_{ortho}$ yielding the orthogonal parameters $\vec{x}^*$. For the orthogonal parameters, $\frac{\partial \vec{M}}{\partial \vec{x}}$ is an orthonormal matrix. This is expected to solve the problem of multicollinearity.

5.2.6. NAIVE-BAYES CLASSIFIER

Explainable supervised classifier are used to guide search towards more lucrative regions in the design space. Naive-Bayes classifier was demonstrated to work reliably [91] and is also considered in this demonstration. The Naive-Bayes classifier predicts labels as

$$y_c = \text{argmax}_j \; \pi_j p(\vec{x}_i | \vec{\mu}_j, \vec{\sigma^2}_j), \tag{18}$$

where j maximizes the probability of $\vec{x}_i$ occurring. The probability of $\vec{x}_i$ occurring given the mean and variance $\vec{\mu}_j, \vec{\sigma^2}_j$ for the j-th cluster is

$$p(\vec{x}_i|\mu_j,\sigma_j^2) = \prod_{d=1}^{k'} \frac{1}{\sqrt{2\pi\sigma^2_{j,d}}}.\exp\left(-\frac{|x_{i,d}-\mu_{j,d}|^2}{2\sigma^2_{j,d}}\right). \tag{19}$$

Where k' is the number of orthogonal parameters, $x_{i,d}$ is the d-th parameter in $\vec{x}_i$, $\mu_{j,d}$ and $\sigma^2{}_{j,d}$ are the mean and variance of $X_{:,d}$. Implementations in SciKit-Learn [107] are used.

Algorithm 6 Assign Labels

```
    Input: $F, \epsilon$
    Output: $\vec{y}_c$
  procedure LABELS
      $N \leftarrow len(F)$
      $\vec{y}_c \leftarrow [0, ..., N-1]$
      $next\_point \leftarrow 0$
      $\sigma \leftarrow \epsilon F$
      $F_{low} \leftarrow F - 1.96\frac{\sigma}{\sqrt{N}}$          ▷ Lower bound of 95%confidence interval
      $F_{high} \leftarrow F + 1.96\frac{\sigma}{\sqrt{N}}$         ▷ Upper bound of 95%confidence interval
      while $next\_point < N$ do
          $mask \leftarrow \{y|F_{high}[i] < F_{low}[next\_point]\ \forall i \in [0, N-1]\}$
          $mask[next\_point] \leftarrow 1$
          $\vec{y}_c \leftarrow \vec{y}_c[mask]$
          $F \leftarrow F[mask]$
          $F_{low} \leftarrow F_{low}[mask]$
          $F_{high} \leftarrow F_{high}[mask]$
          $next\_point \leftarrow \sum_{i=0}^{next\_point-1} mask[i] + 1$
      $\vec{y}_c \leftarrow \{y|\vec{y}_c[i] = 1$ AND $F_i \geq 0.85\eta_{max}\ \forall i \in [0, N-1\}$
      return $\vec{y}_c$
```

For the worked optimization problem in this paper, it is assumed that the clusters are fundamentally of two types (a) interesting designs and (b) un-interesting designs. The former lies on the Pareto front and the latter either lies meaningfully away and/or cannot be evaluated due to limitations of the numerical method. The expectation for propeller design is that the current hypothesis of the two fundamental cluster types holds for most cases. It is, however, recognized that propeller design is a complex problem. It is quite possible that the set of interesting designs will be spread out over multiple clusters with additional constraints/requirements which define an acceptable design. In such cases, One-vs-Rest (OvR) multiclass strategies [103] must be employed to have reliable classifier performance.

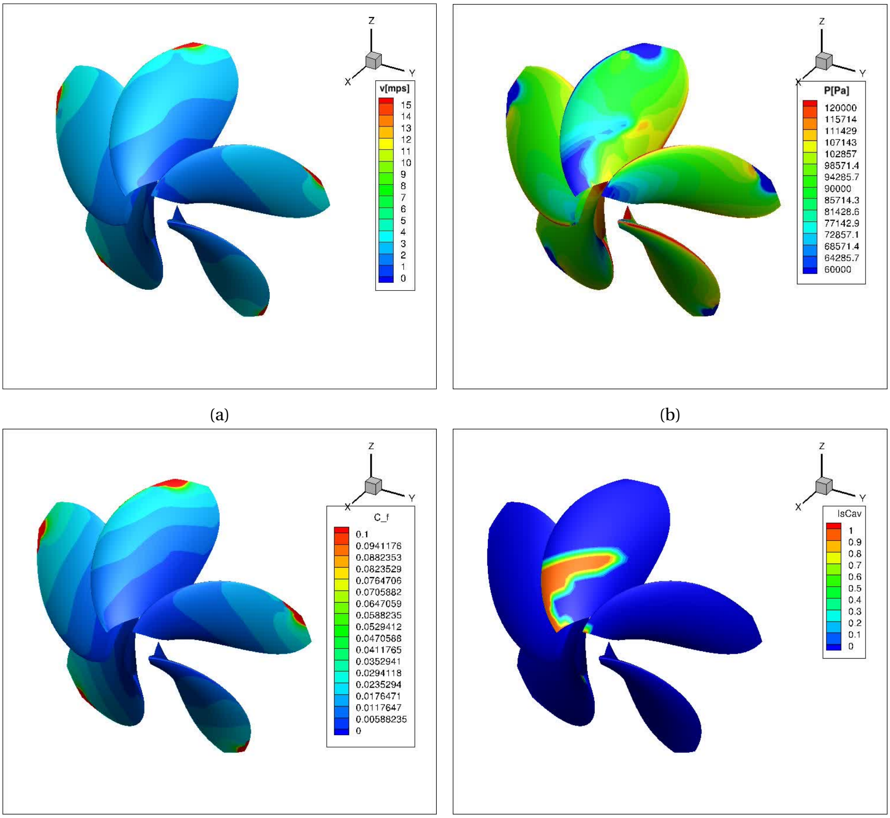
The labels $y_c \in [0,1]$ are determined by whether the designs lie on the Pareto front and whether they have a high efficiency. To account for uncertainties, it is assumed that designs have a mean and variance in performance. A design dominates when the mean of the predicted performance is higher and there is no overlap in the 95 % confidence interval with the compared solution. The 95 % confidence interval is based on the assumption that there is a 15 % uncertainty in predicted performance [83]. Algorithm 6 detail the procedure to assign labels. As inputs it takes objective function values and assumed uncertainty (ϵ). It returns the labels for all sampled points.

5.3. RESULTS

The Pareto fronts resulting from classical NSGA-III and DO optimization are compared. The trade-offs in objectives as observed from the Pareto front are detailed. A solution with minimal suction surface cavitation, no pressure surface cavitation and moderate tip-vortex cavitation is identified and described.

5.3.1. BASELINE PERFORMANCE

The performance of the baseline Wageningen B-4 70 P/D=1.0 propeller operating in the wake of Seiun-Maru is estimated by PROCAL. For this worked example, the baseline performance prediction has local discontinuities in the pressure and velocity fields at the blade tip as illustrated in Figure 5.9. Thus, it is difficult to interpret the results to identify sources of energy losses and cavitation behaviour. These discontinuities are expected to result from panels being relatively more skewed at the blade tip. Consequently, DO is also tasked with identifying designs which can be analysed reliably by PROCAL.

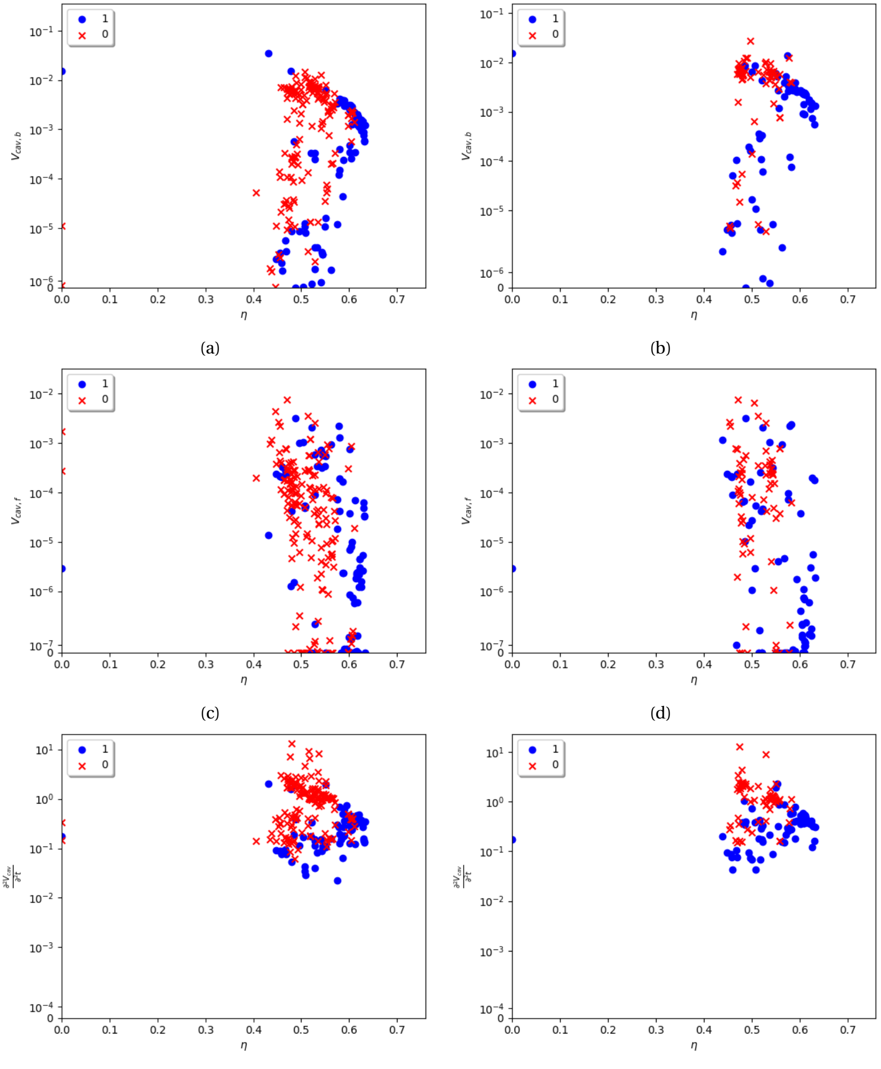


(a) (b)

(c) (d)

Figure 5.9: B4-70 propeller (a) inviscid velocity distribution, (b) pressure field on blade, (c) skin friction coefficient and, (d) fraction of cavitating vertices on mesh panel. $|\vec{v}| = \sqrt{v_x^2 + v_y^2 + v_z^2}$ and IsCav $\in [0.00, 0.25, 0.50, 0.75, 1.00]$.

5.3.2. PARETO FRONT



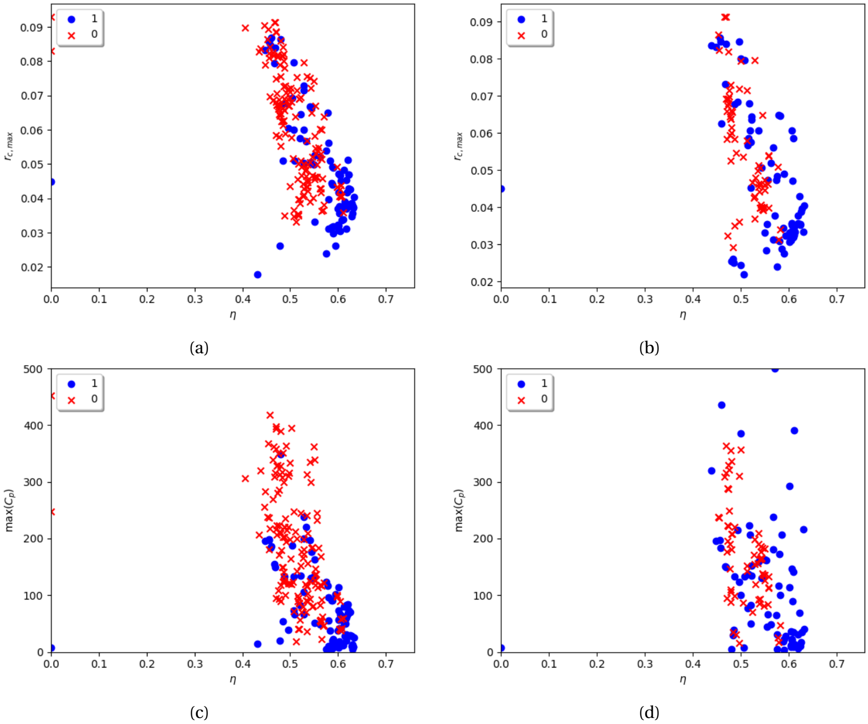
(a) (b)

(c) (d)

(e) (f)

Figure 5.10: Comparison of Pareto fronts of classical NSGA-III optimization (left) with DO (right) after 7 generations. The high dimensional Pareto front is illustrated as a combinations of efficiency vs. $V_{cav,f}$, $V_{cav,b}$ and Cavity volume acceleration ($\frac{\partial^2 V_{cav}}{\partial^2 t}$)

Both classical NSGA-III optimization and DO are iterated through 7 generation with 72 individuals in the population. For DO the classifier is used to guide search from the second generation onwards. Figure 5.10-5.11 compares the Pareto front after the 7-th generation for classical NSGA-III and DO. The high dimensional Pareto front is plotted



(a)

(b)

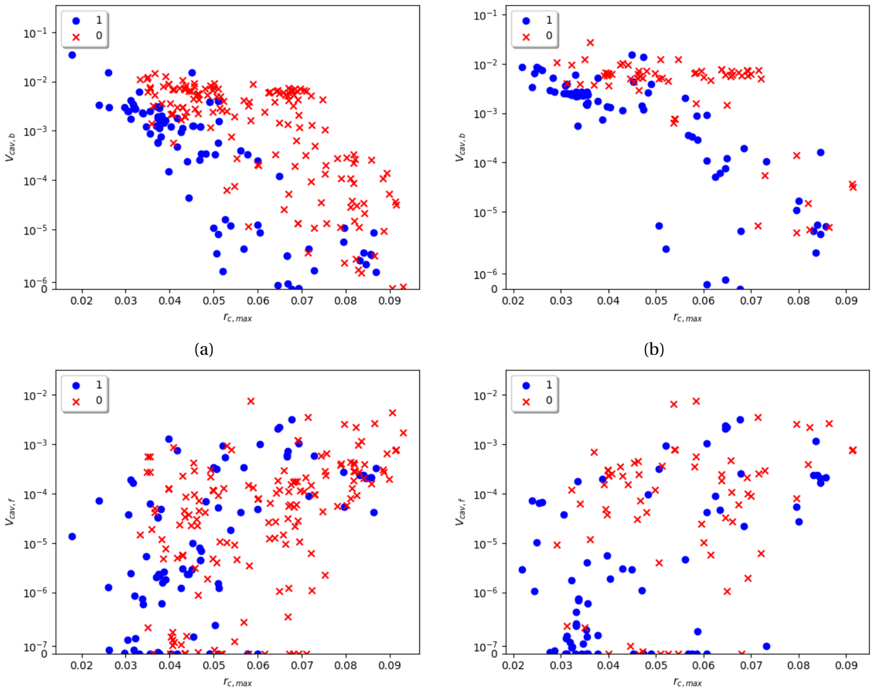
(c)

(d)

Figure 5.11: Comparison of Pareto fronts of classical NSGA-III optimization (left) with DO (right) after 7 generations. The high dimensional Pareto front is illustrated as a combinations of efficiency vs. Maximum vortex cavity radius ($r_{c,max}$) and $max(C_p)$.

with η as the abscissa and $V_{cav,b}$, $V_{cav,f}$, Cavity volume acceleration ($\frac{\partial^2 V_{cav}}{\partial^2 t}$), Maximum vortex cavity radius ($r_{c,max}$), $max(C_p)$ as ordinates to illustrate the trade-off with efficiency. Blue dots represent designs which lie on the Pareto front ($y_c = 1$) and the red crosses represent dominated solutions ($y_c = 0$). It is observed that the Pareto fronts are comparable for both NSGA-III and DO.

Figures 5.10a-5.10d illustrate the trade-off between $V_{cav,b}$, $V_{cav,f}$ and η. It is observed that improving efficiency over 52% is always accompanied by a cavitation penalty. This is because up to η = 52%, efficiency can be increased without significantly loading the mid section of the propeller however, the tip is heavily loaded. This leads to larger tip-vortices as shown in Figure 5.11a. In this region, where $\eta \in [40\%, 52\%]$, lower mid-section loading also suggests lower supervelocities at the mid-section compared to the tip. Consequently, higher rotational losses are expected. Furthermore, up to η = 52%, viscous losses are expected to be reduced since, the gain in efficiency does not lead to an exponential cavitation penalty, . Thus, for region η > 52%, gain in efficiency is expected to result from reduced rotational losses.



(a) (b)

(c) (d)

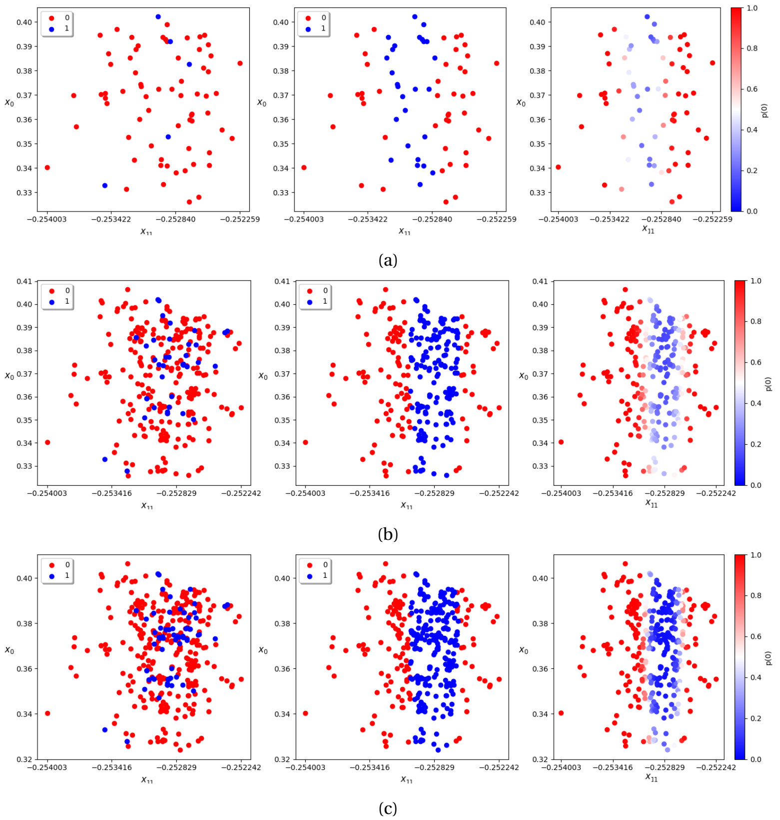
Figure 5.12: Trade-off between V_{cav} and $r_{c,max}$ as found by NSGA-III optimization (left) and DO (right). The Pareto front visualized with (a,b) $V_{cav,b}$ as abscissa and, (c,d) $V_{cav,f}$ as abscissa.

From Figures 5.10e-5.10f, it is observed that the cavity volume acceleration is better contained despite a growth in cavity volume. This is likely because skew is also optimized. From Figures 5.11c-5.11d it can be observed that reducing $max(C_p)$ results in more efficient propellers. This is because higher peak pressures lead to higher induced losses which can be modelled in BEM. In addition to induced losses, higher peak pressures may also lead to flow separation which requires scale resolving simulations, such as LES or Dynamic Eddy Simulations (DES).

Figures 5.11c-5.11d also show that very high values of $max(C_p)$ are observed. The bounds of $max(C_p)$ are determined by the stagnation pressure at the leading edge and the vapour pressure due to cavitation. Excessively high values of $max(C_p)$ may indicate numerical artefacts, which are often observed at nodes on the leading edge when BEM is used. By specifying the minimization of $max(C_p)$ as an objective, search algorithms are encouraged to minimize these unrealistic spikes. However, since same panel distribution is used for each propeller, the relative ranking of the designs is expected to remain unaffected.

From Figures 5.11a-5.11b, it is observed that $r_{c,max}$ can be reduced down to 4cm and efficiency can be increased suggesting a reduction in rotational losses. However, further reducing $r_{c,max}$ results in an efficiency penalty. The penalty is expected to result from growth in cavity volume both on the pressure and suction surface of the propeller. Figure 5.12 illustrates the trade-off between V_{cav} and $r_{c,max}$. It is observed from Figures 5.12a-5.12b that reducing $r_{c,max}$ results in a higher $V_{cav,b}$. This suggests that a trade-off exists between higher tip v/s higher mid-section loading.

5.3.3. NAIVE-BAYES CLASSIFIER



(a)

(b)

(c)

Figure 5.13: Evolution of Naive-Bayes label predictions over generations, with blue dots ($y_c = 1$) indicating solutions that lie on the Pareto front. In all plots, the x-axis is x_0 and y-axis is x_{11}. The plots are for (a) Generation I, (b) Generation IV and (c) Generation VII. Left: Ground reality, Center: The label predicted by Naive-Bayes classifer, Right: The cluster probabilities.

Figure 5.13 illustrates the label predictions made by the trained Naive-Bayes classifier. In Figures 5.13a-5.13c, the left most subplot illustrates the ground reality. Interesting designs which lie on the Pareto front have $y_c = 1$. The centre subplot illustrates the label predicted by the trained classifier and the right most subplot illustrates the $p(\vec{x}|\mu_0, \sigma_0^2)$. It is observed that the classifier learns the location of the Pareto front by the first generation visible as a light blue band in Figure 5.13a. The learning is reinforced in further generations, i.e. the classifier gets more confident of the location of the Pareto front. This is visible as the blue band intensifies. This is also observed from the so-called confusion matrix whose values are plotted in Figure 5.14. It is observed that the true positive predictions(tp) for designs that lie on the Pareto front steadily increases generation over generation. As the model becomes more accurate in predicting the location of the Pareto front, it is possible to reduce evaluations of designs which are expected to lie meaningfully away from the Pareto front reducing the computational cost of optimization.

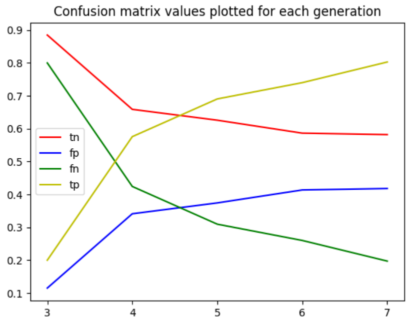


Figure 5.14: Values of the confusion matrix, tp - True Positive, tn - True Negative, fp - False Positive, fn - False Negative

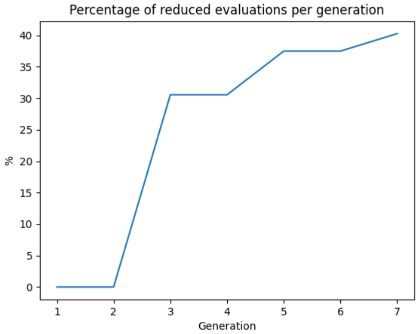


Figure 5.15: Percentage of evaluations reduced per generation

Figure 5.15 shows the percentage of reduced evaluations per generation as the optimization progresses. From the third generation onwards the classifier reduces at least 30 % of evaluations. With online learning, the classifier is better able to identify the Pareto front as the optimization progresses and the percentage of reduced evaluations increases steadily to 40 %.

5.3.4. PARETO SOLUTIONS

Figure 5.16 shows the so-called Parallel Coordinate plot for explored designs (-) and the designs of interest that lie on the Pareto front with $V_{cav,f} = 0.0m^3$ (-). There are 19 diverse designs of interest. Among the designs, increasing efficiency comes with a trade-off with either increasing suction side cavitation, higher cavity volume acceleration or tip-vortex radius as detailed in Subsection 5.3.2. In the Parallel Coordinate plot, if $\Delta\theta, \Delta\gamma, s_{camber} < 0$, the pitch, skew and camber for the hydrofoil cross-section have reduced. Otherwise, they have remained the same or increased compared to the baseline geometry. If $s_{chord} < 1.0$, the hydrofoil chord-length has reduced. Otherwise, the chord-length has either remained the same or increased.

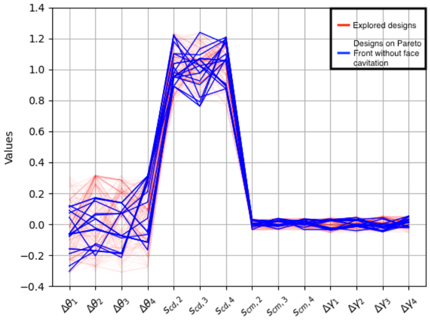


Figure 5.16: Parallel coordinate plot of explored designs and designs that lie on the Pareto front.

Table 5.4 Predicted objective function values.

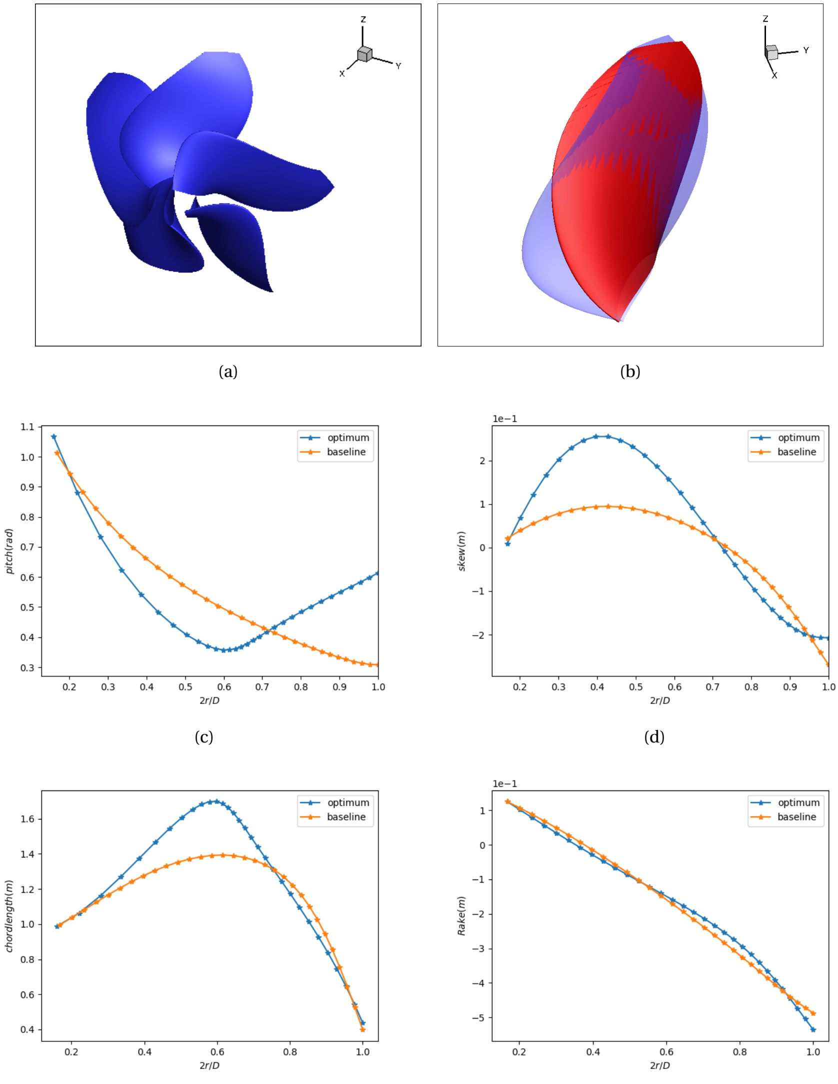
f_i	Units	Value
η	-	5.2023e1 %
$V_{cav,b}$	m^3	3.2972e−4
$V_{cav,f}$	m^3	0.0000
$\frac{\partial^2 V_{cav}}{\partial^2 t}$	$m^3 s^{-1}$	1.7486e−1
$r_{c,max}$	m	5.761e−2
$max(C_p)$	-	8.2648e1

Table 5.6 Design parameters and performance for the least cavitating design.

Variable	Units	Value
$\Delta\theta_1$	rad	$4.1000e{-2}$
$\Delta\theta_2$	rad	$-1.8405e{-1}$
$\Delta\theta_3$	rad	$7.5432e{-2}$
$\Delta\theta_4$	rad	$2.7701e{-1}$
$\Delta\gamma_1$	rad	$3.5201e{-1}$
$\Delta\gamma_2$	rad	$1.3164e{-1}$
$\Delta\gamma_3$	rad	$3.6064e{-1}$
$\Delta\gamma_4$	rad	$-2.2871e{-1}$
$s_{chord,2}$	-	1.1440
$s_{chord,3}$	-	$9.8151e{-1}$
$s_{chord,4}$	-	1.0800
$s_{camber,2}$	-	$1.9978e{-2}$
$s_{camber,3}$	-	$-3.7130e{-4}$
$s_{camber,4}$	-	$2.0544e{-3}$

For the current operating point, a cavitation free solution was not found. Although the maximum efficiency is $max(\eta) = 63.49\%$, the corresponding solution has a large suction surface cavity. Solutions without pressure side cavitation and much lower suction surface cavitation are found on the Pareto front. Among solutions without pressure side cavitation, the design with the minimum suction side cavitation is illustrated in Figure 5.17. Figure 5.17c compares the pitch distribution of this optimized design and baseline. Compared to the baseline design, the pitch at the root section is slightly decreased while the pitch at the blade tip has been increased substantially. Figure 5.17d compares the skew distribution of the optimized design and baseline. The skew at the mid-section and at the tip has been increased. Figure 5.17e compares the chordlength distribution of this optimized design and baseline. The chordlength is measured for the expanded hydrofoil. It is noted that the chordlength of the mid-section has increased substantially. The chordlength of the tip-section has reduced compared to the baseline. From Figure 5.17f it is observed that the rake of the blade has also changed from the baseline while λ has not been varied. As mentioned in Section 5.2.2, the change most likely results from the sequence in which $\theta, \gamma, s_{chord}, s_{camber}$ are varied. Furthermore, it could also result from residuals in fitting of B-Spline curve to the cross-section and lofting operation of a B-Spline surface through the fit curves. Table 5.5 lists the design parameters. In the table, the design parameters provide insight into how the geometry has changed. The chordlength at the root and tip has reduced while that at the mid-section has increased. The camber of the hydrofoil has been reduced at the root and mid-section while it has increased at the tip.

Figure 5.18 illustrates the pressure, velocity, skin-friction coefficient and cavitation behaviour resulting from the previously mentioned changes to the baseline. Table 5.4 lists the objective function values. It is observed that the blade loading is optimized to eliminate discontinuities at the blade tip in velocity and pressure field when compared

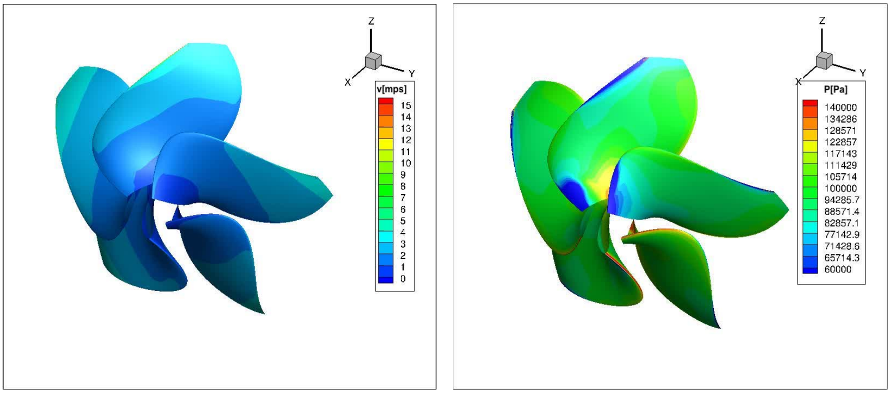


(a) (b)

(c) (d)

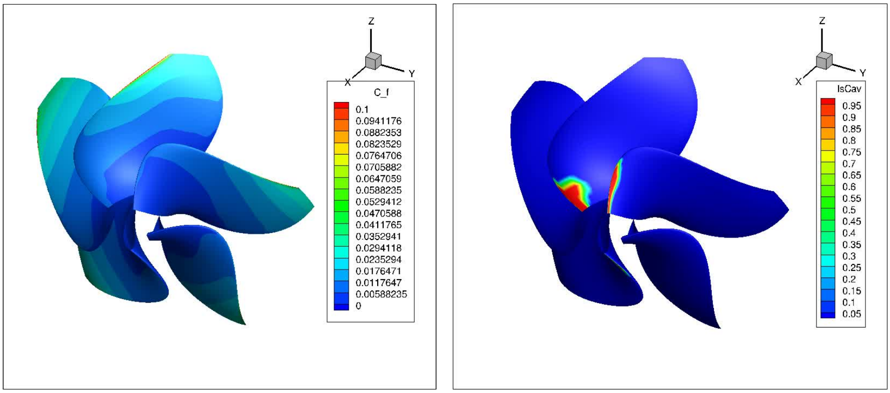
(e) (f)

Figure 5.17: Design solution with minimum suction side cavitation volume:(a) geometry, (b) comparison with baseline geometry, (c) pitch distribution of baseline and design, (d) skew distribution of baseline and design and, (e) chordlength distribution of baseline and design and, (f) rake distribution of baseline and design.



(a) Inviscid velocity distribution

(b) The pressure field on the blade



(c) Skin friction approximation.

(d) Fraction of cavitating vertices on the panel.

Figure 5.18: Design 1 (a) inviscid velocity distribution, (b) pressure filed on the blade, (c) skin friction coefficient and, (d) fraction of cavitating mesh panels. $|\vec{v}| = \sqrt{v_x^2 + v_y^2 + v_z^2}$ and IsCav $\in [0.00, 0.25, 0.50, 0.75, 1.00]$.

with the baseline. Nonetheless, *Design 1* seems to have mitigated adverse pressure gradients. However, suction pressure at the blade root and along the leading edge indicate a risk of cavitation. It is observed observed that loading on the mid-section is reduced as the blade is de-pitched and camber is reduced. In order to meet the thrust constraint, the loading is shifted more towards the blade tip by increasing the tip pitch and camber. Consequently, $V_{cav,b}$ is reduced and $r_{c,max}$ increases. Due to higher tip loading the torque requirement and rotational losses increases resulting in lower hydrodynamic efficiency.

5.4. CONCLUSION

An optimization method to solve multi-objective, constrained optimization problems based on machine learning strategies is proposed. The method referred to as *Dynamic*

Optimization is demonstrated on the hydrodynamic optimization of Wageningen B-4 70 propeller with P/D=1.0 operating in the Seiun-Maru wakefield. It has been demonstrated that compared to classical NSGA-III optimization, DO is able to reduce at least 30 % of evaluations per generation while reproducing a comparable Pareto front. The gain in computational efficiency of optimization is achieved with an online learning strategy using the Naive-Bayes classifier, which identifies the location of the Pareto front and designs that are located close to the front. Both classical NSGA-III and DO identify similar trade-offs between suction side, pressure side, tip-vortex cavitation and efficiency. For the current demonstration case, improving efficiency is accompanied by a cavitation penalty. It is possible to reduce suction side cavitation and eliminate pressure side cavitation fully. Reducing suction side cavitation requires a higher tip loading resulting in a larger tip vortex. Nonetheless, a non-elitist search strategy results in a Pareto front with very diverse designs.

ACKNOWLEDGEMENTS

The authors are also grateful for CRS (Cooperative Research Ships) for supporting this research with PROCAL. The authors also thank Dr. ir. Evert -Jan Foeth and Dr. Thomas Scholcz of MARIN, The Netherlands for their valuable inputs on feedback and support for this research.

6

CONCLUSIONS AND RECOMMENDATIONS

6.1. CONCLUSION

The hydro-acoustic optimization of a propeller blade can be characterized as a multi-objective, constrained optimization problem. The preferred solution often has an optimal blade loading which maximizes efficiency and minimizes acoustic emissions. Even in the preliminary design phase finding solutions is computationally expensive. Thus, there is a requirement to improve the efficiency of optimization while achieving a reasonable trade-off with completeness and number of design variables. To meet the requirement, this research has investigated (a) a new parametric model to quantify geometric variation accurately, (b) machine learning strategies to model and account for uncertainties, (c) assessment of sources of uncertainties in acoustic emission predictions by semi-empirical models and, (d) machine learning strategies to improve efficiency by learning the location of lucrative regions in the design space.

At the heart of the proposed parametric model is Singular Value Decomposition (SVD). It is demonstrated to enable reducing the number of samples in Design of Experiments and classify design variables based on their magnitude of influence on the geometry. For a propeller blade case SVD of the geometry probe data matrix enabled the classification of variables based on the magnitude of influence on the geometry. In the case of the airfoil, it was possible to identify from SVD of the geom probe data matrix that classification of variables is not viable. Furthermore, it was also possible to reduce the number of sample points by 50% while still capturing the performance trends in the design space. This is because the proposed method to select sample points preserve some of the smaller dimensions/modes.

In order to account for uncertainties, it is important that uncertainties in performance predictions of design tools are quantified. Uncertainties for efficiency predictions made in BEM have been extensively investigated. However, sources of uncertainties in acoustic emission predictions made by ETV-2 are yet to be investigated. Thus, in this research, the dynamics of an isolated cavitating vortex without forced excitation is simulated. Similar to Bosscher's observations (Bosschers J., 2018), our simulations find that kinematics of the vortex differ significantly up to the viscous core from the Proctor-Winkelmans vortex model. The Proctor-Winkelmans model is suited for fully wetted vortices. However, the kinematics of cavitating vortices differs from that of a fully wetted vortex. It is observed that the model under-predicts tangential velocity at the core and over-predicts the tangential velocity up to the viscous core. Furthermore, the model does not account for inflection in the tangential velocity. The inflection results in Kelvin-Helmholtz instabilities that grow to excite the vortex. Compressibility effects are expected to play a limited role in determining vortex dynamics or cavity radius as water tends to behave as an incompressible fluid. The cavitating core of the vortex tends to behave as a compressible fluid but given that the density of the vapour is several orders of magnitude lower than that of water, the vapour core is less likely to influence vortex dynamics. Given that the ETV-2 model makes assumptions on the centre frequency of the broad-band hump and is trained on empirical data from specific test cases, the acoustic emission predictions made are expected to have high uncertainties. Based on the simulations, it is expected that the uncertainties in predicting radius of the tip-vortex cavity is also relatively high. However, there is more actionable insight in sources of uncertainty in kinematics. Hence, in this research ETV-2 is applied to minimize noise indirectly with the objective to minimize tip-vortex cavity radius. A smaller cavity is expected to have lower cavity volume acceleration and a weaker breakdown due to lower potential energy thus, it is also expected to have a lower Sound Pressure Levels. However, smaller cavities tend to have higher resonant frequencies, thus the acoustic emissions are also likely to be at higher frequencies.

Two machine learning strategies to improve the efficiency of optimization are proposed. The first is demonstrated on a constrained single objective optimization of propellers. The optimization method referred as *Dynamic Optimization* synergises supervised and unsupervised learning. Compared to a standard DoE based search, DO is almost 50 % more efficient in the current demonstration case due to the proposed orthogonal parametric model and applied classifiers. There is a trade-off between efficient optimization and completeness, nonetheless DO provides sufficient flexibility to intervene and manage the trade-off.

The second optimization method is based on online learning strategies. It is demonstrated on the multi-objective, constrained propeller optimization problem for a propeller operating in a wakefield. It has been demonstrated that compared to classical NSGA-III optimization, DO is able to reduce atleast 30 % of evaluations per generation while reproducing a comparable Pareto front. Dynamic Optimization identifies trade-offs between suction side, pressure side, tip-vortex cavitation and efficiency. Complemented by a non-elitist search strategy results, it yields a Pareto front with a large variety in possible design solutions.

6.2. RECOMMENDATIONS

This research is intended to have a societal and ecological impact with a method to design and optimize green and quiet propellers. The necessary building blocks and prerequisites for such a method are investigated and findings reported in the chapters of this thesis. For recommendations and possible follow-ups specific to the building blocks, the reader is advised to the bulleted recommendations at the end of each chapter. In this section, the author recommends investigating new ideas or concepts to design green and quite propellers.

In addition to shape optimization of rigid propellers another very promising opportunity is to design shape adaptive propellers which passively deform to achieve optimal loads. There are two strategies to achieve this. The *first* is with metal or alloy propellers whose topology has been optimized to shift the inertial axis ahead of the loading axis. This allows for depitching under peak load. Similar examples are available in aerospace engineering, where the wing-box of a swept wing tends to have an inertial axis ahead of the loading axis which tends to depitch the wing. Manufacturing blades with such internal structures is a challenging problem, however additive manufacturing offers interesting opportunities to manufacture such blades. The second strategy is via composite propellers with bend-twist coupling. Composite propellers allow for depitching of heavily loaded blades to achieve optimal loads. Both these strategies need to be further investigated to identify opportunities and risks. Nonetheless, the optimization methods proposed in this research need to be extended to handle design and optimization of deformable propellers.

The first proposed extension is for the orthogonal parametric model. It must be built from both geometric and structural data. For topology optimization, a flattened material density field $\rho(u,v,w)$, where ρ is the material density at equivalent coordinates u,v,w in the solid, can be appended to the geometric design vector. Thus, a design vector will consist of both geometry and structural information. For composite structures, the laminate details need to be appended to the design vector. Laminate properties such as fibre orientation at u,v,w may be sufficient. However, further investigation is required to see if such a design vector correlates well with strains under load.

The second proposed extension is for the ETV-2 cavitation model. Extensions are required to account for deviations in tangential velocity from the kinematic models. This would allow for better estimation of the tip vortex cavity radius. Consequently, the uncertainties in hydroacoustic performance predictions would reduce.

The third proposed extension is for an uncertainty model which can be used as the error term ($\epsilon = N(0,\sigma^2)$) in Gaussian Process Regression. The uncertainty model can be developed by computational experiments where geometry, structure and design points for propellers are perturbed to identify regions where simulations tend to diverge or are known to have high uncertainties. The model can be used to identify the boundaries of application for performance prediction methods such as BEM to focus search in areas where the uncertainties are low.

By extending (a) the orthogonal parametric model to account for structural variables, (b) improving the kinematic model and extending ETV-2 to make dynamic acoustic prediction for deforming blade tips and, (c) interfacing machine learning models with physics in performance prediction tools, the framework of DO can be improved and applied to optimize deformable propellers. This makes green and quiet deformable propellers more accessible.

REFERENCES

[1] Shipmap.org | visualisation of global cargo ships | by kiln and ucl. `https://www.shipmap.org/`. (Accessed on 03/17/2021).

[2] *The problem of multicollinearity*, pages 176–180. Springer US, Boston, MA, 1997.

[3] Z. Abedjan, N. Boujemaa, S. Campbell, P. Casla, S. Chatterjea, S. Consoli, C. Costa-Soria, P. Czech, M. Despenic, C. Garattini, et al. Data science in healthcare: benefits, challenges and opportunities. In *Data science for healthcare*, pages 3–38. Springer, 2019.

[4] K. Abrahamsen. The ship as an underwater noise source. In *Proceedings of Meetings on Acoustics ECUA2012*, volume 17, page 070058. Acoustical Society of America, 2012.

[5] N. Adams and S. Hickel. Implicit large-eddy simulation: Theory and application. In *Advances in turbulence XII*, pages 743–750. Springer, 2009.

[6] A. Arapakopoulos, R. Polichshuk, Z. Segizbayev, S. Ospanov, A. Ginnis, and K. Kostas. Parametric models for marine propellers. *Ocean Engineering*, 192:106595, 2019.

[7] F. P. Arribas and R. P. Fernández. A b-spline design model for propeller blades. *Advances in Engineering Software*, 118:35 – 44, 2018.

[8] P. T. Arveson and D. J. Vendittis. Radiated noise characteristics of a modern cargo ship. *The Journal of the Acoustical Society of America*, 107(1):118–129, 2000.

[9] D. Beattie and P. Whalley. A simple two-phase frictional pressure drop calculation method. *International Journal of Multiphase Flow*, 8(1):83 – 87, 1982.

[10] R. E. Bensow and G. Bark. Implicit LES Predictions of the Cavitating Flow on a Propeller. *Journal of Fluids Engineering*, 132(4):41302–41310, apr 2010.

[11] J. Blank, K. Deb, Y. Dhebar, S. Bandaru, and H. Seada. Generating well-spaced points on a unit simplex for evolutionary many-objective optimization. *IEEE Transactions on Evolutionary Computation*, In press.

[12] P. A. M. Boomkamp. Stability of parallel two-phase flow. *These, Universiteit Twente*, 215:216, 1998.

[13] J. Bosschers. Analysis of inertial waves on inviscid cavitating vortices in relation to low-frequency radiated noise. In *WIMRC Cavitation Forum*. Warwick University UK, 2008.

[14] J. Bosschers. Investigation of hull pressure fluctuations generated by cavitating vortices. In *Proc. First Symposium on Marine Propulsors*, 2009.

[15] J. Bosschers. *Propeller tip-vortex cavitation and its broadband noise*. PhD thesis, University of Twente, Netherlands, Sept. 2018.

[16] J. Bosschers. A semi-empirical prediction method for broadband hull-pressure fluctuations and underwater radiated noise by propeller tip vortex cavitation. *Journal of Marine Science and Engineering*, 6(2):49, 2018.

[17] S. Chandrasekhar. *Hydrodynamic and hydromagnetic stability*. Courier Corporation, 2013.

[18] C.-C. Chang and C.-J. Lin. Libsvm: A library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, 2(3):1–27, 2011.

[19] K. Deb and H. Jain. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: solving problems with box constraints. *IEEE transactions on evolutionary computation*, 18(4):577–601, 2013.

[20] M. Diez, E. F. Campana, and F. Stern. Design-space dimensionality reduction in shape optimization by karhunen–loéve expansion. *Computer Methods in Applied Mechanics and Engineering*, 283:1525 – 1544, 2015.

[21] M. Diez, A. Serani, E. F. Campana, S. Volpi, and F. Stern. Design space dimensionality reduction for single-and multi-disciplinary shape optimization. In *17th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, page 4295, 2016.

[22] P. G. Drazin and W. H. Reid. *Hydrodynamic stability*. Cambridge university press, 2004.

[23] M. Drela. Xfoil: An analysis and design system for low reynolds number airfoils. In T. J. Mueller, editor, *Low Reynolds Number Aerodynamics*, pages 1–12, Berlin, Heidelberg, 1989. Springer Berlin Heidelberg.

[24] C. P. Egerer, S. J. Schmidt, S. Hickel, and N. A. Adams. Efficient implicit les method for the simulation of turbulent cavitating flows. *Journal of Computational Physics*, 316:453–469, 2016.

[25] B. Epps. *Review of Vortex Identification Methods*.

[26] R. B. FlorianVesting. Propeller optimisation considering sheet cavitation and hull interaction. In *Second International Symposium on Marine Propulsors*, 2011.

[27] E.-J. Foeth. Propeller optimization using an unsteady boundary-element method. In *Fourth international symposium on marine propulsors*, pages 27–32, 2015.

[28] S. Friedlander and A. Lipton-Lifschitz. Localized instabilities in fluids. In *Handbook of mathematical fluid dynamics*, volume 2, pages 289–354. Elsevier, 2003.

[29] S. Gaggero, G. Tani, M. Viviani, and F. Conti. A study on the numerical prediction of propellers cavitating tip vortex. *Ocean engineering*, 92:137–161, 2014.

[30] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.

[31] M. Gubisch and S. Volkwein. Proper orthogonal decomposition for linear-quadratic optimal control. *Model Reduction and Approximation: Theory and Algorithms*, 15:1, 2017.

[32] S. Hickel, N. A. Adams, and J. A. Domaradzki. An adaptive local deconvolution method for implicit les. *Journal of Computational Physics*, 213(1):413 – 436, 2006.

[33] S. Hickel, C. P. Egerer, and J. Larsson. Subgrid-scale modeling for implicit large eddy simulation of compressible flows and shock-turbulence interaction. *Physics of Fluids*, 26(10):106101, 2014.

[34] S. Hickel, M. Mihatsch, and S. Schmidt. Implicit large eddy simulation of cavitation in micro channel flows. *arXiv preprint arXiv:1401.6548*, 2014.

[35] J. Huisman and E.-J. Foeth. Automated multi-objective optimization of ship propellers. In *Proceedings of the Fifth International Symposium on Marine Propulsors (SMP'17), Espoo, Finland*, 2017.

[36] S. Knerr, L. Personnaz, and G. Dreyfus. Single-layer learning revisited: a stepwise procedure for building and training a neural network. In *Neurocomputing*, pages 41–50. Springer, 1990.

[37] R. Liebrand, M. Klapwijk, T. Lloyd, and G. Vaz. Transition and Turbulence Modeling for the Prediction of Cavitating Tip Vortices. *Journal of Fluids Engineering*, 143(1), 10 2020. 011202.

[38] P. Maljaars. *Hydro-elastic analysis of flexible marine propellers*. PhD thesis, Delft University of Technology, 2019.

[39] J. Matheis and S. Hickel. Multi-component vapor-liquid equilibrium model for les of high-pressure fuel injection and application to ecn spray a. *International Journal of Multiphase Flow*, 99:294 – 311, 2018.

[40] P. Moriarty. Nafnoise user's guide. *National Wind Technology Center, National Renewable Energy Laboratory*, 2005.

[41] P. Moriarty and P. Migliore. Semi-empirical aeroacoustic noise prediction code for wind turbines. 12 2003.

[42] OECD. Global value chains (gvcs) - oecd, 2019.

[43] F. Örley, S. Hickel, S. J. Schmidt, and N. A. Adams. Large-eddy simulation of turbulent, cavitating fuel flow inside a 9-hole diesel injector including needle movement. *International Journal of Engine Research*, 18(3):195–211, 2017.

[44] A. Oyama, T. Nonomura, and K. Fujii. Data Mining of Pareto-Optimal Transonic Airfoil Shapes Using Proper Orthogonal Decomposition. *Journal of Aircraft*, 47(5):1756–1762, sep 2010.

[45] P. Pennings. *Dynamics of Vortex Cavitation*. PhD thesis, Delft University of Technology, Netherlands, 06 2016.

[46] P. C. Pennings, J. Westerweel, and T. J. C. van Terwisga. Flow field measurement around vortex cavitation. *Experiments in Fluids*, 56(11):206, Nov 2015.

[47] T. v. T. K. V. Pranav Doijode Sumanth, Stefan Hickel. A machine learning approach for propeller design and optimisation : Part - ii. unpublished.

[48] J. L. Proctor, S. L. Brunton, and J. N. Kutz. Dynamic mode decomposition with control. *SIAM Journal on Applied Dynamical Systems*, 15(1):142–161, 2016.

[49] P. G. Saffman. *Vortex dynamics*. Cambridge university press, 1992.

[50] P. J. Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of fluid mechanics*, 656:5–28, 2010.

[51] J. J. Schot, P. C. Pennings, M. Pourquie, and T. J. Van Terwisga. Modelling of tip vortex cavitation for engineering applications in openfoam. In *Proceedings-WCCM XI: 11th World Congress on Computational Mechanics; ECCM V: 5th European Conference on Computational Mechanics; ECFD VI: 6th European Conference on Computational Fluid Dynamics, Barcelona, Spain, 20-25 July 2014*. CIMNE, 2014.

[52] P. D. Sumanth, S. Hickel, and T. van Terwisga. On the design of experiments for propellers-getting to design variables that matter and design points that are important. In *Numerical Towing Tank Symposium, 2018*, 2018.

[53] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz. On dynamic mode decomposition: Theory and applications. *arXiv preprint arXiv:1312.0041*, 2013.

[54] Y. Ukon and H. Yuasa. Pressure distribution and blade stress on a highly skewed propeller. 1994.

[55] UN. Un business - imo, 2019.

[56] T. van Terwisga. Chapter 9, tu delft ocw, cavitation on ship propellers. `https://ocw.tudelft.nl/wp-content/uploads/Chapter_9.pdf`. (Accessed on 03/17/2021).

[57] F. Vesting, R. Gustafsson, and R. E. Bensow. Development and application of optimisation algorithms for propeller design. *Ship Technology Research*, 63(1):50–69, 2016.

[58] W. Wagner and A. Pruß. The iapws formulation 1995 for the thermodynamic properties of ordinary water substance for general and scientific use. *Journal of physical and chemical reference data*, 31(2):387–535, 2002.

[59] Q. Wang, H. Chen, R. Hu, and P. Constantine. Conditional sampling and experiment design for quantifying manufacturing error of transonic airfoil. jan 2011.

[60] K. Washabaugh, D. Amsallem, M. Zahr, and C. Farhat. Nonlinear Model Reduction for CFD Problems Using Local Reduced-Order Bases. In *42nd AIAA Fluid Dynamics Conference and Exhibit*, Fluid Dynamics and Co-located Conferences. American Institute of Aeronautics and Astronautics, jun 2012.

[61] M. M. Barnitsas, D. Ray, and P. Kinley. Kt, kq and efficiency curves for the Wageningen B-series propellers. Technical report, University of Michigan, 1981.

[62] D. Bertetta, S. Brizzolara, S. Gaggero, M. Viviani, and L. Savio. CPP propeller cavitation and noise optimization at different pitches with panel code and validation by cavitation tunnel measurements. *Ocean Engineering*, 53:177 – 195, 2012.

[63] C. M. Bishop. *Pattern recognition and machine learning*. Springer, 2006.

[64] S. Brizzolara and S. Gaggero. Silent propellers with unconventional profile shapes. Examples obtained with a new automatic optimization method. In *Proceedings of the 4th Annual Conference MAST2009 on Maritime Systems and Technology. Stockholm, Sweden*, 2009.

[65] J. Carlton. 3 - propeller geometry. In J. Carlton, editor, *Marine Propellers and Propulsion (Second Edition)*, pages 32–50. Butterworth-Heinemann, Oxford, second edition edition, 2007.

[66] E. Esmailian, H. Ghassemi, and H. Zakerdoost. Systematic probabilistic design methodology for simultaneously optimizing the ship hull–propeller system. *International Journal of Naval Architecture and Ocean Engineering*, 9(3):246 – 255, 2017.

[67] S. Gaggero, G. Tani, D. Villa, M. Viviani, P. Ausonio, P. Travi, G. Bizzarri, and F. Serra. Efficient and multi-objective cavitating propeller optimization: An application to a high-speed craft. *Applied Ocean Research*, 64:31 – 57, 2017.

[68] N. Garg, G. K. Kenway, J. R. Martins, and Y. L. Young. High-fidelity multipoint hydrostructural optimization of a 3-D hydrofoil. *Journal of Fluids and Structures*, 71:15 – 39, 2017.

[69] J. H. Halton. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik*, 2(1):84–90, 1960.

[70] IMO. Reducing greenhouse gas emissions from ships, 2019. *Accessed online on 2019-09-13.* `http://www.imo.org/en/MediaCentre/HotTopics/Pages/Reducing-greenhouse-gas-emissions-from-ships.aspx`.

[71] S. Knerr, L. Personnaz, and G. Dreyfus. Single-layer learning revisited: a stepwise procedure for building and training a neural network. In *Neurocomputing*, pages 41–50. Springer, 1990.

[72] S. M. LaValle. *Planning algorithms*. Cambridge University Press, 2006.

[73] G.-J. Park, T.-H. Lee, K. H. Lee, and K.-H. Hwang. Robust design: An overview. *AIAA Journal*, 44(1):181–191, 2006.

[74] D. Peri. Conformal free form deformation for the optimisation of complex geometries. *Ship Technology Research*, 59(1):36–41, 2012.

[75] H. Prautzsch, W. Boehm, and M. Paluszny. *Bézier and B-spline techniques*, volume 6. Springer, 2002.

[76] C. E. Rasmussen and C. K. Williams. Gaussian processes for machine learning. ISBN 026218253x, 2006. Springer, 2006.

[77] P. J. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53 – 65, 1987.

[78] S. Sarkar, A. Dong, and J. S. Gero. Design Optimization Problem Reformulation Using Singular Value Decomposition. *Journal of Mechanical Design*, 131(8), 07 2009. 081006.

[79] B. Schuiling and T. van Terwisga. Energy analysis of a propeller in open water using a RANS method. In *24th International HISWA Symposium on Yacht Design and Yacht Construction. Amsterdam, The Netherlands*, 2016.

[80] G. Theotokatos and V. Tzelepis. A computational study on the performance and emission parameters mapping of a ship propulsion system. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*, 229(1):58–76, 2015.

[81] C. F. Van Loan and G. Golub. Matrix Computations (Johns Hopkins Studies in Mathematical Sciences). 1996.

[82] G. Vaz and J. Bosschers. Modelling three dimensional sheet cavitation on marine propellers using a boundary element method. In *Sixth International Symposium on Cavitation CAV2006, Wageningen, The Netherlands*, 2006.

[83] G. Vaz, D. Hally, T. Huuva, N. Bulten, P. Muller, P. Becchi, J. L. Herrer, S. Whitworth, R. Macé, and A. Korsström. Cavitating flow calculations for the e779a propeller in open water and behind conditions: code comparison and solution validation. In *Proceedings of the 4th International Symposium on Marine Propulsors (SMP'15), Austin, TX, USA*, pages 330–345, 2015.

[84] F. Vesting. *Marine propeller optimisation-strategy and algorithm development*. PhD thesis, Chalmers University of Technology, 2015.

[85] F. Vesting, R. Johansson, and R. Bensow. Parameter influence analysis in propeller optimisation. In *Proceedings of the Third International Symposium on Marine Propulsors.(Launceston, Tasmania, Australia). Ed. by J. Binns, R. Brown, and N. Bose. University of Tasmania*, pages 397–404, 2013.

[86] P. Doijode Sumanth, S. Hickel, T. van Terwisga, and K. Visser. A machine learning approach for propeller design and optimisation: Part II Under review, 2021.

[87] J. Blank, K. Deb, and P. C. Roy. Investigating the normalization procedure of NSGA-III. In *International Conference on Evolutionary Multi-Criterion Optimization*, volume 1, pages 229–240. Springer, 2019.

[88] J. Bosschers. Investigation of hull pressure fluctuations generated by cavitating vortices. In *Proc. First Symposium on Marine Propulsors*, 2009.

[89] J. Carlton. Chapter 3 - propeller geometry. In J. Carlton, editor, *Marine Propellers and Propulsion (Third Edition)*, pages 29 – 46. Butterworth-Heinemann, Oxford, third edition edition, 2012.

[90] K. Deb and H. Jain. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part I: solving problems with box constraints. *IEEE transactions on evolutionary computation*, 18(4):577–601, 2013.

[91] P. Doijode Sumanth, S. Hickel, T. van Terwisga, and K. Visser. A machine learning approach for propeller design and optimization: Part I. Under review, 2021.

[92] R. Faresi. Full scale measurements. In *ITTC Group Discussions*, 1990.

[93] N. Garg, G. K. Kenway, J. R. Martins, and Y. L. Young. High-fidelity multipoint hydrostructural optimization of a 3D hydrofoil. *Journal of Fluids and Structures*, 71:15–39, 2017.

[94] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.

[95] J. Huisman and E.-J. Foeth. Automated multi-objective optimization of ship propellers. In *Proceedings of the Fifth International Symposium on Marine Propulsors (SMP'17), Espoo, Finland*, 2017.

[96] R. Liebrand, M. Klapwijk, T. Lloyd, and G. Vaz. Transition and turbulence modeling for the prediction of cavitating tip vortices. *Journal of Fluids Engineering*, 143(1):011202, 2021.

[97] Z. Liu and Y. L. Young. Utilization of bend–twist coupling for performance enhancement of composite marine propellers. *Journal of Fluids and Structures*, 25(6):1102–1116, 2009.

[98] P. J. Maljaars and M. L. Kaminski. Hydro-elastic analysis of flexible propellers: an overview. In *Fourth International Symposium on Marine Propulsors*, 2015.

[99] L. Miglianti, F. Cipollini, L. Oneto, G. Tani, S. Gaggero, A. Coraddu, and M. Viviani. Predicting the cavitating marine propeller noise at design stage: A deep learning based approach. *Ocean Engineering*, 209:107481, 2020.

[100] M. Motley, Z. Liu, and Y. Young. Utilizing fluid–structure interactions to improve energy efficiency of composite marine propellers in spatially varying wake. *Composite Structures*, 90(3):304–313, 2009.

[101] M. L. Mulcahy, P. Croaker, D. G. McGuckin, P. A. Brandner, and N. Kississoglou. Optimisation applied to composite marine propeller noise. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, volume 249, pages 4012–4019. Institute of Noise Control Engineering, 2014.

[102] Port of Vancouver. New incentive for cargo and cruise vessels intended to quiet waters around the port of vancouver for at-risk whales – port of vancouver. https://www.portvancouver.com/news-and-media/news/new-incentive-for-cargo-and-cruise-vessels-intended-to-quiet-waters-arou

[103] Scikit Learn. SKLearn Multiclass One-vs-Rest Classifier https://scikit-learn.org/stable/modules/generated/sklearn.multiclass.OneVsRestClassifier.html.

[104] A. I. Oprea. *Prediction of tip vortex cavitation for ship propellers*. PhD Thesis, University of Twente, 2013.

[105] International Maritime Organization. Guidelines for the reduction of underwater noise from commercial shipping to address adverse impacts on marine life. https://cetsound.noaa.gov/Assets/cetsound/documents/MEPC.1-Circ%20883%20Noise%20Guidelines%20April%202014.pdf.

[106] International Maritime Organization. Prevention of air pollution from ships. https://www.imo.org/en/OurWork/Environment/Pages/Air-Pollution.aspx.

[107] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[108] P. Pennings, J. Westerweel, and T. van Terwisga. Cavitation tunnel analysis of radiated sound from the resonance of a propeller tip vortex cavity. *International Journal of Multiphase Flow*, 83:1 – 11, 2016.

[109] M. M. Plucinski, Y. L. Young, and Z. Liu. Optimization of a self-twisting composite marine propeller using genetic algorithms. In *16th International conference on composite materials, Kyoto, Japan*, pages 8–13, 2007.

[110] Y. Tendo, Y. Okada, and A. Okazaki. Developing the high efficiency propeller with optimization of skeg profile for twin-screw LNG carrier. Sixth International Symposium on Marine Propulsors (SMP'19), Rome, Italy May, 2019.

[111] T. van Terwisga, E. van Wijngaarden, J. Bosschers, and G. Kuiper. Achievements and challenges in cavitation research on ship propellers. *International shipbuilding progress*, 54(2-3):165–187, 2007.

[112] F. Vesting and R. E. Bensow. On surrogate methods in propeller optimisation. *Ocean Engineering*, 88:214 – 227, 2014.

[113] Y. L. Young, M. R. Motley, R. Barber, E. J. Chae, and N. Garg. Adaptive composite marine propulsors and turbines: progress and challenges. *Applied Mechanics Reviews*, 68(6):060803, 2016.

[114] H. Kato, Y. Kodama. Microbubbles as a skin friction reduction device–a midterm review of the research. *Proceedings of the 4th Symposium on Smart Control of Turbulence, Tokyo, Japan. 2003.*

APPENDIX

APPENDIX I

Algorithm 7 Clustering Algorithm

```
    Input: $X, \lambda$
    Output:$L0, L1$
procedure GET_CLUSTERS
    $unsup\_clf \leftarrow$ GaussianMixture(n_clusters=2, initialization=kmeans)
    unsup_clf.train(X)
    $y_{c,unsup} \leftarrow unsup\_clf.predict(X)$
    $l1, l2 \leftarrow \emptyset$
    for $i \leftarrow (0 \rightarrow len(X))$ do
        if $y_{c,unsup} == 0$ then
            $l1 \leftarrow l1 \cup i$
        else
            $l2 \leftarrow l2 \cup i$
    $L0 \leftarrow$ RandomChoiseWithoutReplacement($l1, int(len(l1) * \lambda/100)$)    ▷ Getting Indices between 0 and N
    $L1 \leftarrow$ RandomChoiseWithoutReplacement($l2, int(len(l2) * \lambda/100)$)
     return $L0, L1$
```

APPENDIX II

Algorithm 8 Best Cluster

```
    Input: $C0, C1$
    Output: $C\_first, C\_second, \eta_{max}, \vec{s}_{sol}$
procedure BEST_CLUSTER
    $\vec{\eta}_0, \vec{\eta}_1 \leftarrow \vec{0}_{len(C0)}, \vec{0}_{len(C1)}$
    $\eta_{max} \leftarrow 0.0$
    $\eta_{tr}, S_{tr} \leftarrow \emptyset$
    for $i \leftarrow \vec{\eta}_{0,0} .... \vec{\eta}_{0,len(C0)}$ do
        $\eta_{0,i} \leftarrow EVAL(s_i)$
        $\eta_{tr} \leftarrow \eta_{tr} \cup \eta_{0,i}$
        $S_{tr} \leftarrow S_{tr} \cup s_i$
    for $i \leftarrow \vec{\eta}_{1,0} .... \vec{\eta}_{1,len(C1)}$ do
        $\eta_{1,i} \leftarrow EVAL(s_i)$
        $\eta_{tr} \leftarrow \eta_{tr} \cup \eta_{1,i}$
        $S_{tr} \leftarrow S_{tr} \cup s_i$
    $\bar{\eta}^*, \mathbb{V}[\eta*] \leftarrow \mathrm{GPR}(X = S_{tr}, y = \eta_{tr}, var = \sigma^2)$        ▷ See Equations 28a - 28b
    if $\max \bar{\eta}$ in $C0$ then
        $C\_first, C\_second \leftarrow C0, C1$
    else
        $C\_first, C\_second \leftarrow C1, C0$
    $\eta_{max} \leftarrow \max \eta_{tr}$ return $C\_first, C\_second, \eta_{max}, \vec{s}_{sol}$        ▷ $\vec{s}_{sol}$ corresponds to design with maximum efficiency.
```

ACKNOWLEDGEMENTS

This project has been possible thanks to the help and encouragement of many supporters spanning multiple departments within TU Delft and multiple organizations outside TU Delft. This research is the precursor to ECoProp - an NWO funded Joint Industrial Project. Thus, this project was financed by 3mE, TU Delft and supported in kind by LR, TU Delft and MARIN. For this support I am very grateful. I would also like to thank the LR StartUp program which paved a way to convert my Master Thesis into a fully funded PhD program at the faculty of LR and 3mE, TU Delft.

I am very grateful to ir. Klaas Visser, Dr. ir. Stefan Hickel and Dr. ir. Tom van Terwisga for being the first to support my ideas and encourage me to pursue them. From them, I have learnt how to conduct scientific research and make relevant contributions with immediate societal and environmental impact. In addition to research, they also actively involved me in building consensus and consortium for ECoProp. This combined research and entrepreneurial experience has proven immensely valuable and has prepared me well for a future in the industry and academia. Many thanks to the three of them.

I am very thankful to Dineke, Gracia, Anouk, Monique, Patty, Pauline at 3mE, Colette at LR, Zane and Ingrid at MARIN for their full support in the complicated logistics of my PhD. I am specially grateful to Dineke and Anouk for a warm and supportive start to my PhD program. This was very encouraging and kept me motivated through out.

Over the past four years, several supporters have helped shape ideas, provided very valuable feedback and constructive criticism. In an interdisciplinary project, this has brought focus and ensured relevant applications. At LR, TU Delft, I am thankful to Matteo Pini and Richard Dwight for taking the time to provide valuable feedback on the research related to optimization. At MARIN, I am thankful to Evert-Jan Foeth, Thomas Lloyd, Johan Bosschers and Thomas Scholz for their constructive criticism for both CFD and Optimization related research. At 3mE, TU Delft, I am thankful to Matthijs Langelaar for interesting conversations regarding optimization.

I am also very grateful to fellow PhD students - Alina, Carmen, Faisal, Harsh, Marc, Lindert and Loude at 3mE for making the four years memorable. Furthermore, I've had the chance to work and exchange ideas with Gem, Holly, Lina, Maarten, Sören, Swaraj and Themis. This has been both fruitful and fun!

To my dear friends - Alina, Himanshu, Koushik, Mayon, Shwetha, Thej, Vidya and Yaejin, many thanks for being a part of this journey and keeping my spirits high. I am also eternally grateful to my partner for being a dear friend and pillar of support and encouragement through out my PhD.

I hail from a relatively big family which has been instrumental in my upbringing. I'd like to extend my deepest gratitude to my parents, grandparents and siblings who've been extremely supportive and loving. I also take this opportunity to thank my mentors who have helped me greatly in these critical formative years.

Thanks All!

ABOUT AUTHOR

Pranav Doijode Sumanth was borne in Bangalore India, and grew up amides a large, loving and supportive family with many cousins. He obtained is bachelor's degree in 2014 from Bangalore Institute of Technology in Mechanical Engineering before coming to the Netherlands to continue his studies at the TU Delft in the MSc Aerospace, Flight Performance and Propulsion. During his masters, he had an idea to make more efficient engines. Seeing the potential, LR and 3mE Tu Delft offered him a PhD position that resulted in the book at hand. Aside from his research interest in design methods of composite propellers his professional interests lie in pushing the bounds of computing to help make faster and more efficient chips.

Pranav has an aura of enthusiasm and optimism with a *"Can Do!"* attitude and natural ease to network. His way to communicate manages to break down problems into its core features and convey them clearly, making the most complex problems appear straight forward to tackle.

In his free time Pranav is engaged in many side projects. In one project he researches equities to find investment opportunities and has a few subscribers already! In another project he supports Start-Ups with ideas to make best use of AI/ML. Some of his projects have proven their excellence by winning start up awards. He also strives to improve himself, not only by taking courses to develop his knowledge base but also by setting himself goals for physical achievements. He trained as part of the Spartan workout works to continuously beat his 5 min/km running time. At time he made me think that he has the keys to an alternative universe where a day has more than 24 h.

Pranav soaks in knowledge like a sponge, making him a great conversation partner and source of inspiration for a coffee break or when trying out new wines on a relaxed evening. I am certain that Pranav has a steep career ahead of him and am looking forward to seeing where life leads him.

-Alina Colling